European Commission

HANDBOOK
SHE FIGURES
2024

Research and Innovation

She Figures 2024 - Handbook

European Commission
Directorate-General for Research and Innovation
Directorate D — People: Health & Society
Unit D4 — Democracy, Equality & Culture

Contact: Kateřina Svíčková, Head of the Gender Sector D4.001
Oriane Gilloz, Policy Officer, Gender Sector D4.001
Email RTD-GENDERINRESEARCH@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in February 2025
1st edition

PDF ISBN 978-92-68-18634-3 doi:10.2777/218571 KI-09-24-567-EN-N

Luxembourg: Publications Office of the European Union, 2025

EUROPEAN COMMISSION

She Figures 2024

Handbook

Directorate-General for Research and Innovation
Horizon Europe

2025

EN

CONTENTS

1. Tables

2. Acronyms

ASJC	All Science Journal Classification
BES	Business Enterprise Sector
CAGR	Compound Annual Growth Rate
CI	Confidence Interval
DI	Dissimilarity Index
DG	Directorate-General
EFTA	European Free Trade Association
EGGE	Expert Group on Gender and Employment
EIGE	European Institute for Gender Equality
ENP	European Neighbourhood Policy
EPO	European Patent Office
ERA	European Research Area
ETER	European Tertiary Education Register
EU-LFS	Eurostat Labour Force Survey
EU MS	European Union Member States
FORD	Fields of Research and Development
FTE	Full-Time Equivalent
FWCI	Field-Weighted Citation Impact
GCI	Glass Ceiling Index
GOV	Government sector
HC	Head Count
HES	Higher Education Sector
HQP	Highly Qualified Personnel
HRST	Human Resources in Science and Technology
ICT	Information and Communication Technology
ILO	International Labour Organization
ILOSTAT	International Labour Organization Statistics Database
IPC	International Patent Classification (by WIPO)
ISCED	International Standard Classification of Education

ISCED-F	ISCED – Fields of Education and Training
ISCO	International Standard Classification of Occupations
JPO	Japan Patent Office
KIA	Knowledge-Intensive Activities
KIABI	Knowledge-Intensive Activities – Business Industries
MORE	Mobility and Career Paths of Researchers in Europe
NACE	Nomenclature générale des activités économiques dans la communauté européenne (Statistical Classification of economic activities in the European Community)
NPIs	Non-Profit Institutions
NPISH	Non-profit Institutions Serving Households
OECD	Organisation for Economic Co-operation and Development
PATSTAT	EPO Worldwide Patent Statistical Database
PhD	Doctor of Philosophy
PNP	Private Non-Profit
PPS	Purchasing Power Standards
R&D	Research and Development
R&I	Research and Innovation
RFOs	Research Funding Organisations
RPOs	Research Performing Organisations
S&E	Scientists and Engineers
S&T	Science and Technology
GDRIC	Gender Dimension in Research and Innovation Content
SNA	System of National Accounts
STEM	Science, Technology, Engineering and Mathematics
UNESCO	United Nations Educational, Scientific and Cultural Organisation
USPTO	United States Patent and Trademark Office
WIPO	World Intellectual Property Organization
WiS	Women in Science

3. Country codes

Codes for aggregated data

EU-27	European Union – 27 countries (from 2020)

European Union Member States

BE	Belgium
BG	Bulgaria
CZ	Czechia
DK	Denmark
DE	Germany
EE	Estonia
IE	Ireland
EL	Greece
ES	Spain
FR	France
HR	Croatia
IT	Italy
CY	Cyprus
LV	Latvia
LT	Lithuania
LU	Luxembourg
HU	Hungary
MT	Malta
NL	Netherlands
AT	Austria
PL	Poland
PT	Portugal
RO	Romania
SI	Slovenia
SK	Slovakia

FI	Finland
SE	Sweden

European Free Trade Association Countries

IS	Iceland
NO	Norway

Candidate countries

BA	Bosnia and Herzegovina
ME	Montenegro
MD	Moldova
MK	North Macedonia
GE	Georgia
AL	Albania
RS	Serbia
TR	Türkiye
UA	Ukraine

Potential Candidate

XK	Kosovo ([1])

Other countries

AM	Armenia
AR	Argentina
AU	Australia
BR	Brazil
CA	Canada
CN_X_HK	China (except Hong Kong)
FO	Faroe Islands

([1]) This designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

HK	Hong Kong
IL	Israel
IN	India
JP	Japan
KR	South Korea
MX	Mexico
TN	Tunisia
RU	Russia
UK	United Kingdom
US	United States
ZA	South Africa

4. General Introduction

This handbook of indicators has been developed to accompany the She Figures publication. It contains methodological guidance on the collection of data and the calculation of all indicators in the 2024 edition of She Figures. In doing so, it provides further guidelines and recommendations for the collection, processing and use of data on gender equality in research, innovation and science, with the potential to inform organisations at both the national and European level.

She Figures

She Figures provides pan-European, comparable statistics on the state of gender equality in research and innovation (R&I). It covers a wide range of themes, including the gender balance amongst Doctor of Philosophy (PhD) students and academic staff, the relative working conditions of women and men researchers and the steps taken by research institutions to promote gender equality internally. Released every three years since 2003, the report provides a crucial evidence base for policies in this area. It is produced in close collaboration with Member States, Associated Countries and Eurostat. It is recommended reading for policymakers, researchers and anybody with a general interest in these issues.

A large portion of the She Figures publication is dedicated to reporting on a core set of well-established indicators, which serve as the foundation for exposing persistent gender inequalities in the fields of R&I. In addition, each She Figures publication builds on previous versions by introducing new indicators which aim to bring additional and critical gender-based issues to the forefront of the science and technology debate.

This handbook serves as a resource detailing the relevant guidelines for the collection of data pertaining to all indicators of the She Figures 2024 main publication.

Upon future developments and new editions of the She Figures indicators, the handbook will be revised accordingly. As such, it is designed to reflect the state of the art in the mapping and monitoring of gender equality in science and research.

4.1. Aim and scope

Aim

This handbook aims to provide specific guidelines and recommendations concerning the necessary data and indicators for monitoring progress towards gender equality in science, research and innovation.

In particular, the handbook promotes cross-country uniformity in terms of data collection, indicator computation and data-validation procedures. Furthermore, it provides interested stakeholders with detailed information on the data needed to examine gender equality in R&I as well as on the importance given to the gender dimension in R&I content. It serves as a reference document and provides users with the methods needed to calculate the indicators so as to increase the quality and consistency of gender-related indicators across countries and time periods.

Scope

The handbook is not intended to be specific to any version of the She Figures publications. Rather, it is intended to be used as the basis for the computation of indicators in current and future versions of She Figures and related publications.

Current version of the handbook

Although intended to act as a stand-alone document (i.e., untied to any of the specific versions of the She Figures publications), the current version of the handbook was created to accompany the 2024 edition of the publication and thus includes some details specific to that edition. In the 2024 version of She Figures, data are presented at the individual country level as well as the broader EU level for the current 27 EU Member States, the countries associated with Horizon Europe (Albania, Armenia, Bosnia and Herzegovina, Faroe Islands, Georgia, Iceland, Israel, Kosovo, Moldova, Montenegro, North Macedonia, Norway, Serbia, Tunisia,

Türkiye, United Kingdom and Ukraine) ([2]) and G20 countries (Argentina, Australia, Brazil, Canada, China (except Hong Kong), India, Indonesia, Japan, Mexico, Republic of Korea, Russia, Saudi Arabia, South Africa, and United States).

The handbook has been thoroughly cross-referenced and contains an indexed list of indicators aimed at improving accessibility and readability (see Annex 2).

4.2. History and background of She Figures

History

In 1999, the Council of the EU recognised that women were under-represented in the fields of scientific and technical research, describing this as a 'common concern' at the national and European level ([3]). At this time, there were virtually no pan-European statistics on what happened to women after they left university, despite concerns that after graduating from their degrees, 'women frequently encounter[ed] obstacles in their career[s]', which contributed to their under-representation in scientific posts ([4]).

Subsequently, the EU recognised the need for harmonised sex-disaggregated data on women in science and research if governments were to develop effective policies in this area ([5]). Meeting in 1999, the Helsinki Group on Women and Science appointed a sub-group of Statistical Correspondents with responsibility for collecting national data and feeding into the creation of European statistics on these topics.

The end result of this process was She Figures, first released in 2003 and updated every three years ever since. By presenting statistical indicators on a wide range of topics, the report enables readers to develop a comprehensive understanding of the state of gender equality in science, research and innovation.

Changes to She Figures over time

Primarily, the She Figures publication serves as a tool for measuring the impact and effectiveness of gender equality policies in science, research and innovation. The majority of indicators in She Figures present and explore the following themes:

- The presence of women in research across different sectors.
- Horizontal segregation by sex across different fields of study and occupations (in Research and Development (R&D) roles).
- Vertical segregation by sex in academia, i.e., the (under-)representation of women in the highest grades/posts of research and as heads of academic institutions.

Each edition also aims to further understanding of these issues by introducing additional indicators that explore new themes. She Figures 2006 developed new indicators to give a more detailed picture of the labour force as a whole and the patterns of employment for women and men researchers across different sectors, such as the business enterprise sector (BES). The 2009 edition introduced indicators on the gender pay gap and began to break down some data by age group (in addition to sex disaggregation), and the 2012 report added indicators on the mobility of researchers and the proportion of researchers with children.

Similarly, She Figures 2015 included new indicators to match emerging policy priorities. Some provide further insight into the working conditions of researchers and the proportion of Research Performing Organisations

([2]) New Zealand, although associated to Pillar II 'Global Challenges and European Industrial Competitiveness' as from the Work Programmes 2023 onwards, including for the institutionalised European partnerships, was not included in She Figures 2024 because its association took effect when the data for the publication had already been compiled.

([3]) European Council. *Resolution of 20 May 1999 on Women and Science (1999/C 201/01)*. Official Journal of the European Communities, 1999. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31999Y0716(01)

([4]) DG Research, *Women in science and technology: Creating sustainable careers*, Office for Official Publications of the European Communities, Luxembourg, 2009.

([5]) For example, see European Commission, 'Women and science: mobilising women to enrich European research', *COM(1999)76*, 1999. European Parliament, *Resolution on Commission's Communication COM(1999)76 (PE 284.656)*, 1999.

(RPOs) that have adopted Gender Equality Plans, while others measure the relative contribution of women and men to published research and inventorships and the degree to which researchers integrate a gender dimension into their research publications. The 2018 edition, among other things, developed indicators that measure the success of women and men in graduating from tertiary education, as well as the research output and citation impact of women and men as authors and co-authors of scientific publications.

She Figures 2021 included 12 new indicators. One of them measures the proportion of self-employed women within Information and Communication Technologies (ICT) and Science and Engineering (S&E) professionals. Two of the newly introduced indicators were based on the already existing indicator that measures the proportion of women and men who work under precarious contracts in the higher education sector (HES), disaggregating it further by family status and career stage respectively. A fourth indicator provides insight in institutional change in RPOs, more specifically in the promotion of gender equality by implementing measures and actions to support it. The other eight new indicators were all bibliometric: two of them measure the ratio of women to men among authors of scientific publications (for active, and for all, authors, respectively), by level of seniority. A third measures the ratio of the average number of publications by women to that by men, by field of R&D and level of seniority. A fourth measures the ratio of the average Field-Weighted Citation Impact (FWCI) of publications authored by women to that of publications by men, and a fifth indicators measures the average proportion of women among authors by field of R&D. Another new bibliometric indicator measures the average proportion of women among authors on publications that resulted from international collaboration, by field of R&D. Finally, two other new indicators assess the gender dimension or research content and the intersectional aspects of publications in Horizon 2020 projects.

She Figures 2024 includes nine new indicators. These encompass:

- An indicator on research outputs that contribute to SDG 5 (“Achieve gender equality and empower all women and girls”).
- An indicator relating to the gender composition of research teams, to provide insight on gender-based network homophily among researchers.
- Intersectional indicators that explore intersections between gender, disability and migration status, by disaggregating existing She Figures indicators.
- Indicators that replace indicators previously based on the (now discontinued) MORE Survey. The MORE Survey was used in previous editions of She Figures for indicators on the working conditions of researchers, and specifically to explore international mobility, part-time employment and precarious employment. The indicators that aim to replace those formerly based on the MORE Survey focus on international mobility of PhD graduates and authors of publications, and part-time and temporary employment in the Higher Education Sector.

The team had also intended to incorporate an indicator based on the EU-Gender Based Violence survey by Eurostat, EIGE and FRA, however this has not been possible to date due to data availability.

Another important addition to She Figures 2024 is the introduction of the She Figures Index, a composite indicator drawn from representative indicators from the She Figures study. This new index serves as a comprehensive tool to assess gender equality within the European Research Area (ERA) across EU Member States and offers insights into the current state of gender equality in R&I across Europe, tracking progress over time. For a detailed explanation of the methodology behind the She Figures Index, please refer to section 5.12 with comprehensive information available in Annex 4.

Data in She Figures

Most of the She Figures indicators originate from Eurostat (the Statistical Office of the EU), which provides sex-disaggregated data on education, research and development, professional earnings and scientific employment. The Statistical Correspondents enrich this picture by collecting primary data (broken down by sex) on senior academic staff, the heads of universities, funding applicants and beneficiaries and the membership of scientific and advisory boards ([6]). Expansion of She Figures since 2003 has resulted in the use of other sources (e.g., the MORE Survey on the Mobility of Researchers (She Figures 2012 – 2021), the Scopus™ database) and the use of web-scraping techniques.

([6]) This primary data makes up the Women in Science (WiS) database.

4.3. Structure of the handbook

The She Figures Handbook is made up of three sections and four annexes:

- The first (current) section provides a brief overview of the aim and scope of the handbook, as well as background to She Figures.
- The second section describes all indicators used in the She Figures publication, including definitions, rationale as well as computation methods (with the necessary data, data source, formulas and any calculation specifications or comments that may be of relevance).
- The third section details the general quality plan of the She Figures publication, focusing on the methodological principles employed in the verification and validation of data.
- There are four annexes. The first one provides an overview of how key terms are defined. The second one provides an index of the indicators. The third one provides the correspondence of All Science Journal Classification (ASJC) categories with Fields of Research and Development. The last one provides the detailed methodology for the She Figures Index, expanding on the summary provided in 5.12. 5.12

The sections and annexes are followed by the bibliography.

5. Indicators

The indicators presented in this handbook have been selected from a variety of different sources on the basis that they provide important information on gender inequalities in the field of research and innovation (R&I). The development of each new version of She Figures includes several in-depth consultations with key stakeholders to determine how the landscape of data on gender equalities has changed since the previous version of the publication was created and whether the inclusion of new indicators is merited. During this process, indicators from previous versions of the publication are also reassessed to determine whether they are still relevant and to ensure that they adhere to ever increasing quality and coverage standards.

The data required to compute the majority of indicators are drawn from Eurostat databases or from the Women in Science database of data collected by the Statistical Correspondents. Other data sources have been used to develop new indicators in recent years. For example, in the last two editions of She Figures (2018 and 2021), the Scopus™ database has been used to produce bibliometric indicators by sex (e.g., ratio of women to men authorship of scientific papers, proportion of a country's scientific production including a gender dimension), and the EPO Worldwide Patent Statistical Database (PATSTAT) has been used to produce technometric indicators (e.g., proportion of women among inventors). Moreover, web-scraping techniques were used to measure indicators on institutional change for the 2021 and 2024 editions.

In addition, OECD and the International Labour Organization (ILO) were used as supplementary data source for countries not covered by Eurostat. The following Sections (5.1 to 5.12) present the She Figures indicators by data source and subject group. Each section is introduced by a general rationale for the selection of each group of indicators – based on a content perspective – as well as a broad description of the source.

5.1. Eurostat – Education statistics

Content-based rationale

Indicators computed from Eurostat education statistics aim to investigate the level of progress and the persistent barriers that exist for women in the pursuit of postgraduate education, as well as the differences in subject choice and fields of study by gender, particularly in regard to natural science and engineering, within the context of persistent gender stereotypes and the EU's policy agenda. Indicators falling into this category include the proportion of women among ISCED 8 (Doctoral or equivalent) graduates by country, the compound annual growth rate of ISCED 8 graduates by sex, the proportion of women among ISCED 8 graduates by field of study, the distribution of ISCED 8 graduates across the broad fields of study by sex, the ratio of ISCED 6 (Bachelor's or equivalent) graduates to ISCED 6 entrants, the ratio of ISCED 8 entrants to ISCED 7 (Master's or equivalent) graduates, the ratio of ISCED 8 graduates to ISCED 8 entrants, by sex and broad field of study, and the differences in the international credit mobility of women and men ISCED 8 graduates.

Broad overview of the source

These data can be accessed through the Education and Training Statistics database on the Eurostat website (http://ec.europa.eu/eurostat/web/education-and-training/database) and OECD website (http://stats.oecd.org) for AU, BR, CA, JP, MX, US, and IL. Data for G20 countries can be accessed through OECD, depending on the indicator and the data availability. The data are concerned with participation in education and training (including adult learning), learning mobility, education personnel, education finance, education and training outcomes, language learning and self-reported language skills ([7]). Data are collected on an annual basis, based on the academic year (i.e., 2012 refers to the academic year 2011/12) ([8]). 'Education, vocational training and lifelong learning' are ongoing goals towards a sustainable Europe by 2030 ([9]).

Statistics that provide such information are publicly available, regularly updated and accompanied by extensive methodological notes.

The classification of education levels is based on the International Standard Classification of Education (ISCED). This classification was revised in 2011 and data shown in this publication follow the updated version except when it is referred otherwise.

Cut-off date

The cut-off date for data collection for Education statistics downloaded from Eurostat's dissemination database (Eurostat) and OECD's database was 31 July 2023.

([7]) Eurostat, *Statistics Explained: Education statistics*. 2023, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Education_and_training

([8]) Eurostat, Education and Training Reference Metadata, https://ec.europa.eu/eurostat/cache/metadata/en/educ_uoe_enr_esms.htm#ref_period1689000914030

([9]) European Commission. *Towards a sustainable Europe by 2030 (Reflection paper)*. Brussels, 2019, https://commission.europa.eu/publications/sustainable-europe-2030_en

5.1.1. Proportion of women and men among tertiary education graduates

Definition of indicator

This indicator presents the proportion of women and men graduates in tertiary education to the total graduates in tertiary, broken down by country.

Rationale

In line with its ambition to encourage more 'research-intensive' economies, the European Commission had called for more doctoral candidates and argued that efforts must be made to 'tackle stereotyping and dismantling the barriers still faced by women in reaching the highest levels in post-graduate education and research' ([10]). In recent decades, women in the EU have made significant developments in raising their level of education. According to She Figures 2021, the presence of women among doctoral graduates has increased at the EU level and at country level. This indicator sheds light on the level of progress in increasing women's representation in the top levels of education and research, considering their success in ultimately graduating from doctoral degrees, as opposed to their entry as candidates.

Computation method

Data needed

(F_i) Number of women graduates at the i-th education level. **Unit: Number**.

(M_i) Number of men graduates at the i-th education level. **Unit: Number**.

(T_i) Number of total graduates at the i-th education level. **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by age*).

Computation formula

Proportion of women among graduates at the i-th education level = F_i/T_i

Proportion of men among graduates at the i-th education level = M_i/T_i

where:

i denotes the education level according to the International Standard Classification of Education (ISCED-2011).

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level.

- ISCED level 6 corresponds to studies at Bachelor's or equivalent level.
- ISCED level 7 corresponds to studies at Master's or equivalent level.
- ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level.

([10]) European Commission, *Supporting growth and jobs – an agenda for the modernisation of Europe's higher education systems*, COM(2011)567 final, 2011, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A52011DC0567

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

Comments and critical issues

Proportion of women and men among total graduates at ISCED levels 6 and 7 are calculated only at the aggregated EU level, not at country level.

5.1.2. Proportion of women and men among tertiary education students

Definition of indicator

This indicator presents the proportion of women and men among students in tertiary education.

Rationale

Tertiary education, provided by universities and other higher education institutions, seems to play an essential role in society, by fostering innovation, increasing economic development and growth and improving more generally the well-being of citizens. In the EU-27 there were 18.5 million tertiary education students in 2021, with women accounting for 54.2% of them ([11]).

According to She Figures 2021, in 2018, women accounted for 54% of ISCED 6 and 7 students but their presence drops to 48% among ISCED 8 students. This indicator sheds light on the level of progress in increasing women's representation in the top levels of education and research, considering their success in ultimately graduating from doctoral degrees, as opposed to their entry as candidates.

Computation method

Data needed

(F_i) Number of women students at the i-th education level. **Unit: Number**.

(M_i) Number of men students at the i-th education level. **Unit: Number**.

(T_i) Number of total students at the i-th education level. **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_enrt03)*

Computation formula

Proportion of women among students at the i-th education level = F_i/T_i

Proportion of men among students at the i-th education level = M_i/T_i

where:

i denotes the education level according to the International Standard Classification of Education (ISCED-2011).

([11]) Eurostat, *Statistics Explained: Tertiary education statistics*, 2023, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level.

- ISCED level 6 corresponds to studies at Bachelor's or equivalent level.
- ISCED level 7 corresponds to studies at Master's or equivalent level.
- ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level.

The number of students refers to students enrolled in the reference year. It includes all students enrolled in the country, i.e., non-nationals too, but does not include nationals enrolled abroad.

Comments and critical issues

Proportion of women and men among total students are calculated only at the aggregated EU level, not at country level.

5.1.3. Compound annual growth rate (CAGR) of Doctoral graduates, by sex

Definition of indicator

This indicator presents the compound annual growth rate (CAGR) of graduates by sex, meaning the average percentage growth each year for women and men graduates in a given period for graduates at ISCED 8 level.

Rationale

In 2012, the European Commission warned that 'while the proportion of women at the first two levels of tertiary education is higher than that of men, the proportion of women at PhD level is lower' ([12]). The situation is still similar over a decade later, as the share of women among those studying bachelor's and master's degrees are higher than the corresponding share of men in the EU-27 in 2021. However, for people following doctoral studies, the majority of students were men ([13]). This indicator demonstrates the level of progress over time in increasing women's presence amongst those taking doctoral degrees.

Computation method

Data needed

(F) Number of women ISCED 8 graduates in a start and an end year. **Unit: Number**.

(M) Number of men ISCED 8 graduates in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

([12]) DG Research and Innovation, *Researchers' Report 2012*, 2012, https://www.euraxess.tn/sites/default/files/policy_library/121003_the_researchers_report_2012_final_report.pdf

([13]) Eurostat, *Statistics Explained: Tertiary education statistics*, 2023, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02*)

OECD (*http://stats.oecd.org; Graduates by age*).

Computation formula

The CAGR shows the yearly average rate of growth for a given period. For women and men graduates, it is respectively computed as follows:

CAGR for women graduates = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR for men graduates = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period (i.e., e – s);

F_s denotes the number of women graduates in the start year;

F_e denotes the number of women graduates in the end year;

M_s denotes the number of men graduates in the start year;

M_e denotes the number of men graduates in the end year.

For example, if there were 100 women graduates in 2012 and 150 in 2016, the calculation would be:

CAGR for women graduates = $100 \times \left[(150/100)^{1/4} - 1\right]\% = 10.7\%$

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level.

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

5.1.4. Proportion of women among Doctoral graduates, by broad field of study

Definition of indicator

This indicator presents the proportion of women ISCED 8 graduates, within various fields of study.

Rationale

Although there is some disagreement amongst experts ([14]), it is generally accepted that differences in women and men's educational pathways may have some impact on the occupations they pursue at a later stage. More specifically, the EU Council stresses that gender segregation in education leads to inequality in terms of pay, pensions, lifelong earnings, working conditions and the working environment, reinforcing gender stereotypes ([15]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Challenging and dissolving gender stereotypes, especially those that constrain the choices of boys and girls for their field of study' ([16]). Similarly, the new ERA will strengthen focus on the participation of women in Science, Technology, Engineering and Mathematics (STEM) fields ([17]). By breaking down PhD graduates into different fields of study, this indicator enables more in-depth analysis of the extent of gender difference in subject choice.

Computation method

Data needed

(T_b) Number of ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(F_b) Number of women ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(b) Denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org*; Graduates by field).

Computation formula

Proportion of women among students or graduates = F_b / T_b

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level.

([14]) The debate relates to the level and nature of the impact of educational segregation on later segregation in the labour market. For an overview of the debate, consider EIGE (2009), pp. 42–45.

([15]) Council of the European Union, *Council conclusions on enhanced measures to reduce horizontal gender segregation in education and employment*, 2017, https://data.consilium.europa.eu/doc/document/ST-15468-2017-INIT/en/pdf

([16]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025*, COM(2020) 625 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0625

([17]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation*, COM(2020)628 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A628%3AFIN

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e. non-nationals too, but does not include nationals graduating abroad.

The broad fields of study according to the ISCED – Fields of Education and Training (ISCED-F) classification are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services

5.1.5. Distribution of Doctoral graduates across broad fields of study, by sex

Definition of indicator

This indicator presents the distribution of ISCED 8 graduates by sex and broad field of study.

Rationale

As mentioned above, experts generally consider that the differences in women and men's educational pathways may have an impact on the occupations they pursue at a later stage ([18]). This association between education and employment is a core part of the EU policy agenda. For instance, the European Commission promotes gender inclusive STEM education and communication and aims at strengthening the focus of the EU, through the European Education Area ([19]) and the new ERA ([20]), on the participation of women in STEM fields. It, encourages more girls to take science subjects with a view to considering a career in this area (see, for instance, the Hypatia project funded by the European Union's Horizon 2020 Framework Programme for Research and Innovation) and nurtures future women innovators (see, for instance, the shemakes.eu project funded under the Science with and for Society Work Programme, SwafS-26-2020: Innovators of the future: bridging the gender gap).

([18]) There was, however, debate about the level and nature of this impact. For an overview of the debate, consider EGGE (2009), pp.42–45.

([19]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025*, COM(2020) 625 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0625

([20]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation*, COM(2020) 628 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A628%3AFIN

This indicator gives a picture of the overarching differences in women's and men's fields of study at ISCED 8 level. It is slightly different from the indicator 'Proportion of women ISCED 8 graduates by broad field of study' in that it breaks down the fields of study for women ISCED 8 graduates and men ISCED 8 graduates respectively.

Computation method

Data needed

(F) Number of women ISCED 8 graduates (all broad fields of study). **Unit: Number**.

(M) Number of men ISCED 8 graduates (all broad fields of study). **Unit: Number**.

(F_b) Number of women ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(M_b) Number of men ISCED 8 graduates in broad field of study *b*. **Unit: Number**.

(b) Denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

For each sex, this indicator presents the proportion of graduates in each broad field of study, in order to show how women/men graduates (at ISCED 8) are spread out across different subjects.

For each broad field of study, the formula for this indicator is:

Distribution of women graduates across fields of study = $\left(\frac{F_b}{F}\right)$

Distribution of men graduates across fields of study = $\left(\frac{M_b}{M}\right)$

The proportions for each field are shown alongside each other, with a sum total of 100 % for each sex.

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent.

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services

5.1.6. Proportion of women among Doctoral graduates by narrow field of study in natural sciences, ICT and engineering

Definition of indicator

This indicator presents the proportion of women ISCED 8 graduates within the eight subfields of natural sciences, ICT and engineering, falling under the broad fields 'natural sciences, mathematics and statistics', 'Information and Communication Technologies' and 'Engineering, manufacturing and construction'.

Rationale

The EU recognises the existence of horizontal segregation, whereby women and men at the same level of education or employment are concentrated in different fields (full definition available in Annex 2). For example, according to the Gendered Innovations project (in which the European Commission is a partner), 'in both the United States and European Union, women are slightly underrepresented with respect to overall doctoral (ISCED 8) degrees, but substantially underrepresented with respect to S&E doctorates' (Stanford University, 'Disparities between women and men'). The European Education Area ([21]) aims at dismantling the gender stereotypes that affect field choices by boys and girls.

This indicator allows one to measure such segregation at ISCED 8 level, by presenting the proportion of women graduates in certain subfields. By breaking down the graduates by subfield, one can assess variations within broader fields of study.

Computation method

Data needed

(F_n) Number of women ISCED 8 graduates in each narrow field of study *n* in natural sciences, ICT and engineering. **Unit: Number**.

(T_n) Total number of ISCED 8 graduates in each narrow field of study *n* in natural sciences, ICT and engineering. **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

([21]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025*, COM(2020) 625 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0625

Computation formula

Proportion of women among graduates in each narrow field = F_n/T_n

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level.

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

The narrow fields of study in natural sciences and engineering according to the ISCED-F classification of fields of education and training are the following:

- 051 Biological and related sciences
- 052 Environment
- 053 Physical sciences
- 054 Mathematics and statistics
- 061 Information and Communication Technologies
- 071 Engineering and engineering trades
- 072 Manufacturing and processing
- 073 Architecture and construction

5.1.7. Compound annual growth rate (CAGR) of Doctoral graduates by narrow field of study in natural sciences, ICT and engineering, by sex

Definition of indicator

This indicator presents the compound annual growth rate of the number of men and women ISCED 8 graduates within the eight subfields of natural sciences, ICT and engineering, falling under the broad fields 'Natural sciences, mathematics and statistics', 'Information and Communication Technologies' and 'Engineering, manufacturing and construction'.

Rationale

The EU recognises the persistent differences in the educational choices of women and men. In 2020, the European Commission mentions that in the European Education Area, the education and training systems should consider 'Developing a better gender sensitivity ([22]) in education processes and institutions […], challenging and dissolving gender stereotypes, especially those that constrain the choices of boys and girls for their field of study […] and working towards a proper gender balance in leadership positions' ([23]).

This indicator allows one to gauge the extent of such segregation at ISCED 8 level, by calculating the changes in women and men's representation over time. By breaking down the graduations by subfield, one can assess

([22]) According to the European Institute for Gender Equality (EIGE), gender sensitive policies and programmes are those that take into account the particularities pertaining to the lives of both women and men, while aiming to eliminate inequalities and promote gender equality, including an equal distribution of resources, therefore addressing and taking into account the gender dimension.

([23]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025*, COM(2020) 625 final, 2020b.

variations within broader fields of study. The results of this indicator can be compared with those of the similar indicator showing the proportion of women ISCED 8 graduates by narrow field of study in natural sciences, ICT and engineering.

Computation method

Data needed

$(\mathrm{F_n})$ Number of women ISCED 8 graduates in each narrow field of study *n* in a start and an end year. **Unit: Number**.

$(\mathrm{M_n})$ Number of men ISCED 8 graduates in each narrow field of study *n* in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

The compound annual growth rate (CAGR) shows the average rate of growth per year for a given period. In this indicator, it shows the average percentage growth of women and men ISCED 8 graduates in narrow fields of natural science and engineering.

It is respectively computed as follows:

CAGR for women graduates in a narrow field = $\left(\mathrm{F_{n,e}}/\mathrm{F_{n,s}}\right)^{1/N} - 1$

CAGR for men graduates in a narrow field = $\left(\mathrm{M_{n,e}}/\mathrm{M_{n,s}}\right)^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{F_{n,s}}$ denotes the number of ISCED 8 women graduates in narrow field *n* in the start year *s*;

$\mathrm{F_{n,e}}$ denotes the number of ISCED 8 women graduates in narrow field *n* in the end year *e*;

$\mathrm{M_{n,s}}$ denotes the number of ISCED 8 men graduates in narrow field *n* in the start year *s*;

$\mathrm{M_{n,e}}$ denotes the number of ISCED 8 men graduates in narrow field *n* in the end year *e*.

For example, if there were 100 women PhD graduates from physical sciences in 2012 and 150 in 2016, the calculation would be:

CAGR for women PhD graduates in physical sciences =$100 \times \left[(150/100)^{1/4} - 1\right]\% = 10.7\%.$

Specifications

The International Standard Classification of Education (ISCED-2011) categorises education programmes by level. ISCED 8 corresponds to studies at Doctoral (PhD) or equivalent level.

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

The narrow fields of study in natural sciences and engineering according to the ISCED-F classification of fields of education and training are the following:

- 051 Biological and related sciences
- 052 Environment
- 053 Physical sciences
- 054 Mathematics and statistics
- 061 Information and Communication Technologies
- 071 Engineering and engineering trades
- 072 Manufacturing and processing
- 073 Architecture and construction

5.1.8. Ratio of Bachelor graduates to Bachelor entrants, by sex and broad field of study

Definition of indicator

This indicator is the ratio of ISCED 6 graduates to ISCED 6 entrants, broken down by sex, broad field of study and country.

Rationale

The segregation between women and men scientists is already connected to early segregation in education pathways chosen by young women and men. This indicator shows the level of progress in increasing women's representation in the higher levels of education and research, considering their success in graduation at ISCED level 6.

Computation method

Data needed

$(F_{g,b})$ Number of women ISCED 6 graduates in broad field of studies *b*. **Unit: Number**.

$(F_{e,b})$ Number of women ISCED 6 entrants in broad field of studies *b*. **Unit: Number**.

$(M_{g,b})$ Number of men ISCED 6 graduates in broad field of studies *b*. **Unit: Number**.

$(M_{e,b})$ Number of men ISCED 6 entrants in broad field of studies *b*. **Unit: Number**.

Source of data

For entrants:

Eurostat – Education and Training Statistics (online data code: *educ_uoe_ent02*)

OECD (*http://stats.oecd.org*; New entrants by field).

For graduates:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org*; Graduates by field).

Computation formula

Ratio of women ISCED 6 graduates to ISCED 6 entrants for a given field of study = $(F_{g,b}) \Big/ (F_{e,b})$

Ratio of men ISCED 6 graduates to ISCED 6 entrants for a given broad field of study = $(M_{g,b}) \Big/ (M_{e,b})$

where:

b denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Specifications

ISCED level 6 corresponds to studies at Bachelor's or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

New entrants to a level of education are students who, during the course of the reference school or academic year, enter for the first time any programme in a given level of education, irrespective of whether they enter the programme at its beginning or at an advanced stage of it.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services

Comments and critical issues

The indicator compares the same reference year's entrants and graduates. These are two different groups of persons. Ideally one would need to follow up each year's entrants and count those that graduate. The proposed formulation is considered as a proxy.

5.1.9. Ratio of Doctoral entrants to Master graduates, by sex and field of study (broad and narrow)

Definition of indicator

This indicator is the ratio of ISCED 8 entrants to ISCED 7 graduates, broken down by sex, broad field of study and country.

Rationale

The segregation between women and men scientists is already connected to early segregation in education pathways chosen by young women and men. The indicator helps assess the propensity of women and men who graduate from ISCED level 7 to continue to ISCED level 8 studies.

Computation method

Data needed

$(F_{g,b})$ Number of women ISCED 7 graduates in broad field of studies *b*. **Unit: Number**.

$(F_{g,n})$ Number of women ISCED 7 graduates in narrow field of studies *n*. **Unit: Number**.

$(F_{e,b})$ Number of women ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

$(F_{e,n})$ Number of women ISCED 8 entrants in narrow field of studies *n*. **Unit: Number**.

$(M_{g,b})$ Number of men ISCED 7 graduates in broad field of studies *b*. **Unit: Number**.

$(M_{g,n})$ Number of men ISCED 7 graduates in narrow field of studies *n*. **Unit: Number**.

$(M_{e,b})$ Number of men ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

$(M_{e,n})$ Number of men ISCED 8 entrants in narrow field of studies *n*. **Unit: Number**.

Source of data

For entrants:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_ent02)*

OECD (*http://stats.oecd.org*; New entrants by field).

For graduates:

Main data source: Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Ratio of women ISCED 8 entrants to ISCED 7 graduates for a given **broad** field of study

$$= {(F_{e,b})}\Big/{(F_{g,b})}$$

Ratio of women ISCED 8 entrants to ISCED 7 graduates for a given **narrow** field of study

$$= {(F_{e,n})}\Big/{(F_{g,n})}$$

Ratio of men ISCED 8 entrants to ISCED 7 graduates for a given **broad** field of study

$$= {(M_{e,b})}\Big/{(M_{g,b})}$$

Ratio of men ISCED 8 entrants to ISCED 7 graduates for a given **narrow** field of study

$$= {(M_{e,n})}\Big/{(M_{g,n})}$$

where:

b denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields;

n denotes the narrow fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Specifications

ISCED level 7 corresponds to studies at Master's or equivalent level and ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

New entrants to a level of education are students who, during the course of the reference school or academic year, enter for the first time any programme in a given level of education, irrespective of whether they enter the programme at its beginning or at an advanced stage of it.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services

The narrow fields of study in natural sciences and engineering according to the ISCED-F classification of fields of education and training are the following:

051 Biological and related sciences

052 Environment

053 Physical science

054 Mathematics and statistics

061 Information and Communication Technologies

071 Engineering and engineering trades

072 Manufacturing and processing

073 Architecture and construction

Comments and critical issues

The indicator compares the same reference year's ISCED 8 entrants and ISCED 7 graduates. Ideally one would need to follow up each year's ISCED 7 graduates and count those that start ISCED 8 studies at some point in the future. The proposed formulation is considered as a proxy.

5.1.10. Ratio of Doctoral graduates to Doctoral entrants, by sex and broad field of study

Definition of indicator

This indicator is the ratio of ISCED 8 graduates to ISCED 8 entrants, broken down by sex, broad field of study and country.

Rationale

The indicator shows the level of progress in increasing women's representation in the top levels of education and research, considering their success, as well as that of men, in graduation at ISCED level 8.

Computation method

The indicator is calculated from the following data.

Data needed

$(F_{g,b})$ Number of women ISCED 8 graduates in broad field of studies *b*. **Unit: Number**.

$(F_{e,b})$ Number of women ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

$(M_{g,b})$ Number of men ISCED 8 graduates in broad field of studies *b*. **Unit: Number**.

$(M_{e,b})$ Number of men ISCED 8 entrants in broad field of studies *b*. **Unit: Number**.

Source of data

For entrants:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_ent02)*

OECD (*http://stats.oecd.org*; New entrants by field).

For graduates:

Eurostat – Education and Training Statistics *(online data code: educ_uoe_grad02)*

OECD (*http://stats.oecd.org; Graduates by field*).

Computation formula

Ratio of women ISCED 8 graduates to ISCED 8 entrants for a given broad field of study

$$= {(F_{g,b})}\Big/{(F_{e,b})}$$

Ratio of men ISCED 8 graduates to ISCED 8 entrants for a given broad field of study

$$= {(M_{g,b})}\Big/{(M_{e,b})}$$

where:

b denotes the broad fields of study according to the ISCED-F classification of fields of education and training or the total of all fields.

Specifications

ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

New entrants to a level of education are students who, during the course of the reference school or academic year, enter for the first time any programme in a given level of education, irrespective of whether they enter the programme at its beginning or at an advanced stage of it.

The broad fields of study according to the ISCED-F classification of fields of education and training are the following:

00 Generic programmes and qualifications

01 Education

02 Arts and humanities

03 Social sciences, journalism and information

04 Business, administration and law

05 Natural sciences, mathematics and statistics

06 Information and Communication Technologies

07 Engineering, manufacturing and construction

08 Agriculture, forestry, fisheries and veterinary

09 Health and welfare

10 Services

Comments and critical issues

The indicator compares the same reference year's entrants and graduates. These are two different groups of persons. Ideally one would need to follow up each year's entrants and count those that graduate. The proposed formulation is considered as a proxy.

5.1.11. Differences in the international credit mobility of women and men PhD graduates

Definition of indicator

The indicator shows the difference in the percentage of women/ men ISCED 8 graduates who – during their doctorate studies – moved for at least three months to a country other than that where they attained their ISCED 8 degree. Credit mobility includes both study abroad and work placement. Degree mobile graduates are included.

Rationale

One of the main European Commission plans for the new European Research Area is to strengthen the mobility of researchers and the flow of knowledge ([24]). However, there are indications that women may be less mobile than men at certain stages of their life. For example, there are signs that women continue to bear the majority of childbearing and caring for relatives with responsibilities in the EU and – as the European Parliament warns – mobility 'can be difficult to reconcile with family life'. According to the EU funded Gendered Innovations Expert Group report, 'gender roles that limit women's mobility interfere with careers in science and engineering' ([25]).

This indicator aims to identify if there are indeed such differences in the mobility of women and men, focusing on researchers' experiences of mobility during their doctorate study.

Computation method

The indicator is calculated from the following data.

([24]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020c.

([25]) European Commission, *Gendered Innovations 2: How Inclusive Analysis contributes to Research and Innovation*, 2020, https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en

Data needed

(F_i) Number of female ISCED 8 graduates who – during their doctorate studies – moved for at least three months to a country other than that where they attained their ISCED 8 degree. **Unit: Number.**

(F) Number of female ISCED 8 graduates. **Unit: Number**.

(M_i) Number of male ISCED 8 graduates who – during their doctorate studies – moved for at least three months to a country other than that where they attained their ISCED 8 degree. **Unit: Number**.

(M) Number of male ISCED 8 graduates. **Unit: Number**.

Source of data

Eurostat – Credit mobile graduates (at least 3 months abroad) by education level, country of destination, type of mobility and sex (*educ_uoe_mobc02*)

Eurostat – Graduates by education level, programme orientation, sex and field of education (*educ_uoe_grad02*)

Computation formula

Percentage of internationally mobile female ISCED 8 graduates $= {(F_i)}/{(F)}$

Percentage of internationally mobile male ISCED 8 graduates $= {(M_i)}/{(M)}$

Differences between genders in international mobility during PhD $= {(M_i)}/{(M)} - {(F_i)}/{(F)}$

Specifications

ISCED level 8 corresponds to studies at Doctoral (PhD) or equivalent level according to the International Standard Classification of Education (ISCED-2011).

The number of graduates refers to those graduating in the reference year. It includes all persons graduating in the country, i.e., non-nationals too, but does not include nationals graduating abroad.

This indicator focuses on graduates in the higher education sector only. In this indicator, graduates who are 'internationally mobile' are defined as those who have moved abroad for at least three months during their studies to a country other than the one where they completed their degree.

5.2. Eurostat – EU Labour Force Survey

Content-based rationale

Directive 2006/54/EC of 5 July 2006 lays down the principle of equal treatment of men and women in the EU in relation to their working conditions, access to promotion and access to occupational security schemes. Earlier editions of She Figures included indicators on part-time employment of researchers in the higher education sector (HES) and on precarious working contracts of researchers in HES out of total researcher population, by sex, family status and career stage. Those indicators were compiled with data collected in successive surveys on the Mobility and Career Paths of Researchers in Europe (MORE).

The MORE surveys have been discontinued and the EU-LFS has been identified as a source of data for indicators that are proxies of the earlier indicators, albeit without breakdown by family status or career stage.

Broad overview of the source

Statistics on the employment characteristics of the population, compiled from EU-LFS data, can be accessed through the Labour Force Survey (LFS) database on the Eurostat website. The statistics are publicly available, regularly updated and accompanied by extensive methodological notes.

Custom tables can also be produced by Eurostat upon request of interested users, provided that the tables' contents do not breach the confidentiality of personal statistical information of the respondents concerned.

Data are collected on a quarterly basis, although a relatively small subset of the variables are collected only annually or even less frequently. Correspondingly, quarterly and annual statistics are produced. She Figures presents annual statistics. The national statistical institutes across Europe conduct the surveys and are responsible for selecting the sample, preparing the questionnaires, conducting the direct interviews among households, and forwarding the results to Eurostat. At European level, data are made comparable across countries and over time by:

- using the same concepts and definitions
- following the guidelines of the International Labour Organisation
- using common classifications, such as NACE, ISCO, ISCED, NUTS
- recording the same set of characteristics in each country

In 2021, the quarterly EU-LFS sample size was around 1.1 million people ([26]).

EU-LFS covers all industries and occupations. The classification of economic activities of employers is the Statistical classification of economic activities in the European Community (NACE). NACE Rev. 2 is in use since 2008, at 3-digit level. The classification of the occupation of employed persons is the International standard classification of occupations (ISCO) developed by the International Labour Organisation. The classification was last revised in 2008 (ISCO-08).

Cut-off date

The cut-off date for data collection for EU Labour Force Survey data downloaded from Eurostat's dissemination database (Eurostat) was 5 September 2023.

([26]) DG Research and Innovation, European Research Area Policy Agenda – Overview of actions for the period 2022-2024, Publications Office of the European Union, Luxembourg, 2021, https://commission.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

5.2.1. Proportion of researchers employed part-time among researchers in the Higher Education Sector (HES), by sex

Definition of indicator

This indicator compares the proportion of persons employed part-time among women and among men researchers. It covers the higher education sector (HES) only.

Rationale

Part-time work is an important feature of working conditions with noteworthy gender aspects. The predominance of women in part-time work is on the one hand often explained by gender stereotypes related to family responsibilities but is also linked to gender segregation in employment. On the other hand, part-time work might be seen as an instrument to increase the labour market participation – and thus, to a certain extent at least, the economic independence of women ([27]). Different types of work flexibility may have fewer negative, gender-specific consequences, as a critical analysis of part-time work in the Netherlands shows ([28]). The promotion of attractive, sustainable research careers and of gender equality in the research sector are two of the main actions of the new ERA ([29])([30]).

As a first step towards understanding this situation better, this indicator aims to consider the relative propensity of women and men higher education researchers to be employed part-time.

Computation method

Data needed

(F) Number of female researchers in the higher education sector (HES) aged between 25 and 64 who indicated their employment status. **Unit: Number.**

(F_p) Number of female researchers in the higher education sector (HES) aged between 25 and 64 who indicated that they worked part-time. **Unit: Number**

(M) Number of male researchers in the higher education sector (HES) aged between 25 and 64 who indicated their employment status. **Unit: Number**.

(M_p) Number of male researchers in the higher education sector (HES) aged between 25 and 64 who indicated that they worked part-time. **Unit: Number.**

Source of data

Eurostat - Custom data extraction from EU-LFS microdata

Computation formula

This indicator compares the proportion of the women researcher population that work part-time with the proportion of the men researcher population that work part-time.

([27]) Burri, S. and Aune, H., *Sex Discrimination in Relation to Part-Time and Fixed-Term Work*, European Network of Legal Experts in the Field of Gender Equality, 2013.

([28]) Vinkenburg, C.J., van Engen, M., and Peters, C.P., 'Promoting new norms and true flexibility: Sustainability in combining career and care', in de Vos, A., & van der Heijden, B. (eds), *Handbook of research on sustainable careers*, Elgar, London, 2015, pp. 131-145.

([29]) European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation*. COM(2020)0628 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A628%3AFIN

([30]) DG Research and Innovation, *European Research Area Policy Agenda – Overview of actions for the period 2022-2024*, Publications Office of the European Union, Luxembourg, 2021, https://commission.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

The formula for this indicator is as follows:

Proportion of persons employed part time among female researchers in the HES = F_p/F

Proportion of persons employed part time among male researchers in the HES = M_p/M

where:

p denotes part-time employment.

Specifications

This indicator focuses on researchers in the higher education sector only.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([31]).

However, the LFS does not identify researchers as a distinct occupation. The LFS classifies occupations according to the International Standard Classification of Occupations (ISCO-08). For the purposes of the indicator, a person is considered as a researcher in the higher education sector if:

- The person's occupation in the person's main job is one of the following, according to ISCO-08:
    (a) 1223: 'Research and Development Managers'
    (b) 21: 'Science and Engineering Professionals'
    (c) 231: 'University and Higher Education Teachers'
    (d) 221: 'Medical Doctors'

AND

- The economic activity of the person's employer is 854, i.e., 'Higher education' according to the statistical classification of economic activities in the European Community (NACE rev. 2).

Comments and critical issues

The higher education staff selected is a proxy of researchers working in the higher education sector.

5.2.2. Proportion of researchers in HES working under 'precarious' contracts, by sex

Definition of indicator

This indicator compares the proportion of higher education researchers with precarious working contracts by sex.

([31]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Rationale

The existence of, and increase in, precarious employment is subject to debate throughout the EU ([32]). Most affected are junior academic positions or other positions relying on third-party funding. The provision of research jobs associated with precariousness is in sharp conflict with the EU wide goal to provide attractive and secure positions in academia to fully exploit Europe's talent pool for HES within the new ERA ([33])([34]) and Horizon Europe initiatives ([35]).

This indicator aims to provide information on whether there is unequal representation of women and men in precarious forms of employment as a researcher. Women tend to have unequal caring responsibilities which may result in differences in the working conditions.

Computation method

Data needed

F Number of female researchers who indicated their contractual status aged between 25 and 64. **Unit: Number**.

F_p Number of female researchers in the higher education sector (HES) aged between 25 and 64 who indicated that they worked on a 'precarious' working contract. **Unit: Number**.

M Number of male researchers who indicated their contractual status aged between 25 and 64. **Unit: Number**.

M_p Number of male researchers in the higher education sector (HES) aged between 25 and 64 who indicated that they worked on a 'precarious' working contract. **Unit: Number**.

Source of data

Eurostat - Custom data extraction from EU-LFS microdata

Computation formula

The formula for this indicator is as follows:

Proportion of women researchers in HES working under 'precarious' contracts = $\frac{F_p}{F}$

Proportion of men researchers in HES working under 'precarious' contracts = $\frac{M_p}{M}$

where:

p denotes employment on a 'precarious' working contract.

Specifications

This indicator focuses on researchers in the higher education sector only.

([32]) DG for Internal Policies, *Precarious Employment in Europe Part 1: Patterns, Trends and Policy Strategies*, Brussels, 2016, https://www.europarl.europa.eu/RegData/etudes/STUD/2016/587285/IPOL_STU(2016)587285_EN.pdf

([33]) European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation*. COM(2020)0628 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A628%3AFIN

([34]) DG Research and Innovation, *European Research Area Policy Agenda – Overview of actions for the period 2022-2024 [pdf]*, Publications Office of the European Union, Luxembourg, 2021, https://commission.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

([35]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([36]).

However, the LFS does not identify researchers as a distinct occupation. The LFS classifies occupations according to the International Standard Classification of Occupations (ISCO-08). For the purposes of the indicator, a person is considered as a researcher in the higher education sector if:

- The person's occupation in the person's main job is one of the following, according to ISCO-08:

    (e) 1223: 'Research and Development Managers'

    (f) 21: 'Science and Engineering Professionals'

    (g) 231: 'University and Higher Education Teachers'

    (h) 221: 'Medical Doctors'

AND

- The economic activity of the person's employer is 854, i.e., 'Higher education' according to the statistical classification of economic activities in the European Community (NACE rev. 2).

Furthermore, researchers are considered to have precarious working contracts if they have a fixed-term contract of three years or less.

Comments and critical issues

The higher education staff selected is a proxy of researchers working in the higher education sector.

Furthermore, the group of higher education researchers, as defined here, is a very narrow subset of the total population of persons in employment. The corresponding LFS sample sizes are so small that due to data privacy concerns no data can be released about researchers in employment without a contract.

Finally, due to the design of the LFS, data about temporary employment are not available by family status. As a consequence, in She Figures 2024 the indicator is not presented by family status.

5.2.3. Proportion of women among total employment in the EU

Definition of indicator

This indicator presents the proportion of women in total employment as a starting point for considering their participation in different fields and sectors of the labour market.

Rationale

Even though significant progress has been made, gender equality in total employment is far from reality and women are still under-represented in the labour market ([37]). Boosting women's participation in employment is also fundamental to the EU's strategy for sustainable growth ([38]). This indicator considers the current representation of women in the labour market in general.

([36]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([37]) European Commission, *Social Dimension of Europe (Reflection paper)*, 2017, https://commission.europa.eu/publications/reflection-paper-social-dimension-europe_en#:~:text=26%20April%202017-,Description,society%20and%20world%20of%20work.

([38]) European Commission, *Towards a sustainable Europe by 2030 (Reflection paper)*, 2019, https://commission.europa.eu/publications/sustainable-europe-2030_en

Computation method

Data needed

(F) Number of women in employment (aged 25–64). **Unit: Number**.

(T) Total number of people in employment (aged 25–64). **Unit: Number**.

Note that the numbers here are in thousands.

Source of data

Eurostat – Labour Market Statistics *(online data code:* lfsa_egan*)*

Computation formula

The formula for this indicator is as follows:

Proportion of women among total employment = F/T

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

Comments and critical issues

In the body of She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

It is important to ensure the same age range for all indicators in this figure. The age range 25–64 is available only through the detailed Labour Force Survey results, at data code ***lfsa_egan***. There are minor differences between the detailed LFS results and the general LFS results (online data code: ***lfsi_emp_a***).

5.2.4. Compound annual growth rate (CAGR) of people in employment in the EU, by sex

Definition of indicator

This indicator presents the average yearly growth in the number of women and men in total employment.

Rationale

Even though significant progress has been made, gender equality in total employment is far from reality and women are still under-represented in the labour market ([39]). Boosting women's participation in employment is also fundamental to the EU's strategy for sustainable growth ([40]) This indicator considers the current representation of women in the labour market in general.

([39]) European Commission, *Social Dimension of Europe (Reflection paper)*, 2017. https://commission.europa.eu/system/files/2020-07/reflection-paper-social-dimension-europe_en.pdf

([40]) European Commission, Towards a sustainable Europe by 2030 (Reflection paper), 2019. https://commission.europa.eu/system/files/2019-02/rp_sustainable_europe_30-01_en_web.pdf

Computation method

Data needed

(F) Number of women in employment (aged 25–64) in a start and an end year. **Unit: Number**.

(M) Number of men in employment (aged 25–64) in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Note that the numbers of persons here are in thousands.

Source of data

Eurostat – Labour Market Statistics *(online data code:* lfsa_egan*)*

Computation formula

The compound annual growth rate shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of employed women and men in a given period. For women and men, it is respectively computed as follows:

CAGR of women in employment = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of men in employment = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period (in other words, $e - s$);

F_s denotes the number of women in employment in the start year;

F_e denotes the number of women in employment in the end year;

M_s denotes the number of men in employment in the start year;

M_e denotes the number of men in employment in the end year.

For example, if there were 1000 men in employment in 2006 and 1500 in 2016, the calculation would be:

CAGR for men in employment =$100 \times \left[(1500/1000)^{1/10} - 1\right]\% = 4.14\,\%$

Specifications

The EU Labour Force Survey (LFS) defines **employed persons** as ‘all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work’.

Comments and critical issues

In the body of She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled ‘Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers

(S&E), and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

It is important to ensure the same age range for all indicators in this figure. The age range 25-64 is available only through the detailed Labour Force Survey results, at data code lfsa_egan. There are minor differences between the detailed LFS results and the general LFS results (online data code: *lfsi_emp_a*).

5.3. Eurostat – Human resources in science and technology

Content-based rationale

The European Commission has warned that 'gender segregation, or the tendency for men and women to take different jobs, is pervasive across Europe' ([41]). Historically, women have been under-represented in scientific and technical fields. Horizon Europe recognises gender equality and inclusiveness as a key issue. It aims to eliminate gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers, among others ([42]).

She Figures' indicators based on human resources in science and technology (HRST) data explore this situation further. Many are designed to consider the extent to which available human resources in science and technology are being fully utilised, and whether differences by sex persist. These include the following: the proportion of tertiary-educated women employed as professionals or technicians, and the proportion of scientists and engineers in the total labour force, by sex.

Broad overview of the source

These data can be accessed through the Human Resources in Science and Technology (HRST) database on the Eurostat website.

The **Human Resources in Science and Technology database** presents data on '**stocks**' and '**flows**'. Specifically, the data cover the 'demand for and supply of' highly qualified personnel (HQP) in the field of science and technology (S&T) and deal with 'stock', i.e., the current state of the labour force in S&T, and 'flow', i.e., the movement of HQP from job to job and from the academic sector to the public and private sectors. Data are disseminated on a yearly basis ([43]), and are used by both scientists and policymakers ([44]). Data from Eurostat are publicly available, regularly updated and accompanied by extensive methodological notes.

Many data breakdowns are available through the HRST database: sex, age, region, sector of economic activity, occupation, educational attainment and fields of education (however, not all combinations are possible). The HRST database uses some international classifications, including:

- The International Standard Classification of Education (ISCED 2011).
- The International Standard Classification of Occupations (ISCO-08).
- The Statistical classification of economic activities in the European Community (NACE Rev. 2).

In She Figures, indicators based on HRST data consider women and men's employment, including S&T occupations in general, and as scientists and engineers in particular. Additional data are required for these

([41]) European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020) 628 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([42]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Publications Office of the European Union, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([43]) European Commission, *Human resources in science and technology*, 2014. http://ec.europa.eu/eurostat/cache/metadata/en/hrst_esms.htm

([44]) European Commission, *Science, technology and innovation: Overview*, 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/High-tech_statistics#Database

indicators from the Labour Force Survey (LFS) database as indicated in the following subsections. For countries not covered by the EU-LFS (Associated and G20 countries), data were assessed through the International Labour Organization Statistics Database.

Cut-off date

The cut-off date for human resources in science and technology data downloaded from Eurostat's dissemination database (Eurostat) and ILO's database was 31 July 2023.

5.3.1. Proportion of women among tertiary-educated and employed as professionals or technicians (HRSTC) in the EU

Definition of indicator

This indicator presents the proportion of women within the Human Resources in Science and Technology Core group. This category covers those who have completed tertiary education (in any subject) and are employed in a science and technology (S&T) occupation (either as professionals or technicians).

Rationale

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for R&I, the Horizon 2020 and the Horizon Europe (2021-2027) programmes as well as, recognise the economic benefits that science and technology can deliver ([45]) , whilst the Europe 2030 strategy sees this as a priority growth area. The new ERA sees the increase of investment in innovation as critical to assisting the EU in catching up with the main R&D leaders (e.g., US, China) in terms of R&D intensity and in reinforcing its technological and industrial sovereignty (European Commission, 2020c)([46]).

This indicator considers the extent to which the available human resources in science and technology are being fully utilised, broken down by sex.

Computation method

Data needed

(F) Number of tertiary-educated women aged 25–64 who are employed as professionals or technicians (Human Resources in Science and Technology – Core (HRSTC)). **Unit: Number**.

(T) Total number of tertiary-educated people aged 25–64 who are employed as professionals or technicians (Human Resources in Science and Technology – Core (HRSTC)). **Unit: Number**.

Note that the numbers are in thousands.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

([45]) DG Research and Innovation, *HORIZON 2020 in Brief* [pdf], Publications Office of the European Union, Luxembourg, 2014. http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

([46]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, Publications Office of the European Union, Luxembourg, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

Computation formula

The formula for this indicator is:

Proportion of women among the HRSTC group = $F \ / \ T$

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 5) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 **(Professionals)** – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 **(Technicians and Associate Professionals)** – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

Comments and critical issues

In the body of She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E) and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

5.3.2. Compound annual growth rate (CAGR) of tertiary-educated people who are employed as professionals or technicians (HRSTC) in the EU, by sex

Definition of indicator

This indicator presents the average percentage growth each year in the number of women and men in the Human Resources in Science and Technology – Core (HRSTC) group. This covers those who have

completed tertiary education (in any subject) and are employed in a science and technology (S&T) occupation (either as professionals or technicians).

Rationale

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for R&I, the Horizon 2020 and the Horizon Europe (2021-2027) programmes, recognise the economic benefits that science and technology can deliver ([47]) ([48]), whilst the Europe 2030 strategy sees this as a priority growth area. The new ERA sees the increase of investment in innovation as critical to assisting the EU in catching up with its main R&D leaders (e.g., US, China) in terms of R&D intensity and in reinforcing its technological and industrial sovereignty (European Commission, 2020c).

This indicator considers whether there have been any changes to the use of available human resources in science and technology over time (broken down by sex).

Computation method

Data needed

(F) Number of tertiary-educated women, aged 25–64 who are employed as professionals or technicians (HRSTC) in a start and an end year. **Unit: Number**.

(M) Number of tertiary-educated men, aged 25–64 who are employed as professionals or technicians (HRSTC) in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Note that the numbers here are in thousands.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Computation formula

The CAGR shows the yearly average rate of growth for a given period. In this case, it shows the average percentage growth per year in the number of tertiary-educated women and men employed in S&T occupations. For women and men, it is respectively computed as follows:

CAGR of women = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of men = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period;

([47]) DG Research and Innovation, *HORIZON 2020 in Brief*, 2014. http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

([48]) DG Research and Innovation, *Horizon Europe; The EU Research and Innovation investment programme (2021-2027)*, 2019. https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf

F_s denotes the number of women in the HRSTC category in the start year;

F_e denotes the number of women in the HRSTC category in the end year;

M_s denotes the number of men in the HRSTC category in the start year;

M_e denotes the number of men in the HRSTC category in the end year.

For example, if there were 1 000 tertiary-educated women employed in S&T occupations in 2012, and 1 500 in 2016, the calculation would be: $100 \times [(1500/1000)^{1/4} - 1]\% = 10.7\ \%$.

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 5) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 **(Professionals)** – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 **(Technicians and Associate Professionals)** – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

Comments and critical issues

In the body of She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E) and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

5.3.3. Proportion of women among scientists and engineers (S&E) in the EU, by migration status

Definition of indicator

This indicator presents the proportion of women within the total number of scientists and engineers in employment by migration status (native-born, born in another EU member state and born outside the EU).

Rationale

According to a report by the European Commission, 'gender segregation, or the tendency for men and women to take different jobs, is pervasive across Europe. Only 16 % of all employees work in mixed occupations (i.e., where the proportions of men and women are between 40 % and 60 %)' ([49]). Traditionally, women are severely under-represented in the population of scientists and engineers. This is of particular significance given that such professionals are 'often the innovators at the centre of technology-led development' ([50]). Horizon Europe recognises gender equality and inclusiveness as a key issue. It aims to eliminate gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers, among others (DG Research and Innovation 2020)([51]).

By considering the sex breakdown for employed engineers and scientists, this indicator enables one to see whether there have been any advances in equalising the representation of women and men in this area.

For a full explanation of 'gender segregation' and other terms, see Annex 1.

Computation method

Data needed

(F_m) Number of women employed as scientists and engineers, aged 25–64, by migration status. **Unit: Number**.

(T_m) Total number of people employed as scientists and engineers, aged 25–64, by migration status. **Unit: Number**.

Note that the numbers are in thousands.

Source of data

Eurostat - Custom extraction from EU-LFS data.

Computation formula

Proportion of women among persons employed as S&E in migration status 'm' = F_m/T_m

where:

m denotes migration status.

([49]) European Commission, *Report on progress on equality between women and men in 2013* [pdf], SWD(2014) 142 final, 2014. https://op.europa.eu/en/publication-detail/-/publication/cae38103-6104-4a90-9771-17996c3d9762/language-en#:~:text=The%20Staff%20Working%20Document%20on,of%20equal%20value%3B%20equality%20in

([50]) European Commission, *Human resources in science and technology: Statistics explained*, 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/Human_resources_in_science_and_technology#Women_in_science_and_technology

([51]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

S&E (Scientists and engineers) are people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to the fields which are covered by one of the following occupations defined in the ISCO-08:

science and engineering professionals (ISCO-08 code: 21)

health professionals (ISCO-08 code: 22)

information and communications technology professionals (ISCO-08 code: 25)

- ISCO Sub-major Group 21 **(Science and engineering professionals)** – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes and on efficiency of production and work organisation; preparing scientific papers and reports.

- ISCO Sub-major Group 22 **(Health professionals)** – Occupations whose main tasks usually include: conducting research and obtaining scientific knowledge through the study of human and animal disorders and illnesses and ways of treating them; advising on or applying preventive and curative measures, or promoting health; preparing scientific papers and reports.

- ISCO Sub-major Group 25 **(Information and communications technology professionals)** – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

Comments and critical issues

In the body of She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E) and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

5.3.4. Compound annual growth rate (CAGR) of scientists and engineers (S&E) in the EU, by sex

Definition of indicator

This indicator presents the average percentage growth each year in the number of people employed as scientists and engineers, broken down by sex.

Rationale

The European Commission has reported on the persistence of 'gender segregation' in the labour market (the concentration of women and men in particular fields and at particular levels) ([52]). Traditionally, women are severely under-represented in the population of scientists and engineers. Horizon Europe recognises gender equality and inclusiveness as a key issue. It aims to eliminate gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers, among others ([53]).

This indicator enables one to see the rate at which women and men's employment as scientists and engineers has been growing over time. To reduce the gender imbalance, it is likely that women's representation will need to be growing at a faster rate than that for men.

For a full explanation of 'gender segregation' and other terms, see Annex 1

Computation method

Data needed

(F) Number of women, aged 25–64, employed as scientists and engineers in a start and an end year. **Unit: Number**.

(M) Number of men, aged 25–64, employed as scientists and engineers in a start and an end year. **Unit: Number**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Note that the numbers here are in thousands.

Source of data

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

Computation formula

The CAGR shows the yearly average rate of growth for a given period. In this case, it shows the average percentage growth of women and men employed as scientists and engineers. For women and men, it is respectively computed as follows:

CAGR of women scientists and engineers = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of men scientists and engineers = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

N denotes the number of years in the reference period;

([52]) European Commission, *Report on progress on equality between women and men in 2013*, SWD(2014) 142 final, 2014. https://op.europa.eu/en/publication-detail/-/publication/cae38103-6104-4a90-9771-17996c3d9762/language-en#:~:text=The%20Staff%20Working%20Document%20on,of%20equal%20value%3B%20equality%20in

([53]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

F_s denotes the number of women S&E in the start year;

F_e denotes the number of women S&E in the end year;

M_s denotes the number of men S&E in the start year;

M_e denotes the number of men S&E in the end year.

For example, if there were 100 women S&E in 2012, and 150 in 2016, the calculation would be: $CAGR\ for\ women\ S\&E = 100 \cdot [(150/100)^{1/4} - 1]\% = 10.7\ \%$.

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

S&E (Scientists and engineers) are people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to the fields which are covered by one of the following occupations defined in the ISCO-08:

science and engineering professionals (ISCO-08 code: 21)

health professionals (ISCO-08 code: 22)

information and communications technology professionals (ISCO-08 code: 25)

- ISCO Sub-major Group 21 **(Science and engineering professionals)** – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes, and on efficiency of production and work organisation; preparing scientific papers and reports.

- ISCO Sub-major Group 22 **(Health professionals)** – Occupations whose main tasks usually include: conducting research and obtaining scientific knowledge through the study of human and animal disorders and illnesses and ways of treating them; advising on or applying preventive and curative measures, or promoting health; preparing scientific papers and reports.

- ISCO Sub-major Group 25 **(Information and communications technology professionals)** – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

Comments and critical issues

In the body of She Figures, this indicator is presented in a figure that shows multiple indicators alongside each other. This figure is entitled 'Proportion of women in the EU among total employment; the population of tertiary educated professionals and technicians (HRSTC); and the population of scientists and engineers (S&E) and compound annual growth rate (CAGR) and trends in the numbers of women and men in the EU in the same populations'.

5.3.5. Proportion of tertiary educated and employed as professionals and technicians (HRSTC) among tertiary-educated population (HRSTE), by sex

Definition of indicator

This indicator identifies the proportions of highly educated men and women who are employed as professionals or technicians. Specifically, it presents the proportions of women and men who are tertiary educated and working in a science and technology occupation, out of the number of women and men who are tertiary educated. Those in science and technology occupations are those working as 'Professionals' or 'Technicians and Associate Professionals'.

Rationale

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for R&I, the Horizon 2020 and the Horizon Europe (2021-2027) programmes, recognise the economic benefits that science and technology can deliver ([54]) ([55]), whilst the Europe 2030 strategy sees this as a priority growth area. Furthermore, the new ERA sees the increase of investment in innovation as critical to assisting the EU in catching up with its main R&D leaders (e.g., US, China) in terms of R&D intensity and in reinforcing its technological and industrial sovereignty ([56]).

Computation method

Data needed

(F) Number of tertiary-educated women, aged 25–64 (HRSTE). **Unit: Number**.

(F_c) Number of tertiary-educated women, aged 25–64, who are also employed in S&T occupations (HRSTC). **Unit: Number**.

(M) Number of tertiary-educated men, aged 25–64 (HRSTE). **Unit: Number**.

(M_c) Number of tertiary-educated men, aged 25–64, who are also employed in S&T occupations (HRSTC). **Unit: Number**.

Note that the numbers here are in thousands.

Source of data

Eurostat - Custom extraction from EU-LFS data.

Computation formula

The formula for this indicator is:

Percentage of tertiary educated women working in S&T occupations = $F_c \ / \ F$

Percentage of tertiary educated men working in S&T occupations = $M_c \ / \ M$

([54]) DG Research and Innovation, HORIZON 2020 in Brief [pdf], Publications Office of the European Union, Luxembourg, 2014. http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

([55]) DG Research and Innovation, *Horizon Europe; The EU Research and Innovation investment programme (2021-2027)*, 2019. https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf

([56]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

where:

c denotes tertiary-educated people who are also working in an S&T occupation.

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 5) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 **(Professionals)** – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 **(Technicians and Associate Professionals)** – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

5.3.6. Proportion of tertiary educated and employed as professionals and technicians (HRSTC) among tertiary-educated population (HRSTE), by sex and migration status

Definition of indicator

This indicator identifies the proportions of highly educated men and women by migration status (native-born, born in another EU member state and born outside the EU), and who are employed as professionals or technicians. Specifically, it presents the proportions of women and men by migration status who are tertiary educated and working in a science and technology occupation, out of the number of women and men who are tertiary educated. Those in science and technology occupations are those working as 'Professionals' or 'Technicians and Associate Professionals'.

Rationale

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for R&I, the Horizon 2020 and the Horizon Europe (2021-2027) programmes,

recognise the economic benefits that science and technology can deliver ([57]) ([58]), whilst the Europe 2030 strategy sees this as a priority growth area. Furthermore, the new ERA sees the increase of investment in innovation as critical to assisting the EU in catching up with its main R&D leaders (e.g., US, China) in terms of R&D intensity and in reinforcing its technological and industrial sovereignty ([59]).

Computation method

Data needed

(F_m) Number of tertiary-educated women aged 25–64 (HRSTE), by migration status. **Unit: Number**.

$(F_{c,m})$Number of tertiary-educated women, aged 25–64, who are also employed in S&T occupations (HRSTC), by migration status. **Unit: Number**.

(M_m)Number of tertiary-educated men, aged 25–64 (HRSTE), by migration status. **Unit: Number**.

$(M_{c,m})$Number of tertiary-educated men, aged 25–64, who are also employed in S&T occupations (HRSTC), by migration status. **Unit: Number**.

Note that the numbers here are in thousands.

Source of data

Eurostat - Custom extraction from EU-LFS data.

Computation formula

The formula for this indicator is:

Percentage of tertiary educated women working in S&T occupations in migration status 'm' = $F_{c,m}/F_m$

Percentage of tertiary educated men working in S&T occupations in migration status 'm' = $M_{c,m}/M_m$

where:

c denotes tertiary-educated people who are also working in an S&T occupation

m denotes migration status.

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 5) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

([57]) DG Research and Innovation, *HORIZON 2020 in Brief* [pdf], Publications Office of the European Union, Luxembourg, 2014. http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

([58]) DG Research and Innovation, *Horizon Europe; The EU Research and Innovation investment programme (2021-2027)*, 2019. https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf

([59]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 **(Professionals)** – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 **(Technicians and Associate Professionals)** – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

5.3.7. Proportion of tertiary educated and employed as professionals and technicians (HRSTC) among tertiary-educated population (HRSTE), by sex and disability status

Definition of indicator

This indicator identifies the proportions of highly educated men and women with disabilities (severely limited, limited but not severely and not limited at all), and who are employed as professionals or technicians. Specifically, it presents the proportions of women and men with disabilities who are tertiary educated and working in a science and technology occupation, out of the number of women and men who are tertiary educated. Those in science and technology occupations are those working as 'Professionals' or 'Technicians and Associate Professionals'.

Rationale

Fostering greater investment in science and technology is a core part of the European vision for growth. The EU's main funding instruments for R&I, the Horizon 2020 and the Horizon Europe (2021-2027) programmes, recognise the economic benefits that science and technology can deliver ([60]) ([61]), whilst the Europe 2030 strategy sees this as a priority growth area. Furthermore, the new ERA sees the increase of investment in innovation as critical to assisting the EU in catching up with its main R&D leaders (e.g., US, China) in terms of R&D intensity and in reinforcing its technological and industrial sovereignty ([62]).

([60]) DG Research and Innovation, *HORIZON 2020 in Brief*, Publications Office of the European Union, Luxembourg, 2014. http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

([61]) DG Research and Innovation, *Horizon Europe; The EU Research and Innovation investment programme (2021-2027)*, 2019. https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf

([62]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

Computation method

Data needed

(F_d)Number of tertiary-educated women, aged 25–64 (HRSTE),by disability status. **Unit: Number**.

($F_{c,d}$)Number of tertiary-educated women, aged 25–64, who are also employed in S&T occupations (HRSTC), by disability status. **Unit: Number**.

(M_d)Number of tertiary-educated men, aged 25–64 (HRSTE), by disability status. **Unit: Number**.

($M_{c,d}$)Number of tertiary-educated men, aged 25–64, who are also employed in S&T occupations (HRSTC), by disability status. **Unit: Number**.

Note that the numbers here are in thousands.

Source of data

Eurostat - Custom extraction from EU-LFS data.

Computation formula

The formula for this indicator is:

Percentage of tertiary educated women working in S&T occupations in disability status 'd' = $F_{c,d}/F_d$

Percentage of tertiary educated men working in S&T occupations in disability status 'd' = $M_{c,d}/M_d$

where:

c denotes tertiary-educated people who are also working in an S&T occupation

d denotes disability status.

Specifications

According to the EU Labour Force Survey (LFS), **employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

HRST are persons that have successfully completed tertiary education (at least ISCED2011 level 5) or persons employed in S&T occupations even if not with tertiary education. Thus, the three categories of HRST are:

HRSTO: Persons employed in an S&T occupation;

HRSTE: Persons with tertiary education in any field of study;

HRSTC: Persons with tertiary education in any field of study and employed in an S&T occupation (the intersection of the former two groups).

S&T occupations are all occupations classified into major group 2, 'Professionals', or 3, 'Technicians and Associate Professionals', of the ISCO-08 International Standard Classification of Occupations.

- ISCO Major Group 2 **(Professionals)** – Occupations whose main tasks usually include: conducting analysis and research and developing concepts, theories and operational methods; advising on or applying existing knowledge related to physical sciences, mathematics, engineering and technology, life sciences, medical and health services, social sciences and humanities; teaching the theory and practice

of one or more disciplines at different educational levels; teaching and educating persons with learning difficulties or special needs; providing various business, legal and social services; creating and performing works of art; providing spiritual guidance; preparing scientific papers and reports.

- ISCO Major Group 3 **(Technicians and Associate Professionals)** – Occupations whose main tasks usually include: undertaking and carrying out technical work connected with research and the application of concepts and operational methods in the fields of physical sciences including engineering and technology, life sciences including the medical profession and social sciences and humanities; initiating and carrying out various technical services related to trade, finance and administration including administration of government laws and regulations and to social work; providing technical support for the arts and entertainment; participating in sporting activities; executing some religious tasks.

5.3.8. Proportion of scientists and engineers (S&E) among the total labour force, by sex

Definition of indicator

This indicator presents the proportion of scientists and engineers in the labour force, by sex.

Rationale

According to a report by the European Commission, 'gender segregation, or the tendency for men and women to take different jobs, is pervasive across Europe. Only 16 % of all employees work in mixed occupations (i.e., where the proportions of men and women are between 40 % and 60 %)' ([63]). Traditionally, women are severely under-represented in the population of scientists and engineers. This is of particular significance given that such professionals are 'often the innovators at the centre of technology-led development' ([64]). Horizon Europe recognises gender equality and inclusiveness as a key issue and aims to eliminate gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers among others ([65]).

By comparing the proportion of women and men engineers and scientists in the entire labour force, this indicator offers one measure of the level of segregation in this area (which is sometimes seen as connected to earlier segregation in the education pathways chosen by young women and men).

Computation method

Data needed

(F) Number of women working as scientists and engineers aged 25–64. **Unit: Number**.

(M) Number of men working as scientists and engineers aged 25–64. **Unit: Number**.

(T) Total number of people (both men and women) in the labour force, aged 25–64. **Unit: Number**.

Note that the number of people in the labour force (T) is in thousands.

([63]) European Commission, *Report on progress on equality between women and men in 2013*, SWD(2014) 142 final, 2014. https://op.europa.eu/en/publication-detail/-/publication/cae38103-6104-4a90-9771-17996c3d9762/language-en#:~:text=The%20Staff%20Working%20Document%20on,of%20equal%20value%3B%20equality%20in

([64]) European Commission, *Human resources in science and technology: Statistics explained*, 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/Human_resources_in_science_and_technology#Women_in_science_and_technology

([65]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

Source of data

For F and M:

Eurostat – Human Resources in Science & Technology *(online data code: hrst_st_ncat)*

For T:

Eurostat – Labour Market Statistics (online data code: *lfsa_agan)*

Computation formula

Proportion of women scientists and engineers among the total labour force = F/T

Proportion of men scientists and engineers among the total labour force = M/T

Specifications

According to the EU Labour Force Survey (LFS), **the labour force** (also termed 'active population') is defined as the sum of employed and unemployed persons. **Employed persons** are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'. **Unemployed persons** are 'all persons aged 15 to 74 who were not employed during the reference week, were available to start work within the two weeks following the reference week and had been actively seeking work in the four weeks preceding the reference week or had already found a job to start within the next three months'.

S&E (Scientists and engineers) are people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to the fields which are covered by one of the following occupations defined in the ISCO-08:

science and engineering professionals (ISCO-08 code: 21)

health professionals (ISCO-08 code: 22)

information and communications technology professionals (ISCO-08 code: 25)

- ISCO Sub-major Group 21 **(Science and engineering professionals)** – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes and on efficiency of production and work organisation; preparing scientific papers and reports.

- ISCO Sub-major Group 22 **(Health professionals)** – Occupations whose main tasks usually include: conducting research and obtaining scientific knowledge through the study of human and animal disorders and illnesses and ways of treating them; advising on or applying preventive and curative measures, or promoting health; preparing scientific papers and reports.

- ISCO Sub-major Group 25 **(Information and communications technology professionals)** – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture

specifications, data models and diagrams in the development, configuration and integration of computer systems.

Comments and critical issues

Note that this indicator is calculated differently to the 'Percentage of active population' unit at the data code *hrst_st_ncat*. In She Figures, the denominator for this indicator is the total labour force (women and men combined), whereas on Eurostat the denominator is restricted to either women or men.

5.3.9. Unemployment rate of tertiary educated people, by sex

Definition of indicator

This indicator presents the rate of unemployment for persons who have completed tertiary education, by sex.

Rationale

The education level does have significant influence on the employment level in most EU countries, but the differences of this influence vary crucially among the countries due to some historical reasons, labour market structural differences and unemployment insurance system peculiarities ([66]). The indicator helps to compare the employment outlook of women and men science and technology resources and to reveal potential disadvantages of one of the two sexes.

Computation method

Data needed

(F_u) Number of unemployed women, aged 25-64, with tertiary education. **Unit: Number**.

(F) Number of women in the labour force, aged 25-64, with tertiary education. **Unit: Number**.

(M_u) Number of unemployed men, aged 25-64, with tertiary education. **Unit: Number**.

(M) Number of men in the labour force, aged 25-64, with tertiary education. **Unit: Number**.

(T_u) Number of unemployed persons, aged 25-64, with tertiary education. **Unit: Number**.

(T) Number of persons in the labour force, aged 25-64, with tertiary education. **Unit: Number**.

Source of data

Eurostat - Human resources in Science and Technology (*online data code: hrst_st_nunesex*)

International Labour Organization (*https://ilo.org/global/statistics-and-databases*; Unemployment by sex, age and education).

Computation formula

Unemployment rate of tertiary educated women = F_u / F

([66]) Snieska V. et al., *Education and unemployment in European Union economic cycles*, 2015. https://www.sciencedirect.com/science/article/pii/S1877042815057833

Unemployment rate of tertiary educated men = M_u/M

Unemployment rate of tertiary educated persons = T_u/T

Specifications

According to the EU Labour Force Survey (LFS), **unemployed persons** are 'all persons aged 15 to 74 who were not employed during the reference week, were available to start work within the two weeks following the reference week and had been actively seeking work in the four weeks preceding the reference week or had already found a job to start within the next three months'. **The labour force** (also termed 'active population') is defined as the sum of employed and unemployed persons.

Comments and critical issues

Data from Eurostat do not need computation and can be extracted from the above-mentioned dataset. The indicator based on data from International Labour Organisation (ILO) needs to be computed with the formula shown above. Definition of unemployed persons does not change in ILO.

5.3.10. Proportion of women among self-employed individuals within Information and Communication Technology (ICT) and Science and Engineering (S&E) professionals

Definition of indicator

This indicator calculates the proportion of women who are self-employed within the occupations of Information and Communication Technology Professionals (ISCO-08 Division 25) and Science and Engineering Professionals (ISCO-08 Division 21).

Rationale

The European Commission's Entrepreneurship Action Plan 2020 points to the gender gap in the share of women who are self-employed in the EU ([67]). Furthermore, the European Commission's 'Women in Digital' policy aims to foster women's labour market participation in technology-oriented occupations and in knowledge-intensive sectors including ICT ([68]). Taking these priorities into account, this indicator sheds light on women's share of self-employment specifically within the ICT sector.

Computation method

Data needed

(F) Number of self-employed women with or without employees. **Unit: Number**.

(T) Total number of self-employed individuals with or without employees. **Unit: Numbe**r.

Source of data

Eurostat –European Labour Force Survey (EU-LFS).

([67]) European Commission, *Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of The Regions Entrepreneurship 2020 Action Plan COM(2012)0795*, 2013. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0795&from=EN

([68]) European Commission, *Women in Digital - Shaping Europe's Digital Future*, 2020. https://digital-strategy.ec.europa.eu/en/library/women-digital#:~:text=The%20future%20is%20about%20digital,equitable%20digital%20economy%20and%20society

Computation formula

Proportion of self-employed women among ICT and S&E professionals = F/T

Specifications

According to the EU Labour Force Survey, **self-employed persons with employees** refer to persons who work in their own business or professional practice for the purpose of earning a profit and who employ at least one other person. Similarly, **self-employed persons without employees** refer to persons who work in their own business or professional practice for the purpose of earning a profit and who do not employ any other person.

ISCO Sub-major Group 25 (**Information and communications technology professionals**) – Occupations whose main tasks usually include: researching information technology use in business functions; identifying areas for improvement and researching the theoretical aspects and operational methods for the use of computers; evaluating, planning and designing hardware or software configurations for specific applications including for Internet, Intranet and multimedia systems; designing, writing, testing and maintaining computer programs; designing and developing database architecture and database management systems; developing and implementing security plans and data administration policy and administering computer networks and related computing environments; analysing, developing, interpreting and evaluating complex system design and architecture specifications, data models and diagrams in the development, configuration and integration of computer systems.

ISCO Sub-major Group 21 (**Science and Engineering professionals**) – Occupations whose main tasks usually include: conducting research, enlarging, advising on or applying scientific knowledge obtained through the study of structures and properties of physical matter and phenomena, chemical characteristics and processes of various substances, materials and products, all forms of human, animal and plant life and of mathematical and statistical concepts and methods; advising on, designing and directing construction of buildings, towns and traffic systems, or civil engineering and industrial structures, as well as machines and other equipment; advising on and applying mining methods and ensuring their optimum use; surveying land and see and making maps; studying and advising on technological aspects of particular materials, products and processes and on efficiency of production and work organisation; preparing scientific papers and reports.

5.4. Eurostat – High-tech industry and knowledge-intensive services

Content-based rationale

These indicators were originally developed in line with Europe's 2020 vision of 'smart growth' to determine the extent to which women's full educational capacities are being exploited and as a way to gauge the EU's use of available human capital and women's role within a priority area of the economy. The European Commission is building a European Education Area by 2025 to harness the full potential of education, training and culture as drivers for job creation, economic growth and social fairness ([69]). The fast-changing labour market and societal transitions require people with high-level skills and advanced knowledge.

The indicators for high-tech industry and knowledge-intensive services include employment in knowledge-intensive activities (KIA) and employment in knowledge-intensive activities – business industries (KIABI).

Broad overview of the source

These data can be accessed from the Science, Technology and Innovation Statistics database on Eurostat's website.

'Statistics on high-tech industry and knowledge-intensive services' (sometimes referred to simply as 'high-tech statistics') cover statistics concerning employment, economic indicators, patents and products in the

([69]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

high-tech categories of the manufacturing sector, as well as the knowledge-intensive service sector ([70]). Data from Eurostat are publicly available, regularly updated and accompanied by extensive methodological notes.

Cut-off date

The cut-off date for data collection for high-tech industry and knowledge-intensive services data downloaded from Eurostat's dissemination database (Eurostat) was 31 July 2023.

5.4.1. Proportion of employment in knowledge-intensive activities (KIA) among total employment, by sex

Definition of indicator

This indicator presents the relative presence of employed women and men in KIA covering all sectors of the economy.

Rationale

Europe's strategy for sustainable growth aims to foster an economy based on people with advanced knowledge and high-level skills ([71]). This indicator reveals the extent to which women's full educational capacities are being utilised, by measuring the relative proportion of women and men in KIA.

Computation method

Data needed

(F) Number of women employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{F_k})$ Number of women employed specifically in KIA. **Unit: Number**.

(M) Number of men employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{M_k})$ Number of men employed specifically in KIA. **Unit: Number**.

(T) Total number of people employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{T_k})$ Total number of people employed specifically in KIA. **Unit: Number**.

Source of data

Eurostat – High-tech industry and knowledge-intensive services *(online data code: htec_kia_emp2)*

Note that this data code (*htec_kia_emp2*) provides the KIA employment rates and the numerators of the computation formula but not the denominators.

Computation formula

Respectively, the formulas for this indicator are:

([70]) European Commission, *High-tech statistics*, 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/High-tech_statistics#Database

([71]) European Commission, *Towards a sustainable Europe by 2030 (Reflection paper)*, 2019. https://commission.europa.eu/system/files/2019-02/rp_sustainable_europe_30-01_en_web.pdf

Proportion of women employed in knowledge intensive activities = F_k / F

Proportion of men employed in knowledge intensive activities = M_k / M

Proportion of persons employed in knowledge intensive activities = T_k / T

where:

k denotes knowledge-intensive-activities specifically.

Specifications

The **International Standard Classification of Education (ISCED 2011)** categorises education programmes by level. Tertiary-educated people are those who have graduated from the following stages:

- The first stage, which includes largely theory-based programmes to provide sufficient qualifications to gain entry to advanced research programmes and professions with high skills requirements and programmes which are practically, technically or occupationally specific (ISCED 5, 6 and 7).
- The second stage, which leads to the award of an advanced research qualification (e.g., PhD, non-PhD programmes with an advanced research component, etc.). The programmes are devoted to advanced study and original research (ISCED 8).

An activity is classified as '**knowledge-intensive'** if tertiary-educated people employed in this activity represent more than 33% of the total employment in the activity. The definition is based on the average number of employed persons aged 25–64 at aggregated EU level.

Women may be over-represented in some knowledge-intensive activities that are not related to science and research. This indicator does not enable one to analyse differences in representation across the activities.

The activities come from the NACE Rev. 2 categories (2-digit level) based on EU Labour Force Survey (LFS) data. NACE refers to the European Community's statistical classification of economic activities.

In this indicator, there is one aggregate in use based on the following classification: total knowledge-intensive activities (KIA). The lists of activities included in each aggregate, according to NACE Rev. 2 (2-digit level), are presented in the table below:

Table 1 Total knowledge-intensive activities (KIA), NACE Rev. 2

Codes	Description
09	Mining support service activities
19	Manufacture of coke and refined petroleum products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
26	Manufacture of computer, electronic and optical products
51	Air transport
58	Publishing activities
59	Motion picture, video and television programme production, sound recording
60	Programming and broadcasting activities
61	Telecommunications

Codes	Description
62	Computer programming, consultancy and related activities
63	Information service activities
64	Financial service activities, except insurance and pension funding
65	Insurance, reinsurance and pension funding, except compulsory social security
66	Activities auxiliary to financial services and insurance activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
72	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities
75	Veterinary activities
78	Employment activities
79	Travel agency, tour operator reservation service and related activities
84	Public administration and defence; compulsory social security
85	Education
86	Human health activities
90	Creative, arts and entertainment activities
91	Libraries, archives, museums and other cultural activities
94	Activities of membership organisations
99	Activities of extraterritorial organisations and bodies

5.4.2. Proportion of employment in knowledge-intensive activities – Business industries (KIABI) out of total employment, by sex

Definition of indicator

Similar to the previous indicator, this indicator shows the relative proportion of employed women and men in knowledge-intensive activities (KIA) (activities where more than one third of the workforce is tertiary educated), although it is restricted to business industries only.

Rationale

KIA are key to the EU's vision of fostering a knowledge-based economy. The term itself encompasses a wide range of activities, although the present indicator restricts the focus to business industries (KIABI). This is a

particularly important sector of the economy to examine, given that the EU considers 'innovation through business activities' to represent a strength of national R&I systems ([72]). Furthermore, the new ERA sees the increase of investment in innovation, especially in the business sector, as critical to assisting the EU in catching up with its main R&D leaders (e.g., US, China) in terms of R&D intensity and in reinforcing its technological and industrial sovereignty ([73]).

Assessing the relative proportion of women and men's employment in KIABI is thus a keyway of gauging the EU's use of available human capital, as well as the foundation for considering women's role within a priority area of the economy.

Computation method

Data needed

(F) Number of women employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{F_b})$ Number of women employed specifically in KIABI. **Unit: Number**.

(M) Number of men employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{M_b})$ Number of men employed specifically in KIABI. **Unit: Number**.

(T) Total number of people employed in all sectors of the economy. **Unit: Number**.

$(\mathrm{T_b})$ Total number of people employed in specifically in KIABI. **Unit: Number**.

Source of data

Eurostat – High-tech industry and knowledge-intensive services *(online data code: htec_kia_emp2)*

Note that this data code (*htec_kia_emp2*) provides the numerators ($\mathrm{F_b}$, $\mathrm{M_b}$, $\mathrm{T_b}$) but not the denominators (T, F or M) of the above computation formula. However, it provides the KIA employment rates.

Computation formula

Respectively, the formulas for this indicator are:

Proportion of women employed in KIABI = $\mathrm{F_b} / \mathrm{F}$

Proportion of men employed in KIABI = $\mathrm{M_b} / \mathrm{M}$

Proportion of persons employed in KIABI = $\mathrm{T_b} / \mathrm{T}$

Specifications

The **International Standard Classification of Education (ISCED 2011)** categorises education programmes by level. Tertiary-educated people are those who have graduated from the following stages:

- The first stage, which includes largely theory-based programmes to provide sufficient qualifications to gain entry to advanced research programmes and professions with high skills requirements and programmes which are practically, technically or occupationally specific (ISCED 5, 6 and 7).

([72]) This is evident in, for example, European Commission (2014), *Innovation Union Scoreboard 2014: The Innovation Union's Performance Scoreboard for Research and Innovation [Executive Summary].*

([73]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

- The second stage, which leads to the award of an advanced research qualification (e.g., PhD, non-PhD programmes with an advanced research component, etc.). The programmes are devoted to advanced study and original research (ISCED 8).

An activity is classified as '**knowledge-intensive'** if tertiary-educated people employed in this activity represent more than 33% of the total employment in the activity. The definition is based on the average number of employed persons aged 25–64 at aggregated EU level.

The activities come from the NACE Rev. 2 categories (2-digit level) based on EU Labour Force Survey (LFS) data. NACE refers to the European Community's statistical classification of economic activities.

In this indicator, there is one aggregate in use based on the following classification: Knowledge-intensive activities – Business industries (KIABI). The list of activities included in this aggregate, according to NACE Rev. 2 (2-digit level), is given below:

Table 2 Knowledge-intensive activities – Business industries (KIABI), NACE Rev. 2

Codes	Description
09	Mining support service activities
19	Manufacture of coke and refined petroleum products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
26	Manufacture of computer, electronic and optical products
51	Air transport
58	Publishing activities
59	Motion picture, video and television programme production, sound recording
60	Programming and broadcasting activities
61	Telecommunications
62	Computer programming, consultancy and related activities
63	Information service activities
64	Financial service activities, except insurance and pension funding
65	Insurance, reinsurance and pension funding, except compulsory social security
66	Activities auxiliary to financial services and insurance activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
72	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities

75	Veterinary activities
78	Employment activities
79	Travel agency, tour operator reservation service and related activities
90	Creative, arts and entertainment activities

5.5. Eurostat – Research and development statistics

Content-based rationale

Research and development (R&D) is central to the EU's vision for growth, as demonstrated by the Horizon 2020 programme, the Horizon Europe programme and the Europe 2020 and Europe 2030 strategies. Although, on average, a higher proportion of women in the EU are completing degrees than ever before, there are signs that they continue to lag behind men when it comes to their representation amongst the researcher population. This situation persists across all sectors, particularly in the business enterprise sector (BES). She Figures indicators based on R&D statistics explore women's presence in R&D personnel and especially as researchers, broken down by sector, as well as by field of R&D and age group.

In addition, some indicators consider R&D expenditure, to provide an insight into whether there are any correlations between spending levels and other factors.

Broad overview of the source

These data can be accessed through the Research and Development (R&D) node of the Science, Technology and Innovation database on the Eurostat website.

Eurostat's Statistics on R&D provide data on R&D spending and R&D personnel working in the main sectors of the economy: the business enterprise (BES), government (GOV), higher education (HES) and the private non-profit (PNP) sectors. R&D personnel data can be viewed in full-time equivalent (FTE), in head count (HC), as a percentage of employment and as a percentage of the labour force. Amongst other things, the data are disaggregated by occupation, qualification, sex, citizenship, age group, fields of R&D and economic activity (NACE Rev. 2).

Data from Eurostat is publicly available, regularly updated and accompanied by extensive methodological notes. The data are collected through samples, census surveys or administrative registers – or through a combination of sources.

Cut-off date

The cut-off date for R&D statistics downloaded from Eurostat's dissemination database (Eurostat) and OECD's database was 27 October 2023.

5.5.1. Proportion of women among researchers

Definition of indicator

This indicator presents the proportion of women out of the researcher population in all sectors of the economy, broken down by country.

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to EIGE, boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and

an increase in the relevance and quality of R&I outputs for all members of society ([74]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers, among others ([75]).

This indicator aims to shed light on whether there have been any improvements in the gender balance within the EU's researcher population.

Computation method

Data needed

(F) Number of female researchers in all sectors of the economy. **Unit: Head count/Full-time Equivalent**.

(T) Number of researchers in all sectors of the economy. **Unit: Head count/Full-time Equivalent**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org; R&D personnel by sector and function*)

Computation formula

Proportion of women among researchers = F/T

Specifications

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([76]).

Full-time equivalent (FTE) of R&D personnel is defined as the ratio of working hours actually spent on R&D during a specific reference period (usually a calendar year) divided by the total number of hours conventionally worked in the same period by an individual or by a group. ([77]).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([78]).

Comments and critical issues

For the countries that have data available in Eurostat, the indicator can be found calculated in the data code *rd_p_femres*.

The proportion of women among researchers is calculated both in Head Count (HC) and in Full-time Equivalent (FTE). The proportion of women among researchers that is presented in different tables and

([74]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

([75]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([76]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([77]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([78]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

figures of Chapter 4 of the publication is calculated in HC while the same proportion presented in Chapter 5 in combination with R&D Expenditure per capita researcher is calculated in FTE.

5.5.2. Compound annual growth rate (CAGR) for researchers, by sex

Definition of indicator

This indicator compares the average annual percentage change in the number of women and men in the researcher population over a particular period.

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to the European Institute for Gender Equality (EIGE), boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([79]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers, among others ([80]).

This indicator aims to capture the relative changes in women's and men's participation in the researcher population.

Computation method

Data needed

(F) Number of female researchers in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in a start and an end year. **Unit: Head count**.

(N) Number of years in reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number of years**.

Source of data

Eurostat – Statistics on research and *development (online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of women researchers and men researchers each year in a given period. It is calculated in the following way:

CAGR of female researchers = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of male researchers = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

([79]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

([80]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

s refers to the start year;

e refers to the end year;

$\mathrm{F_s}$ denotes the number of female researchers in the start year;

$\mathrm{F_e}$ denotes the number of female researchers in the end year;

$\mathrm{M_s}$ denotes the number of male researchers in the start year;

$\mathrm{M_e}$ denotes the number of male researchers in the end year.

For example, if there were 200 women researchers in 2016 and 150 in 2020, the calculation would be:

CAGR of women researchers = $100 \times [(150/200)^{1/4} - 1]\% = -6.9\ \%$.

Specifications

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([81]).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([82]).

Comments and critical issues

Numbers (HC) of men researchers are not given directly and need to be calculated as the difference of women researchers from total researchers.

In the EU, men are more likely to be employed as researchers than women. For a reduction of the gender gap in employment rates, the CAGR needs to be higher for women than it is for men.

5.5.3. Proportion of researchers per thousand labour force, by sex

Definition of indicator

This indicator presents the number of researchers for every thousand people in the labour force in a given country, broken down by sex.

Rationale

This indicator is another measure of the level of gender balance amongst the researcher population, given the historic tendency for this field to be dominated by men. Fostering equality in the representation of women and men amongst researchers demonstrates the EU's wider desire to 'reduce gender segregation at all levels in education and employment, as it contributes to inequalities in terms of the economic independence of women and men' ([83]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality from R&I systems through addressing unconscious biases and

([81]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([82]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([83]) Council of the European Union, *Council conclusions on women and the economy: Economic independence from the perspective of part-time work and self-employment*, 2014. http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/lsa/143269.pdf

systemic structural barriers, among others ([84]). Similarly, in line with the new ERA Communication ([85]) and the European Strategy for Universities (European Commission, 2022) adopted in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

Computation method

Data needed

(F) Number of female researchers. **Unit: Head count**.

(M) Number of male researchers. **Unit: Head count**.

$(\mathrm{F_i})$ Number of women in the labour force, aged 15 and over. **Unit: Number**.

$(\mathrm{M_i})$ Number of men in the labour force, aged 15 and over. **Unit: Number**.

Source of data

For F and M:

Eurostat – Statistics on research and development *(online data code: rd_p_persocc*)

For $\mathrm{F_i}$ and $\mathrm{M_i}$:

Eurostat – Labour Force Survey *(online data code: lfsa_agan*)

Note that the numbers from the Labour Force Survey are in thousands.

Computation formula

The formula for this indicator is:

Female researchers per thousand female labour force = $\mathrm{F} \,/\, \mathrm{F_i}$

Male researchers per thousand male labour force = $\mathrm{M} \,/\, \mathrm{M_i}$

Specifications

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([86]).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([87]).

According to the EU Labour Force Survey (LFS) the labour force (also termed 'active population') is defined as the sum of employed and unemployed persons.

([84]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([85]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([86]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([87]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Employed persons are 'all persons aged 15 years or more who worked at least one hour for pay or profit or family gain during the reference week or were temporarily absent from such work'.

Unemployed persons are 'all persons aged 15 to 74 who were not employed during the reference week, were available to start work within the two weeks following the reference week and had been actively seeking work in the four weeks preceding the reference week or had already found a job to start within the next three months'.

Comments and critical issues

F and *M* are in head count, i.e. number of persons, whereas F_i and M_i are in thousand persons. It is for this reason that the indicator states 'per thousand labour force'.

5.5.4. Proportion of women among researchers, by sector

Definition of indicator

This indicator presents the proportion of women among researchers in three broad economic sectors: the higher education sector (HES), the government sector (GOV) and the business enterprise sector (BES).

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates ([88]). Despite this, the EU's researcher population has continued to be dominated by men ([89]). According to EIGE, boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([90]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality from R&I systems through addressing unconscious biases and systemic structural barriers, among others ([91]).

This indicator enables greater analysis by considering the situation for researchers within different sectors.

Computation method

Data needed

(F_i) Number of female researchers in sector i, where the sector can be higher education, government and business enterprise. **Unit: Head count**.

(T_i) Total number of researchers in sector i, where the sector can be higher education, government and business enterprise. **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

([88]) Eurostat, *Statistics Explained: Tertiary education statistics*, 2023. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics

([89]) DG Research and Innovation, *She Figures 2018*, Publications Office of the European Union, Luxembourg, 2019. https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en

([90]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

([91]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

Computation formula

Proportion of women among researchers in sector i = F_i / T_i

where:

i denotes the sector (either HES, GOV or BES).

Specifications

The Frascati Manual (OECD, 2015) identifies and defines four **sectors of the economy, namely** the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES), the private non-profit sector (PNP) and the Rest of the world. The definitions for the first three of these (included in this indicator) are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market non-profit institutes (NPIs) that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e. units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([92]).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([93]).

5.5.5. Distribution of researchers across sectors of employment, by sex

Definition of indicator

This indicator presents the distribution of women and men researchers across four broad sectors of activity: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES) and the private non-profit sector (PNP).

Rationale

This indicator enables one to compare the sectors in which women and men researchers work. There are many reasons why this may be of interest, partly arising from the economic changes currently affecting much

([92]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([93]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

of the EU. In 2017, almost half (45.65 %) of the research population worked in the business enterprise sector([94]). At the same time, the share of R&D expenditure in the same sector was 66 % (Eurostat, 2020). However, women have higher shares among researchers in the government or higher education sector ([95]).

Computation method

Data needed

(F) The number of female researchers in each of the four economic sectors: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES) and the private non-profit sector (PNP). **Unit: Head count**.

(M) The number of male researchers in each of the four economic sectors: the higher education sector (HES), the government sector (GOV), the business enterprise sector (BES) and the private non-profit sector (PNP). **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

Computation formula

This indicator shows how researchers are spread out across different sectors, broken down by sex.

To compute this indicator, perform these two calculations for *each* sector:

Distribution of female researchers across sectors = $\mathrm{F_i} / \mathrm{F_a}$

Distribution of male researchers across sectors = $\mathrm{M_i} / \mathrm{M_a}$

where:

$\mathrm{F_i}$ denotes the number of female researchers in a particular sector;

$\mathrm{M_i}$ denotes the number of male researchers in a particular sector;

$\mathrm{F_a}$ denotes the number of female researchers in all sectors;

$\mathrm{M_a}$ denotes the number of male researchers in all sectors.

For each sex, the proportions for the sectors are shown alongside one another (with a sum total of 100%).

For example, suppose there are 1,000 women researchers. Of these, 350 are in the HES, 224 are in the GOV sector, 326 are in the BES and 100 are in the PNP. The proportion of women researchers in each sector would be as follows:

HES: 350 / 1000 = 35 %

GOV: 224 / 1000 = 22.4 %

BES: 326 / 1000 = 32.6 %

([94]) See Eurostat, total R&D personnel by sectors of performance, occupation and sex [*rd_p_persocc*].

([95]) DG Research and Innovation, *She Figures 2018*, Publications Office of the European Union, Luxembourg, 2019. https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en

PNP: 100 / 1000 = 10 %

Sum total of 100 %.

Specifications

The Frascati Manual [96] identifies and defines **four sectors of the economy**, namely HES, GOV, BES, PNP and Rest of the world. The definitions for sectors included in this indicator are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e. units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

PNP (§3.75): 'The Private non-profit sector comprises all non-profit institutions serving households (NPISH), as defined in the System of National Accounts (SNA) 2008, except those classified as part of the Higher education sector. For completeness of presentation it comprises also, households and private individuals engaged or not engaged in market activities, as explained in the section 'Criteria for the classification of institutional sectors for R&D statistics' earlier in this chapter'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([97]).

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([98]).

5.5.6. Distribution of researchers in the higher education sector (HES) across fields of Research and Development, by sex

Definition of indicator

This indicator focuses on the higher education sector (HES) and presents the distribution of women and men researchers across the six major fields of R&D: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; humanities.

([96]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([97]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([98]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Rationale

Although women are more likely than men to have a higher education degree, they remain over-represented in fields of study that are linked to traditional women's roles such as care-related fields and are under-represented in science, mathematics, IT, engineering and related careers. As a result, inequality in occupations is taking new forms and, despite their investment in education, young women are still twice as likely as young men to be economically inactive ([99]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([100]). Similarly, in line with the new ERA Communication ([101]) and the European Strategy for Universities (European Commission, 2022)([102]) issued in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

Consistent with the indicators on ISCED 8 graduates, this indicator sheds light on the extent of gender segregation across different fields of R&D in the higher education sector (HES). It is particularly important to consider this sector, given that it is one of the main sources of employment for researchers in the EU. According to the most recent data, a great proportion of the researchers in the EU in 2021 (31.9 %) are employed in the higher education sector ([103]).

Computation method

Data needed

(F) Number of female researchers in the Higher Education Sector (HES). **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the Higher Education Sector (HES), in each field of Research and Development. **Unit: Head count**.

(M) Number of male researchers in the Higher Education Sector (HES). **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the Higher Education Sector (HES), in each field of Research and Development. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

Computation formula

This indicator shows how researchers are spread out across different fields of Research and Development, broken down by sex.

To compute this indicator, perform these two calculations for each field of R&D in turn:

([99]) European Commission, *Gender Equality Strategy 2020-2025*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

([100]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([101]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([102]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

([103]) See Eurostat, R&D personnel (Statistics Explained article)

Distribution of female researchers across fields of R&D = $\mathrm{F_i} \ / \ F$

Distribution of male researchers across fields of R&D = $\mathrm{M_i} \ / \ M$

For each sex, the proportions for the fields of Research and Development are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1,000 women researchers in the HES. Of these, 150 are in natural sciences, 170 in engineering and technology, 200 in medical sciences, 82 in agricultural and veterinary sciences, 250 in social sciences and 148 in humanities. The proportion of women researchers in the HES in each field of Research and Development would be as follows:

Natural sciences: 150 / 1000 = 15 %

Engineering and technology: 170 / 1000 = 17 %

Medical sciences: 200 / 1000 = 20 %

Agricultural and veterinary sciences: 82 / 1000 = 8.2 %

Social sciences: 250 / 1000 = 25 %

Humanities and arts: 148 / 1000 = 14.8 %

The total is 100 %.

Specifications

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D** (Table 2.2, p. 59) that are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H)

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([104]).

Comments and critical issues

The breakdown of researchers by field of R&D is performed according to the field in which they work and not according to the field of their qualification.

([104]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

5.5.7. Compound annual growth rate (CAGR) of women researchers in the higher education sector (HES), by field of Research and Development

Definition of indicator

This indicator presents the compound annual growth rate of women researchers in the Higher education sector (HES) in six major fields of R&D: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; humanities.

Rationale

Although women are more likely than men to have a higher education degree, they remain over-represented in fields of study that are linked to traditional women roles such as care-related fields and are under-represented in science, mathematics, IT, engineering and related careers. As a result, inequality in occupations is taking new forms and, despite their investment in education, young women are still twice as likely as young men to be economically inactive ([105]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([106]). Similarly, in line with the new ERA Communication ([107]) and the European Strategy for Universities ([108]) issued in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

Consistent with the indicators on ISCED 8 graduates, this indicator sheds light on the extent of gender segregation across different fields of R&D in the higher education sector (HES). It is particularly important to consider this sector, given that it is one of the main sources of employment for researchers in the EU. According to the most recent data, a great proportion of the researchers in the EU in 2021 (31.9 %) are employed in the higher education sector ([109]).

Computation method

Data needed

(F) Number of women researchers in each of the fields of Research and Development in the higher education sector, in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

Computation formula

The CAGR shows the average rate of growth per year**,** for a given period. In this case, it shows the average percentage growth of women researchers in each main field of Research and Development in the higher education sector (HES) in a given period.

([105]) European Commission, *Gender Equality Strategy 2020-2025*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

([106]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([107]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([108]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

([109]) See Eurostat, R&D personnel (Statistics Explained article)

For each field of Research and Development respectively, perform this calculation:

CAGR of female researchers in each field of R&D = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{F_s}$ the number of women researchers in the chosen field of R&D (HES) in the start year;

$\mathrm{F_e}$ the number of women researchers in the chosen field of R&D (HES) in the end year.

Specifications

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D**, which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H)

The breakdown of researchers by field of Research and Development is performed according to the field in which they work and not according to the field of their qualification.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([110]).

Comments and critical issues

In fields of Research and Development where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance in that field.

5.5.8. Proportion of women among researchers, by main field of Research and Development (FORD) and by sector (HES, GOV and BES)

Definition of indicator

This indicator presents the proportion of women researchers in each of the six fields of Research and Development: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; and humanities. It does so for the higher education sector (HES), the government sector (GOV) and the business enterprise sector (BES) in turn.

([110]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Rationale

The EU's commitment to tackling 'gender segregation at all levels in education and employment' encompasses the research fields in which women and men work ([111]). In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to EIGE, boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([112]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([113]). Similarly, in line with the new ERA Communication ([114]) and the European Strategy for Universities ([115]) issued in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

This indicator sheds light on the extent of gender segregation across different fields of research in the HES, GOV and BES sectors.

Computation method

Data needed

(F) Number of female researchers, broken down by sector (HES, GOV, BES) and field of Research and Development (FORD). **Unit: Head count**.

(T) Total number of researchers, broken down by sector (HES, GOV, BES) and field of Research and Development (FORD). **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code rd_p_perssci)*

Computation formula

For each field of Research and Development, perform this calculation:

Proportion of women among researchers in a FORD in the HES = $\mathrm{F_{hi}} / \mathrm{T_{hi}}$

Proportion of women among researchers in a FORD in the GOV = $\mathrm{F_{gi}} / \mathrm{T_{gi}}$

Proportion of women among researchers in a FORD in the BES = $\mathrm{F_{bi}} / \mathrm{T_{bi}}$

where:

i denotes a particular field of R&D (FORD);

([111]) Council of the European Union, *Council conclusions on women and the economy: Economic independence from the perspective of part-time work and self-employment*, 2014. http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/lsa/143269.pdf

([112]) European Commission, *Gender Equality Strategy 2020-2025*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

([113]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([114]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([115]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

h denotes the higher education sector;

g denotes the government sector;

b denotes the business enterprise sector;

F_{hi} denotes the number of female researchers working in the HES in a particular field of R&D;

T_{hi} denotes the total number of researchers working in the HES in the same field of R&D as that in F_{hi} ;

F_{gi} denotes the number of female researchers working in GOV in a particular field of R&D;

T_{gi} denotes the total number of researchers working in GOV in the same field of R&D as that in F_{gi} ;

F_{bi} denotes the number of female researchers working in the BES in a particular field of R&D;

T_{bi} denotes the total number of researchers working in the BES in the same field of R&D as that in F_{bi} .

For example, in a particular sector, suppose there are 1,200 people working as researchers. Of these, 150 work in natural sciences (68 of them women), 245 work in engineering and technology (80 of them are women), 300 work in medical sciences (178 of them are women), 95 work in agricultural and veterinary sciences (34 of them are women), 140 work in social sciences (75 are women) and finally, 270 work in humanities (125 are women).

The proportion of women among researchers in each field of R&D is as follows:

natural sciences: 68 / 150 = 45.3 %

engineering and technology: 80 / 245 = 32.7 %

medical sciences: 178 / 300 = 59.3 %

agricultural and veterinary sciences: 34 / 95 = 35.8 %

social sciences: 75 / 140 = 53.6 %

humanities and arts: 125 / 270 = 46.3 %

Specifications

The Frascati Manual (OECD, 2015) provides definitions for the six **main fields of R&D** (p. 59), which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H)

The breakdown of researchers by field of Research and Development is according to the field in which they work and not according to the field of their qualification.

The Frascati Manual ([116]) identifies and defines **four sectors of the economy**, namely HES, GOV, BES, PNP and Rest of the world. The definitions for the first three of these (included in this indicator) are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e. units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([117]).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([118]).

Comments and critical issues

In the body of She Figures, this indicator is presented for two reference years ito show the evolution of the proportion of women researchers in different fields and sectors (i.e. the extent of change over time).

5.5.9. Distribution of researchers in the government sector (GOV) across fields of Research and Development, by sex

Definition of indicator

This indicator focuses on the government sector (GOV) and presents the distribution of women and men researchers across the six fields of Research and Development: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; and humanities.

Rationale

The EU is committed to reducing 'gender segregation at all levels in education and employment', which includes the research fields in which women and men work. Indicators on horizontal segregation tend to focus on the higher education sector. However, in 2017, the government sector employed about 10 % of researchers in the EU ([119]), making it another sector of interest when considering researchers' career patterns

([116]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([117]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([118]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([119]) See Eurostat, total R&D personnel by sectors of performance, occupation and sex [*rd_p_persocc*].

and the extent of horizontal segregation ([120]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([121]). Similarly, in line with the new ERA Communication ([122]) and the European Strategy for Universities ([123]) issued in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

Computation method

Data needed

(F) Number of female researchers in the Government sector (GOV). **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the Government sector (GOV), in each field of Research and Development. **Unit: Head count**.

(M) Number of male researchers in the Government sector (GOV). **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the Government sector (GOV), in each field of Research and Development. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

Computation formula

This indicator shows how researchers are spread out across different fields of Research and Development, broken down by sex.

To compute this indicator, perform these calculations for each field of R&D in turn:

Distribution of female researchers across fields of R&D= $\mathrm{F_i} / \mathrm{F}$

Distribution of male researchers across fields of R&D = $\mathrm{M_i} / \mathrm{M}$

where:

i refers to a particular field of R&D;

$\mathrm{F_i}$ denotes the number of female researchers in the GOV sector, in a given field of R&D;

$\mathrm{M_i}$ denotes the number of male researchers in the GOV sector, in a given field of R&D.

For each sex, the proportions for the fields of R&D are shown alongside one another (with a sum total of 100 %).

([120]) For a definition of horizontal segregation, please refer to Annex 1.

([121]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([122]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([123]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

For each sex, the proportions for the fields of Research and Development are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1,000 women researchers in the GOV sector. Of these, 150 are in natural sciences, 170 in engineering and technology, 200 in medical sciences, 82 in agricultural and veterinary sciences, 250 in social sciences and 148 in humanities. The proportion of women researchers in the GOV sector in each field of Research and Development would be as follows:

Natural sciences: 150 / 1000 = 15 %

Engineering and technology: 170 / 1000 = 17 %

Medical sciences: 200 / 1000 = 20 %

Agricultural and veterinary sciences: 82 / 1000 = 8.2 %

Social sciences: 250 / 1000 = 25 %

Humanities and arts: 148 / 1000 = 14.8 %

The total of is 100 %.

Specifications

The Frascati Manual ([124]) provides definitions for the six main fields of R&D (Table 2.2, p. 59) that are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H)

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([125]).

Comments and critical issues

The breakdown of researchers by field of R&D is performed according to the field in which they work and not according to the field of their qualification.

([124]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([125]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

5.5.10. Compound annual growth rates (CAGR) of women researchers in the government sector (GOV) by field of Research and Development

Definition of indicator

This indicator presents the compound annual growth rate of women researchers in the Government sector (GOV) in six major fields of R&D: natural sciences; engineering and technology; medical sciences; agricultural and veterinary sciences; social sciences; humanities.

Rationale

The EU is committed to reducing 'gender segregation at all levels in education and employment', which includes the research fields in which women and men work. Indicators on horizontal segregation tend to focus on the higher education sector. However, in 2017, the government sector employed about 10 % of researchers in the EU ([126]), making it another sector of interest when considering researchers' career patterns and the extent of horizontal segregation ([127]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([128]). Similarly, in line with the new ERA Communication ([129]) and the European Strategy for Universities ([130]) adopted in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

Computation method

Data needed

(F) Number of women researchers in each of the fields of Research and Development in the Government sector, in a start and an end year. **Unit: Head count.**

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number.**

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

Computation formula

The CAGR shows the average rate of growth per year**,** for a given period. In this case, it shows the average percentage growth of women researchers in each main field of Research and Development in the Government sector (GOV) in a given period.

For each field of Research and Development respectively, perform this calculation:

CAGR of female researchers in each field of R&D = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

where:

([126]) See Eurostat, total R&D personnel by sectors of performance, occupation and sex [rd_p_persocc].

([127]) For a definition of horizontal segregation, please refer to Annex 1.

([128]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([129]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([130]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

s refers to the start year;

e refers to the end year;

F_s the number of women researchers in the chosen field of R&D (GOV) in the start year;

F_e the number of women researchers in the chosen field of R&D (GOV) in the end year.

Specifications

The Frascati Manual ([131]) provides definitions for the six **main fields of R&D** (p. 59), which are included in this indicator:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities and arts (H)

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([132]).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([133]).

Comments and critical issues

In fields of Research and Development where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance in that field.

5.5.11. Distribution of researchers in the business enterprise sector across economic activities (NACE Rev. 2), by sex

Definition of indicator

This indicator presents the distribution of men and women researchers across specific economic activities in the business enterprise sector: manufacturing; services of the business economy; and all other economic activities.

Rationale

Previous editions of She Figures (2006, 2009, 2012, 2015, 2018, 2021) have shown that, of the three main sectors of the economy (HES, GOV and BES), women researchers are worst represented in the business enterprise sector (BES), making up less than a fifth of such employees. Whilst other She Figures indicators

([131]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([132]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([133]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

give a picture of women's overall representation in this sector, this indicator provides an insight into the economic activities being pursued by women and men researchers within the sector (thereby exploring whether there is also gender segregation within BES).

Computation method

Data needed

(F) Number of female researchers in the business enterprise sector (BES). **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the business enterprise sector (BES), in each economic activity. **Unit: Head count**.

(M) Number of male researchers in the business enterprise sector (BES). **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the the business enterprise sector (BES), in each economic activity. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_bempoccr2)*

Note that this data code from Eurostat already combines codes G–N as 'Services of the business economy', as well as some of the 'Other NACE codes'.

Computation formula

This indicator covers three types of economic activities in the NACE Rev. 2 classification:

- 'Manufacturing' – Code C;
- 'Services of the business economy – Codes G–N combined;
- 'Other NACE codes' – Codes A, B, D–F, O–U.

The formula for this indicator is:

Distribution of female researchers (BES) across economic activities = Fi/F

Distribution of male researchers (BES) across economic activities = Mi/M

where:

i denotes a particular economic activity (for this indicator, either 'Manufacturing', 'Services of the business economy' or 'Other NACE codes');

For each sex, the proportions for the three types of economic activity are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 1,000 women researchers in the BES. Of these, 240 work in manufacturing, 340 in 'services of the business economy' and 420 in the remaining economic activities ('Other NACE codes'). The proportion of women researchers in the BES in each type of economic activity would be as follows:

manufacturing: 240 / 1000 = 24 %

services of the business economy: 340 / 1000 = 34 %

other NACE codes: 420 / 1000 = 42 %

The total is 100 %.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([134]).

Researchers in the business enterprise sector are categorised using the Statistical **Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev. 2)**. This has 21 main sections:

A Agriculture, forestry and fishing

B Mining and quarrying

C Manufacturing

D Electricity, gas, steam and air conditioning supply

E Water supply, sewerage, waste management and remediation activities

F Construction

G Wholesale and retail trade; repair of motor vehicles and motorcycles

H Transportation and storage

I Accommodation and food service activities

J Information and communication

K Financial and insurance activities

L Real estate activities

M Professional, scientific and technical activities

N Administrative and support service activities

O Public administration and defence; compulsory social security

P Education

Q Human health and social work activities

R Arts, entertainment and recreation

S Other service activities

T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

U Activities of extraterritorial organisations and bodies

For a full listing of the NACE Rev. 2 categories (including divisions and groups), please see Nace Rev.2 – Statistical classification of economic activities in the European Community

([134]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([135]).

5.5.12. Proportion of women among researchers in the business enterprise sector, by selected economic activities (NACE Rev. 2)

Definition of indicator

This indicator allows comparison of the proportion of women researchers across five different economic activities in the business enterprise sector: manufacturing; manufacturing of chemicals and chemical products; manufacturing of basic pharmaceutical products and pharmaceutical preparations; services of the business economy; other NACE codes.

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to EIGE, boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([136]).

Previous editions of She Figures (2006, 2009, 2012, 2015, 2018, 2021) have shown that, of the three main sectors of the economy (HES, GOV and BES), women researchers are worst represented in the business enterprise sector, making up less than a fifth of such employees. By considering individual economic activities, this indicator enables one to assess if this picture also holds in key sections of the sector.

Computation method

Data needed

(F) Number of female researchers in the business enterprise sector (BES). **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the business enterprise sector (BES), in each economic activity. **Unit: Head count**.

(T) Total number of researchers in the business enterprise sector (BES). **Unit: Head count**.

$(\mathrm{T_i})$ Total number of researchers in the business enterprise sector (BES), in each economic activity. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_bempoccr2)*

Note that this data code from Eurostat *rd_p_bempoccr2*) already combines codes G–N as 'Services of the business economy', as well as some of the 'Other NACE codes'.

Computation formula

This indicator covers five types of economic activities/divisions in the NACE Rev. 2 classification:

- 'Manufacturing' – Code C

([135]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([136]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

- ‘Manufacturing of chemicals and chemical products’- Code C20
- ‘Manufacture of basic pharmaceutical products and pharmaceutical preparations’ – Code C21
- ‘Services of the business economy – Codes G–N combined
- ‘Other NACE codes’ – Codes A, B, D–F, O–U

Applied to each activity/division in turn, the formula for this indicator is:

Proportion of women among researchers in a given economic activity = Fi/Ti

where:

i denotes a particular economic activity (one of the five covered by this indicator);

Note: Please ensure that the economic activity covered by Fi and Ti is the same.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([137]).

Researchers in the business enterprise sector are categorised using the Statistical **Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev. 2)**. This has 21 main sections. For this indicator, the most relevant sections are:

C Manufacturing, which includes two divisions:

C20: Manufacture of chemicals and chemical products;

C21: Manufacture of basic pharmaceutical products and pharmaceutical preparations.

G Wholesale and retail trade; repair of motor vehicles and motorcycles

H Transportation and storage

I Accommodation and food service activities

J Information and communication

K Financial and insurance activities

L Real estate activities

M Professional, scientific and technical activities

N Administrative and support service activities

The remaining sections (covered in ‘Other NACE codes’) are:

A Agriculture, forestry and fishing

([137]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

B Mining and quarrying

D Electricity, gas, steam and air conditioning supply

E Water supply, sewerage, waste management and remediation activities

F Construction

O Public administration and defence; compulsory social security

P Education

Q Human health and social work activities

R Arts, entertainment and recreation

S Other service activities

T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

U Activities of extraterritorial organisations and bodies

For a full listing of the NACE Rev. 2 categories (including divisions and groups), please see Nace Rev.2 – Statistical classification of economic activities in the European Community

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([138]).

5.5.13. Compound annual growth rate (CAGR) for researchers in the higher education sector (HES), by sex

Definition of indicator

This indicator compares the average annual rate of growth in women and men's employment as researchers in the higher education sector (HES) over a particular period.

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to EIGE, boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([139]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([140]). Similarly, in line with the new ERA Communication ([141]) and the European Strategy for Universities

([138]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([139]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

([140]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([141]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([142]) adopted in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

This indicator enables one to gauge changes in the patterns of women and men's employment as researchers over time, in the higher education sector (HES). Through comparing these results with those of the equivalent indicators for the government (GOV) sector and business enterprise sector (BES), it is also possible to consider whether increases/decreases in one sector are offset by those in another.

Computation method

Data needed

(F) Number of female researchers in the higher education sector in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in the higher education sector in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of women researchers and men researchers in the higher education sector (HES) each year in a given period.

It is calculated in the following way:

CAGR of female researchers in the HES = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of male researchers in the HES = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

Fs denotes the number of female researchers in the HES in the start year;

Fe denotes the number of female researchers in the HES in the end year;

$\mathrm{M_s}$ denotes the number of male researchers in the HES in the start year;

$\mathrm{M_e}$ denotes the number of male researchers in the HES in the end year.

([142]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

For example, if there were 100 women researchers in the HES in 2012 and 150 in 2016, the calculation would be:

CAGR of female researchers = $100 \times [(150/100)^{1/4} - 1]\% = 10.7\ \%$.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([143]).

The Frascati Manual (OECD, 2015) ([144]) identifies and defines **four sectors of the economy**, namely HES, GOV, BES, PNP and Rest of the world. The definition for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([145]).

Comments and critical issues

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

When calculating the CAGR for women and men, it is important to ensure the same reference period.

5.5.14. Compound annual growth rate (CAGR) of researchers in the government sector (GOV), by sex

Definition of indicator

This indicator compares the average annual rate of growth in women and men's employment as researchers in the government sector, over a particular period.

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to (EIGE), boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([146]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([147]). Similarly, in line with the new ERA Communication ([148]) and the European Strategy for Universities

([143]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([144]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([145]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([146]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

([147]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([148]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([149]) adopted in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

This indicator enables one to gauge changes in the patterns of women and men's employment as researchers over time, in the government sector (GOV). Through comparing these results with those of the equivalent indicators for the higher education sector (HES) and business enterprise sector (BES), it is also possible to consider whether increases/decreases in one sector are offset by those in another.

Computation method

Data needed

(F) Number of female researchers in the government sector in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in the government sector in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

Computation formula

The compound annual growth rate (CAGR) shows the average rate of growth per year**,** for a given period. In this case, it shows the average percentage growth of women researchers and men researchers in the government sector (GOV) each year in a given period.

It is calculated in the following way:

CAGR of female researchers in the GOV = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of male researchers in the GOV = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

Fs denotes the number of female researchers in the GOV in the start year;

Fe denotes the number of female researchers in the GOV in the end year;

Ms denotes the number of male researchers in the GOV in the start year;

Me denotes the number of male researchers in the GOV in the end year.

([149]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

For example, if there were 100 women researchers in the GOV in 2012 and 150 in 2016, the calculation would be:

CAGR of female researchers $=100 \times [(150/100)^{1/4} - 1]\% = 10.7\ \%$.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([150]).

The Frascati Manual ([151]) identifies and defines four sectors of the economy, namely HES, GOV, BES, PNP and Rest of the world. The definition for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([152]).

Comments and critical issues

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

When calculating the CAGR for women and men, it is important to ensure the same reference period.

5.5.15. Compound annual growth rate (CAGR) of researchers in the business enterprise sector (BES), by sex

Definition of indicator

This indicator compares the average annual rate of growth in women and men's employment as researchers in the business enterprise sector, over a particular period.

Rationale

In recent decades, women in the EU have made significant advances in raising their level of educational qualification, now making up a majority of all tertiary education graduates. Despite this, the EU's researcher population has continued to be dominated by men. According to (EIGE), boosting the proportion of women in the R&I workforce could have many benefits, including greater use of available talent, economic growth and an increase in the relevance and quality of R&I outputs for all members of society ([153]). Horizon Europe recognises gender equality and inclusiveness as a crosscutting priority and aims at tackling gender inequality in R&I systems through addressing unconscious biases and systemic structural barriers, among others ([154]). Similarly, in line with the new ERA Communication ([155]) and the European Strategy for Universities

([150]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([151]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([152]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([153]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

([154]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([155]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([156]) adopted in January 2022, the European Commission is committed to strengthening women's and girls' participation in Science, Technology, Engineering and Mathematics (STEM) studies and careers through a roadmap of activities.

This indicator enables one to gauge changes in the patterns of women and men's employment as researchers over time, in the business enterprise sector (BES). Through comparing these results with those of the equivalent indicators for the higher education sector (HES) and government sector (GOV), it is also possible to consider whether increases/decreases in one sector are offset by those in another.

Computation method

Data needed

(F) Number of female researchers in the business enterprise sector in a start and an end year. **Unit: Head count**.

(M) Number of male researchers in the business enterprise sector in a start and an end year. **Unit: Head count**.

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

Computation formula

The CAGR shows the average rate of growth per year for a given period. In this case, it shows the average percentage growth of women researchers and men researchers in the business enterprise sector (BES) each year in a given period.

It is calculated in the following way:

CAGR of female researchers in the BES = $(\mathrm{F_e}/\mathrm{F_s})^{1/N} - 1$

CAGR of male researchers in the BES = $(\mathrm{M_e}/\mathrm{M_s})^{1/N} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{F_s}$ denotes the number of female researchers in the BES in the start year;

$\mathrm{F_e}$ denotes the number of female researchers in the BES in the end year;

$\mathrm{M_s}$ denotes the number of male researchers in the BES in the start year;

$\mathrm{M_e}$ denotes the number of male researchers in the BES in the end year.

([156]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

For example, if there were 100 women researchers in the BES in 2012 and 150 in 2016, the calculation would be:

CAGR of female researchers = $100 \times [(150/100)^{1/4} - 1]\% = 10.7\ \%$

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([157]).

The Frascati Manual ([158]) identifies and defines **four sectors of the economy**, namely HES, GOV, BES, PNP and Rest of the world. The definition for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([159]).

Comments and critical issues

In areas where one sex is under-represented, a higher CAGR for that sex may signal a reduction in the gender imbalance.

When calculating the CAGR for women and men, it is important to ensure the same reference period.

5.5.16. Distribution of researchers in the higher education sector (HES) across age groups, by sex

Definition of indicator

This indicator shows the distribution of both women and men researchers in the higher education sector (HES) across different age groups.

Rationale

This indicator focuses on the higher education sector (HES) and can be compared with the results of the equivalent indicator for the government sector (GOV).

Considering the age distribution of researchers, it may reveal differences in the career patterns of women and men. For example, according to Eurostat, a higher proportion of women are outside of the labour force due to caring responsibilities, including for children ([160]). This may reduce their participation in the labour market during the key childbearing years of a particular country. On another level, by taking older age as a 'proxy' for seniority, this indicator can be used to gauge women and men's relative presence in the top research positions, against a backdrop of far-reaching under-representation of women in decision-making roles.

([157]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([158]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([159]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([160]) In 2019, in the EU, 37.3 % of women (aged 25 to 49) who were outside of the labour force were in the position due to looking after children or incapacitated adults. For men of the same age group outside of the labour force, the rate was 3.9 %. See Eurostat, 'Inactive Population – Main reason for not seeking employment – Distributions by sex and age (%)', data table *lfsa_igar*.

Computation method

Data needed

(F) Number of female researchers in the higher education sector (HES) aged 25 and over. **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the higher education sector (HES), in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

(M) Number of male researchers in the higher education sector (HES) aged 25 and over. **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the higher education sector (HES), in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persage)*

Computation formula

The formula for this indicator is:

Distribution of female researchers across age groups = $\mathrm{F_i}/F$

Distribution of male researchers across age groups = $\mathrm{M_i}/M$

where:

i denotes a particular age group;

For each sex, the proportions for the age groups are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 100 men researchers (aged 25 and over) in the HES in one country. Of these, 12 are aged 25–34; 26 aged 35–44; 38 aged 45–54; and 24 aged 55 and over. The proportion of men in each age group would be as follows:

aged 25-34: 12 / 100 = 12 %

aged 35-44: 26 / 100 = 26 %

aged 45-54: 38 / 100 = 38 %

aged 55 and over: 24 / 100 = 24 %

The total is 100 %.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([161]).

([161]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

The Frascati Manual ([162]) identifies and defines **four sectors of the economy**, namely HES, GOV, BES, PNP and Rest of the world. The definition for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([163]).

5.5.17. Distribution of researchers in the government sector (GOV) across age groups, by sex

Definition of indicator

This indicator shows the distribution of both women and men researchers in the government sector (GOV) across different age groups.

Rationale

This indicator focuses on the government sector (GOV) and can be compared with the results of the equivalent indicator for the higher education sector (HES).

Considering the age distribution of researchers may reveal differences in the career patterns of women and men. For example, according to Eurostat, a higher proportion of women are outside of the labour force due to caring responsibilities, including for children ([164]). This may reduce their participation in the labour market during the key childbearing years of a particular country. On another level, by taking older age as a 'proxy' for seniority, this indicator can be used to gauge women and men's relative presence in the top research positions, against a backdrop of far-reaching under-representation of women in decision-making roles.

Computation method

Data needed

(F) Number of female researchers in the GOV sector aged 25 and over. **Unit: Head count**.

$(\mathrm{F_i})$ Number of female researchers in the GOV sector, in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

(M) Number of male researchers in the GOV sector, aged 25 and over. **Unit: Head count**.

$(\mathrm{M_i})$ Number of male researchers in the GOV sector, in each of these age categories: 25–34; 35–44; 45–54; 55 and over. **Unit: Head count**.

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persage)*

Computation formula

([162]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([163]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([164]) Ibid.

The formula for this indicator is:

Distribution of female researchers across age groups = $\mathrm{F_i}/F$

Distribution of male researchers across age groups = $\mathrm{M_i}/M$

where:

i denotes a particular age group;

For each sex, the proportions for the age groups are shown alongside one another (with a sum total of 100 %).

For example, suppose there are 100 men researchers (aged 25 and over) in the GOV sector in one country. Of these, 12 are aged 25–34; 26 aged 35–44; 38 aged 45–54; and 24 aged 55 and over. The proportion of men in each age group would be as follows:

aged 25-34: 12 / 100 = 12 %

aged 35-44: 26 / 100 = 26 %

aged 45-54: 38 / 100 = 38 %

aged 55 and over: 24 / 100 = 24 %

The total is100 %.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([165]).

The Frascati Manual ([166]) identifies and defines **four sectors of the economy**, namely HES, GOV, BES, PNP and Rest of the world. The definition for the HES sector included in this indicator is:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([167]).

5.5.18. Dissimilarity Index for researchers in the higher education sector (HES) and government sector (GOV)

Definition of indicator

The Dissimilarity Index (DI) provides a theoretical measurement of the percentage of women and men in a field of R&D who would have to move to another field of R&D to ensure that the proportions of women were

([165]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([166]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([167]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

the same across all the possible fields of R&D. It can therefore be interpreted as the hypothetical distance from a balanced sex distribution across fields of R&D, based upon the overriding proportion of women.

Rationale

Although women are more likely than men to have a higher education degree, they remain over-represented in fields of study that are linked to traditional women's roles such as care-related fields and are under-represented in natural sciences, mathematics, IT, engineering and related careers ([168]). The new ERA will strengthen focus on the participation of women in Science, Technology, Engineering and Mathematics (STEM) fields ([169]).

This indicator shows the proportion of one sex or all employees that would need to change field of R&D in order to achieve a gender balance across those fields.

Computation method

Data needed

(F) Number of female researchers across all fields of R&D. **Unit: Head count**.

(F_i) Number of female researchers in each field of R&D. **Unit: Head count**.

(M) Number of male researchers across all fields of R&D. **Unit: Head count**.

(M_i) Number of male researchers in each field of R&D. **Unit: Head count**.

Source of data

Eurostat – Research and development statistics *(online data code: rd_p_perssci)*

Computation formula

This table presents the values of the Dissimilarity Index (DI) in the different countries for researchers in two sectors: higher education and government. Seven fields were considered in computing the DI: natural sciences; engineering and technology; medical and health sciences; agricultural and veterinary sciences; social sciences; humanities; and any other field of Research and Development. The full calculation method is explained under 'Specifications' below.

The formula for the Dissimilarity Index is:

$$DI = \frac{\sum_i | Fi / F - Mi / M |}{2}$$

where:

i denotes a particular R&D field.

For example, if we have three fields, A, B and C, with 17, 37 and 91 women and 108, 74, 182 men respectively, the overall proportion of women is 28.5 %. We therefore need to calculate:

$$\frac{\left|\frac{17}{145} - \frac{108}{364}\right| + \left|\frac{37}{145} - \frac{74}{364}\right| + \left|\frac{91}{145} - \frac{182}{364}\right|}{2} = \frac{0.1795 + 0.0519 + 0.1276}{2} = 0.1795$$

([168]) European Commission, *Gender Equality Strategy 2020-2025*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

([169]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

This means that 18 % of researchers will have to change field in order to maintain the background proportion of 28.5 % women in each field.

Specifications

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([170]).

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([171]).

Comments and critical issues

In order to interpret the DI correctly, it is important to know which sex is in the majority overall. The maximum value is 1, which indicates the presence of either women or men only in each of the fields, depending on the majority gender. The minimum value of 0 indicates a distribution between women and men within each occupation which is equal to the overall average proportion of women. If the same occupational categories are used for different countries, the DI yields a comparable, descriptive statistic that reflects the extent to which the two sexes are differently distributed. The results also depend on the number of categories. If more categories are used, the indicator will reflect greater variability in the distribution, which in turn will yield results indicating a higher level of segregation.

The index shown in She Figures is the Duncan and Duncan Index of Dissimilarity, first developed in the 1950s and now used extensively for international comparisons of inequality and dissimilarity (not solely between the sexes but also between other groups).

5.5.19. Distribution of R&D personnel across occupations, by sector of the economy and sex

Definition of indicator

This indicator presents the distribution of R&D personnel across three occupations (researchers, technicians, and other supporting staff), by sex in the three main sectors of the economy: higher education sector (HES), government (GOV) sector and business enterprise sector (BES).

Rationale

This indicator focuses on R&D personnel across all three sectors, namely the higher education sector, the government sector, and the business enterprise sector. Since this indicator corrects for the total number of personnel for each sex, it allows for a comparison of the presence of each sex across the different occupations.

Computation method

Data needed

$(M_{s,i})$ Number of men in a given R&D occupation and sector. **Unit: Head count**.

$(F_{s,i})$ Number of women in a given R&D occupation and sector. **Unit: Head count**.

(M_s) Number of men in all R&D occupations in a given sector. **Unit: Head count**.

([170]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([171]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

(F_s) Number of women in all R&D occupations in a given sector. **Unit: Head count**.

(i) Denotes a particular R&D occupation:

- Researchers
- Technicians
- Other supporting staff
- Technicians and other supporting staff (in cases where technicians and other supporting staff are not provided separately, see 'Comments and critical issues' in section 5.5.19)

(s) Denotes a sector of activity:

- higher education sector (HES)
- government sector (GOV)
- business enterprise sector (BES)
- total of all sectors (HES, GOV and BES)

Source of data

Eurostat – Statistics on research and development *(online data code: rd_p_persocc)*

OECD (*http://stats.oecd.org*; R&D personnel by sector and function)

Computation formula

This indicator presents the relative proportion of personnel per occupation by sex.

Distribution of female personnel across occupations by sector = $\mathrm{F_{s,i}}/\mathrm{F_s}$

Distribution of female personnel across occupations by sector = $\mathrm{M_{s,i}}/\mathrm{M_s}$

For each sex, the proportions for the occupations are shown alongside one another (with a total of 100 %).

For example, suppose there are 1,000 women R&D personnel in these three sectors. Of these, 390 work as researchers, 260 work as technicians and 350 work as other supporting staff. The proportion of women in each occupation would be as follows:

researchers: 390 / 1000 = 39 %

technicians: 260 / 1000 = 26 %

other supporting staff: 350 / 1000 = 35 %

The total is 100 %.

Specifications

The Frascati Manual ([172]) provides an international definition for **R&D personnel**, §5.6: 'All persons employed directly on R&D should be counted, as well as those providing direct services such as R&D managers, administrators, and clerical staff.' R&D personnel comprise three categories of occupations:

([172]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Researchers are professionals engaged in the conception or creation of new knowledge. They conduct research and improve or develop concepts, theories, models, techniques instrumentation, software or operational methods ([173]).

Technicians (and equivalent staff) are persons whose main tasks require technical knowledge and experience in one or more fields of engineering, physical and life sciences or social sciences and humanities. They participate in R&D by performing scientific and technical tasks involving the application of concepts and operational methods, normally under the supervision of researchers. Equivalent staff performs the corresponding R&D tasks under the supervision of researchers in the social sciences and humanities ([174]).

Other supporting staff includes skilled and unskilled craftsmen, secretarial and clerical staff participating in R&D projects or directly associated with such projects ([175]).

The Frascati Manual ([176]) identifies and defines **four sectors of the economy**, HES, GOV, BES, PNP and Rest of the world. The definitions for the first three of these (included in this indicator) are:

HES (§3.67): 'It comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions'.

GOV (§3.60): 'The Government sector consists of the following groups of resident institutional units: all units of central (federal), regional (state) or local (municipal) government including social security funds, except those units that provide higher education services or fit the description of higher education institutions provided in this manual. It consists also of all non-market NPIs that are controlled by government units that are not part of the Higher education sector'.

BES (§3.51): 'The Business enterprise sector comprises all resident corporations, including not only legally incorporated enterprises, regardless of the residence of their shareholders. This group also includes all other types of quasi-corporations, i.e. units capable of generating a profit or other financial gain for their owners that are recognised by law as separate legal entities from their owners and set up for purposes of engaging in market production at prices that are economically significant. It comprises also the unincorporated branches of non-resident enterprises that are deemed to be resident because they are engaged in production on the economic territory on a long-term basis and all resident NPIs that are market producers of goods or services or serve business'.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year ([177]).

Comments and critical issues

From the reference year 2012 onwards, it is not compulsory for countries to report technicians separately from other supporting staff when providing data for their R&D personnel ([178]). Therefore, distribution of R&D personnel across occupations is presented for the categories each country provides. In cases where data for the aggregate of technicians and other supporting staff are not provided, in any way, these are calculated with the subtraction of researchers from the total R&D personnel.

([173]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([174]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([175]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([176]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([177]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([178]) European Parliament, *Commission Implementing Regulation on statistics science and technology No 995/2012*, 2012. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0018:0030:EN:PDF

5.5.20. Total intramural R&D expenditure in purchasing power standards (PPS) per capita researcher in FTE, by sector of the economy

Definition of indicator

This indicator breaks down R&D expenditure per capita researcher in FTE by sector of economy (business enterprise, government, higher education or private non-profit) for a given year. To account for differences in prices, currency and exchange rates, the data are expressed in purchasing power standards (PPS).

Rationale

Although it does not provide any gender-specific information, the indicator should be viewed in conjunction with the indicator that addresses the distribution of researchers across sectors by sex, in order to see if there is any correlation between R&D spending and women researchers' presence.

Computational method

Data needed

(T_s) The overall number of researchers, in full-time equivalent (FTE), by sector of the economy (HES, GOV, BES, PNP, Total). **Unit: Full-Time Equivalent**.

(E_s) R&D expenditure in millions of purchasing power standards (PPS), by sector of the economy (HES, GOV, BES, PNP, Total). **Unit: Million PPS**.

(s) Denotes a particular sector:

- higher education sector (HES)
- government sector (GOV)
- business enterprise sector (BES)
- private non-profit (PNP)
- sum of the sectors (HES+GOV+BES+PNP)

Source of data

For T_s: Eurostat – Statistics on research and development *(online data: rd_p_persocc)*

For E_s: Eurostat – Statistics on research and development *(online data: rd_e_gerdtot)*

Computation formula

R&D expenditure in PPS per capita researcher (in a given sector) = $(E_s \times 1{,}000{,}000)/T_s$

Specifications

The definition of the **full-time equivalent (FTE)** unit of measurement of personnel employed in R&D, as proposed by the Frascati Manual corresponds to one year's work by one person on R&D.

The Frascati Manual defines intramural expenditures on **R&D as all expenditures for R&D performed** within a statistical unit or sector of the economy during a specific period, whatever the source of funds. It recommends using purchasing power parities (PPP) to express R&D statistics in monetary terms.

The **PPPs** are currency conversion rates that convert to a common currency and equalise the purchasing power of different currencies. They eliminate the differences in price levels between countries because economic indicators expressed in a national currency are converted into an artificial common currency, called the purchasing power standards (PPS).

5.6. Eurostat – Structure of Earnings Survey (SES)

Content-based rationale

The right of women and men to equal pay for work of equal value is enshrined in the European Pillar of Social Rights, which was endorsed in 2017. This is further supported by the Directive on pay transparency, proposed in 2021 ([179]) and adopted in 2023 ([180]).

The concentration of women and men in different economic activities ('sectoral gender segregation') or occupations ('occupational gender segregation') are some of the factors that can explain differences in average remuneration.

Considering that women are more often employed in low-paying sectors of the economy, sectoral gender segregation can help explain the wage gap between men and women. For example, women are more likely to work in education and health, while men are more likely to work in finance and IT. The earnings gap between women and men can also be explained by occupational gender segregation. The fact that men are more likely to be promoted to managerial and supervisory positions than women due to discrimination or self-policing may also contribute to occupational gender segregation. The term "glass ceiling" is a metaphor for a barrier that prevents women from moving up the career ladder.

However, differences in wages are not the only cause of gender earnings gaps. Different rates of part-time employment or of employment in general between women and men can also have a significant contribution. For this reason, She Figures 2024 has adopted a new indicator in the place of the Gender Pay Gap (GPG) used in earlier editions. The gender overall earnings gap (GOEG) is a synthetic indicator measuring the combined impact of three factors on the difference in average earnings between all women and men of working age, employed and not employed: (1) average hourly earnings, (2) monthly average of the number of hours paid, and (3) employment rate. As such, it aims to measure a more extended concept of the earnings gap between women and men. For more information, please refer to the presentation of the indicator in the rest of this section and to section 4.3 of Eurostat's report on gender pay gaps in the EU ([181]).

Broad overview of the source

The GOEG is published every four years by Eurostat. Its reference years – 2002, 2006, 2010, 2014 and 2018 – are calculated with data from the Structure of Earnings Survey (SES). The SES provides comparable data regarding earnings, and employee and employer characteristics such as gender, age, occupation, economic activity and enterprise size. The data are collected once every four years (most recently in 2022) and are released two years after the reference year. The indicator covers: (1) NACE economic activity sections B to S with exclusion of section O, "public administration, defense, and compulsory social security" ([182]); (2) businesses with at least 10 employees.

([179]) European Commission, *Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL to strengthen the application of the principle of equal pay for equal work or work of equal value between men and women through pay transparency and enforcement mechanisms*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0093

([180]) European Parliament, *Gender pay gap: Parliament adopts new rules on binding pay-transparency measures (Press release)*, 2023. 20230327IPR78545_en.pdf (europa.eu)

([181]) Eurostat, *Gender pay gaps in the European Union – a statistical analysis*, 2021, https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/KS-TC-22-002

([182]) Based on NACE Rev. 2. Further details can be found here: https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

Cut-off date

The cut-off date for Structures of Earnings Survey data downloaded from Eurostat's dissemination database (Eurostat) was 29 February 2024.

5.6.1. Gender overall earnings gap (%), by economic activity

Definition of indicator

This indicator presents the gender overall earnings gap in the economic activity 'Scientific research and development services' (NACE Rev. 2, Section M, Division 72) and in 'Total economy' (NACE Rev. 2, Sections B to S, excluding Section O).

Rationale

Despite more than 30 years of equal pay legislation, the gender pay gap has remained persistent across all Member States regardless of the overall level of women's employment, national welfare models or equality legislation. A gender-segregated labour market, the difficulty of balancing work and family life and the undervaluation of women's skills and work are some of the complex causes of the persistent gender pay gap ([183]).

Computation method

Data needed

$(\mathrm{E}_{f,i})$ Average hourly earnings of female employees in economic activity i. **Unit: Euro**.

$(\mathrm{H}_{f,i})$ Average monthly hours paid to female employees in economic activity i. **Unit: Number**.

$(\mathrm{R}_{f,i})$ Employment rate of females in economic activity i. **Unit: Percentage.**

$(\mathrm{E}_{m,i})$ Average hourly earnings of male employees in economic activity i. **Unit: Euro**.

$(\mathrm{H}_{m,i})$ Average monthly hours paid to male employees in economic activity i. **Unit: Number**.

$(\mathrm{R}_{m,i})$ Employment rate of males in economic activity i. **Unit: Percentage.**

The employment rate is the ratio of persons employed in a given economic activity to the labour force in the same economic activity. The labour force is the sum of persons employed in this activity and unemployed persons whose last employment was in this activity.

Computation formula

$$GOEG = \frac{\left(\mathrm{E}_{m,i} \cdot \mathrm{H}_{m,i} \cdot \mathrm{R}_{m,i}\right) - \left(\mathrm{E}_{f,i} \cdot \mathrm{H}_{f,i} \cdot \mathrm{R}_{f,i}\right)}{\left(\mathrm{E}_{m,i} \cdot \mathrm{H}_{m,i} \cdot \mathrm{R}_{m,i}\right)} \times 100$$

Source of data

Eurostat –Structure of earnings survey 2018 for earnings and number of hours paid

Eurostat - Labour Force Survey (LFS) for employment rates

([183]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research*, Publications Office of the European Union, Luxembourg, 2016. https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

Specifications

The Statistical Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev. 2) is used. This has 21 main sections:

A Agriculture, forestry and fishing

B Mining and quarrying

C Manufacturing

D Electricity, gas, steam and air conditioning supply

E Water supply, sewerage, waste management and remediation activities

F Construction

G Wholesale and retail trade; repair of motor vehicles and motorcycles

H Transportation and storage

I Accommodation and food service activities

J Information and communication

K Financial and insurance activities

L Real estate activities

M Professional, scientific and technical activities

N Administrative and support service activities

O Public administration and defence; compulsory social security

P Education

Q Human health and social work activities

R Arts, entertainment and recreation

S Other service activities

T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

U Activities of extraterritorial organisations and bodies

Division 72 'Scientific and development research' is in section M.

For a full listing of the NACE Rev. 2 categories (including divisions and groups), please see: http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2

A local unit is classified by its principal economic activity. A local unit may, however, also perform secondary activities. This means that NACE division 72 includes local units having their principal activity in research and development as defined in the NACE classification. Note that those local units could have some secondary activities. It is also possible that research and development is a secondary activity for some local units having another principal activity (e.g. in education). SES does not provide information on secondary activities. See more on the principal and secondary activities in the document on NACE Rev. 2 classification (pages 22-30):

http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

5.6.2. Gender overall earnings gap (%), by age group and economic activity

Definition of indicator

This indicator presents the gender overall earnings gap by age group (less than 35 years, 35-44 years, 45-54 years and 55 years or more) in the economic activity 'Scientific research and development services' (NACE Rev. 2, Section M, Division 72) and in 'Total economy' (NACE Rev. 2, Sections B to S, excluding Section O).

Rationale

Despite more than 30 years of equal pay legislation, the gender pay gap has remained persistent across all Member States regardless of the overall level of women's employment, national welfare models or equality legislation. A gender-segregated labour market, the difficulty of balancing work and family life, the undervaluation of women's skills and work are some of the complex causes of the persistent gender pay gap ([184]).

Computation method

Data needed

$(\mathrm{E}_{f,g,i})$ Average hourly earnings of female employees in age group g and economic activity i. **Unit: Euro**.

$(\mathrm{H}_{f,g,i})$ Average monthly hours paid to female employees in age group g and economic activity i. **Unit: Number**.

$(\mathrm{R}_{f,g,i})$ Employment rate of females in age group g and economic activity i. **Unit: Percentage.**

$(\mathrm{E}_{m,g,i})$ Average hourly earnings of male employees in age group g and economic activity i. **Unit: Euro**.

$(\mathrm{H}_{m,i})$ Average monthly hours paid to male employees in age group g and economic activity i. **Unit: Number**.

$(\mathrm{R}_{m,i})$ Employment rate of males in age group g and economic activity i. **Unit: Percentage.**

The employment rate is the ratio of persons of a given age group employed in a given economic activity to the labour force of the same age group and economic activity. The labour force is the sum of persons employed in this activity and unemployed persons whose last employment was in this activity.

Computation formula

$$GOEG = \frac{\left(\mathrm{E}_{m,g,i} \cdot \mathrm{H}_{m,g,i} \cdot \mathrm{R}_{m,g,i}\right) - \left(\mathrm{E}_{f,g,i} \cdot \mathrm{H}_{f,g,i} \cdot \mathrm{R}_{f,g,i}\right)}{\left(\mathrm{E}_{m,g,i} \cdot \mathrm{H}_{m,g,i} \cdot \mathrm{R}_{m,g,i}\right)} \times 100$$

Source of data

Eurostat –Structure of earnings survey 2018 for earnings and number of hours paid

Eurostat - Labour Force Survey (LFS) for employment rates

([184]) European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research* [pdf], Publications Office of the European Union, Luxembourg, 2016, https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

Specifications

The Statistical **Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev.2)** is used. This has 21 main sections:

A Agriculture, forestry and fishing

B Mining and quarrying

C Manufacturing

D Electricity, gas, steam and air conditioning supply

E Water supply, sewerage, waste management and remediation activities

F Construction

G Wholesale and retail trade; repair of motor vehicles and motorcycles

H Transportation and storage

I Accommodation and food service activities

J Information and communication

K Financial and insurance activities

L Real estate activities

M Professional, scientific and technical activities

N Administrative and support service activities

O Public administration and defence; compulsory social security

P Education

Q Human health and social work activities

R Arts, entertainment and recreation

S Other service activities

T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

U Activities of extraterritorial organisations and bodies

Division 72 'Scientific and development research' is in section M.

For a full listing of the NACE Rev. 2 categories (including divisions and groups), please see: http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2

A local unit is classified by its principal economic activity. A local unit may, however, also perform secondary activities. This means that NACE division 72 includes local units having their principal activity in research and development as defined in the NACE classification. Note that those local units could have some secondary activities. It is also possible that research and development is a secondary activity for some local units having another principal activity (e.g. in education).

SES does not provide information on secondary activities. See more on the principal and secondary activities in the document on NACE Rev. 2 classification (pages 22–30): http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

5.7. Institutional change indicator through web-scraping techniques

Content-based rationale

Institutional change is embedded in the European Commission's Gender Equality Strategy 2020-2025 ([185]), Strategic Plan for Horizon Europe 2021-2024 ([186]), and Strategic Plan for Horizon Europe 2025-2027 ([187]). A key instrument for creating institutional change in research organisations are Gender Equality Plans (GEPs). For calls with deadlines in 2022 and onwards, having a GEP in place is an eligibility criterion for all public bodies, higher education institutions and research organisations from EU Member States and associated countries wishing to participate in Horizon Europe ([188]).

Institutional change is also a key priority of the new European Research Area (ERA), with the Commission's 2020 Communication on the new ERA recognising the need to address diversity by opening policy to intersections with other social categorisations (e.g. ethnicity, disability and sexual orientation). As such, the Communication proposed to build on Horizon Europe and develop inclusive gender equality plans and policies with Member States and stakeholders.

Overview of the source

The data source is the textual content of the institutional websites of higher education institutes and other public research organisations. This content is searched with the help of specialised software (web-scraping) for the occurrence of particular terms or phrases which indicate the incidence of a state of interest.

The organisations covered comprise:

a) the higher education institutions (HEIs) listed in the European Tertiary Education Register (ETER);

b) the public bodies and research organisations that applied for funding under Framework Programme 7 (FP7), Horizon 2020 (H2020) or Horizon Europe framework programmes for R&I, irrespective of whether their application was successful. Initial lists were extracted by DG RTD from internal data on CORDA ([189]) and were cleaned of duplicate records. National Statistical Correspondents provided input with their countries' updated lists of public organisations from their national R&D survey.

The web-scraping was performed using SerpApi, a Google search application programming interface (API) though Python scripts. The detailed descriptions of the indicators based on web-scraping techniques follow below.

A limitation is that information included as PDF documents within a website, as it is often the case for GEPs, would not be captured by this approach. As a result, some organisations might describe relevant measures

([185]) European Commission, *Gender Equality Strategy 2020-2025*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

([186]) DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)*, Luxembourg, 2020. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

([187]) European Commission: Directorate-General for Research and Innovation, Horizon Europe strategic plan 2025-2027, Publications Office of the European Union, 2024, https://data.europa.eu/doi/10.2777/092911

([188]) For more information on the Gender Equality Plans as an eligibility criterion in Horizon Europe, see https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/democracy-and-rights/gender-equality-research-and-innovation_en#:~:text=As%20detailed%20in%20the%20last,research%20organisations%20from%20EU%20Member

([189]) Common Research Data Warehouse

in a format that remains undetected, potentially understating the extent of gender equality initiatives at the institutional level.

Cut-off date

The cut-off date for web-scraping was 1 February 2024.

5.7.1. Proportion of Research Performing Organisations (RPOs) that have taken measures and actions to promote Gender Equality, by type of organisation

Definition of indicator

The proportion of Research Performing Organisations (RPOs) that indicate in their websites that they have taken actions and measures to promote gender equality, by type of organisation (higher education institutions and public research organisations).

Rationale

A key instrument for institutional change is the development and implementation of targeted actions and measures towards gender equality in RPOs. For calls with deadlines in 2022 and onwards, having a GEP in place is an eligibility criterion for all public bodies, higher education institutions and research organisations from EU Member States and associated countries wishing to participate in Horizon Europe ([190]).

This indicator relies on web-scraping techniques to capture the proportion of RPOs that indicate in their website that they have taken actions and measures towards gender equality. This indicator builds on a similar indicator that was published in previous editions of She Figures, which is the 'Proportion of RPOs that have adopted GEPs'. Web-scraping detects websites whose content includes specific words and phrases and does not contain other specific words and phrases. Since it was not possible to specify sets of words and phrases which would match the existence of GEPs within the strict definition of the term ([191]), a broader term of 'measures and actions towards gender equality' was used.

Computation method

Data needed

(X) The number of RPOs in whose website was detected at least one specific search term indicating the presence of measures and actions taken by it to promote Gender Equality.

(T) Total number of RPOs whose website's content was scraped and searched for the specific terms.

Source of data

a) The higher education institutions (HEIs) listed in the European Tertiary Education Register (ETER).

b) The public bodies and research organisations that applied for funding under Framework Programme 7 (FP7), Horizon 2020 (H2020) or Horizon Europe framework programmes for R&I, irrespectively of whether their application was successful.

c) Input from statistical correspondents potentially revising list of organisations to be scraped.

([190]) DG Research and Innovation, *European Research Area Policy Agenda – Overview of actions for the period 2022-2024*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/490ee6ca-aa58-11ec-83e1-01aa75ed71a1

([191]) A GEP is a set of measures and actions aiming to create institutional and cultural change towards gender equality. It is put in place following a self-assessment of the organisation. GEPs sets concrete objectives and targets to be achieved, the strategies and practices adopted for their achievement, and the creation of effective monitoring and evaluation systems using indicators.

Computation formula

Proportion of RPOs that have taken actions and measures towards Gender Equality = X/T

Specifications

The organisations' websites were scraped using a specific list of terms and phrases, translated in each country's official language(s). The initial list of search phrases was developed in English though manual searches of a random sample of websites (their English versions) from all countries and used to identify common terms and phrases in the websites indicating actions and measures taken by the organisation to promote gender equality. Afterwards, exploratory web-scrapes were used to identify which of these terms and phrases capture more accurately actions and measures towards gender equality. The final list of search phrases (in English) is:

- Gender equality plan
- Equal opportunities officer
- Equal participation officer
- Eliminate/prevent sex discrimination
- Eliminate/prevent harassment
- Harassment policy
- Gender diversity committee
- Gender diversity office
- Gender diversity task force
- Gender equality body
- Gender equality policy
- Gender equality action
- Gender equality training
- Gender Action Policy
- Gender Action Plan
- Gender Mainstreaming
- Gender Focal Point
- Gender balance
- Gender dimension
- Gender-responsive innovation
- Gender-transformative
- Gender perspective
- Gender empowerment

- Women's careers
- Work-life balance
- Gender-based violence
- Equal treatment office
- Equal treatment officer
- Anti-discrimination action
- Diversity office
- Diversity officer
- Equality office
- Equality officer
- Inclusion office
- Inclusion officer
- Gender equality charter

The final list of search phrases was translated in all official languages of the participating countries and was sent to the Statistical Correspondents to review and correct if needed. For countries that have more than one official language (e.g., Belgium), the Statistical Correspondents indicated which organisations should be scraped in each language.

The search query combining these terms was formed after testing the results of the exploratory scrapes and is the following:

["Gender equality plan" OR "Equal opportunities officer" OR
"Equal participation officer" OR "Eliminate sex discrimination" OR
"Prevent sex discrimination" OR "Eliminate harassment" OR
"Prevent harassment" OR "Harassment policy" OR
"Gender diversity committee" OR "Gender diversity office"
OR "Gender diversity task force"
OR
["Gender equality body" AND "Gender mainstreaming"] OR
["Gender equality policy" AND "Gender mainstreaming"] OR
["Gender equality action" AND "Gender mainstreaming"] OR
["Gender equality training" AND "Gender mainstreaming"] OR
["Gender Action Policy" AND "Gender mainstreaming"] OR
["Gender Action Plan" AND "Gender mainstreaming"] OR
["Gender Focal Point" AND "Gender mainstreaming"] OR
["Gender-responsive innovation" AND "Gender mainstreaming"] OR
["Gender-based violence" AND "Gender mainstreaming"] OR
["Equal treatment office" AND "Gender mainstreaming"] OR
["Equal treatment officer" AND "Gender mainstreaming"] OR
["Anti-discrimination action" AND "Gender mainstreaming"] OR
["Diversity office" AND "Gender mainstreaming"] OR
["Diversity officer" AND "Gender mainstreaming"] OR
["Equality office" AND "Gender mainstreaming"] OR

[“Equality officer” AND “Gender mainstreaming”] OR
[“Inclusion office” AND “Gender mainstreaming”] OR
[“Inclusion officer” AND “Gender mainstreaming”] OR
[“Gender equality charter” AND “Gender mainstreaming”] OR
OR
[“Gender equality body” AND “Gender balance”] OR
[“Gender equality policy” AND “Gender balance”] OR
[“Gender equality action” AND “Gender balance”] OR
[“Gender equality training” AND “Gender balance”] OR
[“Gender Action Policy” AND “Gender balance”] OR
[“Gender Action Plan” AND “Gender balance”] OR
[“Gender Focal Point” AND “Gender balance”] OR
[“Gender-responsive innovation” AND “Gender balance”] OR
[“Gender-based violence” AND “Gender balance”] OR
[“Equal treatment office” AND “Gender balance”] OR
[“Equal treatment officer” AND “Gender balance”] OR
[“Anti-discrimination action” AND “Gender balance”] OR
[“Diversity office” AND “Gender balance”] OR
[“Diversity officer” AND “Gender balance”] OR
[“Equality office” AND “Gender balance”] OR
[“Equality officer” AND “Gender balance”] OR
[“Inclusion office” AND “Gender balance”] OR
[“Inclusion officer” AND “Gender balance”] OR
[“Gender equality charter” AND “Gender balance”] OR
OR
[“Gender equality body” AND “Gender dimension”] OR
[“Gender equality policy” AND “Gender dimension”] OR
[“Gender equality action” AND “Gender dimension”] OR
[“Gender equality training” AND “Gender dimension”] OR
[“Gender Action Policy” AND “Gender dimension”] OR
[“Gender Action Plan” AND “Gender dimension”] OR
[“Gender Focal Point” AND “Gender dimension”] OR
[“Gender-responsive innovation” AND “Gender dimension”] OR
[“Gender-based violence” AND “Gender dimension”] OR
[“Equal treatment office” AND “Gender dimension”] OR
[“Equal treatment officer” AND “Gender dimension”] OR
[“Anti-discrimination action” AND “Gender dimension”] OR
[“Diversity office” AND “Gender dimension”] OR
[“Diversity officer” AND “Gender dimension”] OR
[“Equality office” AND “Gender dimension”] OR
[“Equality officer” AND “Gender dimension”] OR
[“Inclusion office” AND “Gender dimension”] OR
[“Inclusion officer” AND “Gender dimension”] OR
[“Gender equality charter” AND “Gender dimension”] OR
OR
[“Gender equality body” AND “Gender-transformative”] OR
[“Gender equality policy” AND “Gender-transformative”] OR
[“Gender equality action” AND “Gender-transformative”] OR
[“Gender equality training” AND “Gender-transformative”] OR
[“Gender Action Policy” AND “Gender-transformative”] OR
[“Gender Action Plan” AND “Gender-transformative”] OR
[“Gender Focal Point” AND “Gender-transformative”] OR

[“Gender-responsive innovation” AND “Gender-transformative”] OR
[“Gender-based violence” AND “Gender-transformative”] OR
[“Equal treatment office” AND “Gender-transformative”] OR
[“Equal treatment officer” AND “Gender-transformative”] OR
[“Anti-discrimination action” AND “Gender-transformative”] OR
[“Diversity office” AND “Gender-transformative”] OR
[“Diversity officer” AND “Gender-transformative”] OR
[“Equality office” AND “Gender-transformative”] OR
[“Equality officer” AND “Gender-transformative”] OR
[“Inclusion office” AND “Gender-transformative”] OR
[“Inclusion officer” AND “Gender-transformative”] OR
[“Gender equality charter” AND “Gender-transformative”] OR
OR
[“Gender equality body” AND “Gender perspective”] OR
[“Gender equality policy” AND “Gender perspective”] OR
[“Gender equality action” AND “Gender perspective”] OR
[“Gender equality training” AND “Gender perspective”] OR
[“Gender Action Policy” AND “Gender perspective”] OR
[“Gender Action Plan” AND “Gender perspective”] OR
[“Gender Focal Point” AND “Gender perspective”] OR
[“Gender-responsive innovation” AND “Gender perspective”] OR
[“Gender-based violence” AND “Gender perspective”] OR
[“Equal treatment office” AND “Gender perspective”] OR
[“Equal treatment officer” AND “Gender perspective”] OR
[“Anti-discrimination action” AND “Gender perspective”] OR
[“Diversity office” AND “Gender perspective”] OR
[“Diversity officer” AND “Gender perspective”] OR
[“Equality office” AND “Gender perspective”] OR
[“Equality officer” AND “Gender perspective”] OR
[“Inclusion office” AND “Gender perspective”] OR
[“Inclusion officer” AND “Gender perspective”] OR
[“Gender equality charter” AND “Gender perspective”] OR
OR
[“Gender equality body” AND “Gender empowerment”] OR
[“Gender equality policy” AND “Gender empowerment”] OR
[“Gender equality action” AND “Gender empowerment”] OR
[“Gender equality training” AND “Gender empowerment”] OR
[“Gender Action Policy” AND “Gender empowerment”] OR
[“Gender Action Plan” AND “Gender empowerment”] OR
[“Gender Focal Point” AND “Gender empowerment”] OR
[“Gender-responsive innovation” AND “Gender empowerment”] OR
[“Gender-based violence” AND “Gender empowerment”] OR
[“Equal treatment office” AND “Gender empowerment”] OR
[“Equal treatment officer” AND “Gender empowerment”] OR
[“Anti-discrimination action” AND “Gender empowerment”] OR
[“Diversity office” AND “Gender empowerment”] OR
[“Diversity officer” AND “Gender empowerment”] OR
[“Equality office” AND “Gender empowerment”] OR
[“Equality officer” AND “Gender empowerment”] OR
[“Inclusion office” AND “Gender empowerment”] OR
[“Inclusion officer” AND “Gender empowerment”] OR
[“Gender equality charter” AND “Gender empowerment”] OR

OR
[“Gender equality body” AND “Women’s careers”] OR
[“Gender equality policy” AND “Women’s careers”] OR
[“Gender equality action” AND “Women’s careers”] OR
[“Gender equality training” AND “Women’s careers”] OR
[“Gender Action Policy” AND “Women’s careers”] OR
[“Gender Action Plan” AND “Women’s careers”] OR
[“Gender Focal Point” AND “Women’s careers”] OR
[“Gender-responsive innovation” AND “Women’s careers”] OR
[“Gender-based violence” AND “Women’s careers”] OR
[“Equal treatment office” AND “Women’s careers”] OR
[“Equal treatment officer” AND “Women’s careers”] OR
[“Anti-discrimination action” AND “Women’s careers”] OR
[“Diversity office” AND “Women’s careers”] OR
[“Diversity officer” AND “Women’s careers”] OR
[“Equality office” AND “Women’s careers”] OR
[“Equality officer” AND “Women’s careers”] OR
[“Inclusion office” AND “Women’s careers”] OR
[“Inclusion officer” AND “Women’s careers”] OR
[“Gender equality charter” AND “Women’s careers”] OR
OR
[“Gender equality body” AND “Work-life balance”] OR
[“Gender equality policy” AND “Work-life balance”] OR
[“Gender equality action” AND “Work-life balance”] OR
[“Gender equality training” AND “Work-life balance”] OR
[“Gender Action Policy” AND “Work-life balance”] OR
[“Gender Action Plan” AND “Work-life balance”] OR
[“Gender Focal Point” AND “Work-life balance”] OR
[“Gender-responsive innovation” AND “Work-life balance”] OR
[“Gender-based violence” AND “Work-life balance”] OR
[“Equal treatment office” AND “Work-life balance”] OR
[“Equal treatment officer” AND “Work-life balance”] OR
[“Anti-discrimination action” AND “Work-life balance”] OR
[“Diversity office” AND “Work-life balance”] OR
[“Diversity officer” AND “Work-life balance”] OR
[“Equality office” AND “Work-life balance”] OR
[“Equality officer” AND “Work-life balance”] OR
[“Inclusion office” AND “Work-life balance”] OR
[“Inclusion officer” AND “Work-life balance”] OR
[“Gender equality charter” AND “Work-life balance”]]

The list of public research organisations for each country had as basis the corresponding list used for the 2021 edition of She Figures. That list had been approved by each country’s respective Statistical Correspondent.

The 2021 list was updated as follows:

- DG RTD extracted from CORDA lists of all public bodies and research organisations that applied for funding under Framework Programme 7 (FP7), Horizon 2020 (H2020) or Horizon Europe framework programmes for R&I, irrespective of whether their application was successful.

- These initial lists were cleaned to remove duplicate records and records that were already part of the lists used in 2021.

- The 2021 list and the ‘cleaned’ list of each country were merged, and the resulting merged list was sent to the Statistical Correspondents of each country for review.

- After review, the Statistical Correspondents replied with a revised list. These versions of the lists were used to conduct the web-scraping.

No sensitive information about any organisation was shared with the Statistical Correspondents. The only information shared was the name and URL of each organisation.

Comments and critical issues

The results of this indicator are estimates only. The accuracy of this indicator was calculated at 80% during the exploratory web-scraping phase. This means that the indicator correctly assigned organisations as having or not having taken actions and measures towards gender equality in 80% of the cases. The indicator’s sensitivity, specificity and precision were estimated at 70.4%, 88.6% and 85.2% respectively.

Moreover, it is known that not all organisations that applied for funding by FP7, H2020 or Horizon Europe are research organisations (e.g., municipal authorities). Statistical correspondents’ input contributed greatly to the removal of such organisations from the lists.

Finally, during the exploratory web-scraping analysis, it was noted that scraping PDF files attached to the websites produced many false positive results and decreased the accuracy of the indicator. Therefore, the indicator was calculated based only on the content of the websites and not of any attached PDFs. This represents a limitation, as GEPs are often published in PDF format. Consequently, some organisations may describe relevant measures in a format that remains undetected, potentially underreporting the extent of gender equality initiatives at the institutional level.

5.8. Women in Science (WiS) questionnaire

Content-based rationale

The indicators that stem from this source investigate the under-representation of women at the higher levels of the academic career path and in positions of power (known as the ‘glass ceiling’ phenomenon – whereby the representation of women decreases as the seniority of the role increases). They cover a wide range of sectors, particularly in science and technology, as well as the differences in success in obtaining research funding, by sex. Indicators computed include: the proportion of women academic staff by grade and in total; the proportion of women grade A staff by main field of Research and Development; the distribution of grade A staff across fields of Research and Development by gender; and the Glass Ceiling Index.

Broad overview of the source

The Women in Science (WiS) questionnaires were sent to appointed Statistical Correspondents and provide data in support of the sets of indicators investigating the under-representation of women at the higher levels of the academic career path, as detailed in the rationale above.

The detailed descriptions of the indicators based on the Women in Science questionnaire follow below.

5.8.1. Proportion of women among academic staff, by grade

Definition of indicator

This indicator presents the proportion of women among the persons occupying positions at different grades of an academic career for a given year.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women leave those careers. Overall, women are under-represented at the higher levels of the academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of academic career path as only 23.6 % of heads of higher education institutions, 31% of board members in research decision-making are women and 24.5% of board leaders are women ([192]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([193]).

As such, monitoring the number of women present at each level of academia is necessary to observe whether there is progress towards reducing vertical segregation ('the leaky pipeline'), defined as the under- or over-representation of a clearly identifiable group of workers in occupations or sectors at the top of an ordering based on 'desirable' attributes ([194]).

Computation method

Data needed

(F_{GY}) Number of female academic staff at a given grade G (G = A, B, C, D, or T [Total]) for a given year Y. **Unit: Head count**.

(M_{GY}) Number of male academic staff at a given grade G (G = A, B, C, D, or T [Total]) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among academic staff, by grade = $\frac{\mathrm{F_{GY}}}{\mathrm{F_{GY}+M_{GY}}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

([192]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([193]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

([194]) Gasser C.E. and Shaffer K.S., 'Career Development of Women in Academia: Traversing the Leaky Pipeline', *The Professional Counselor*, Vol. 4, No. 4, 2014, pp. 332-352. https://mdsoar.org/bitstream/handle/11603/5410/GASSER%20&%20SHAFFER%20-%20WOMEN%20IN%20ACADEMIA.pdf?sequence=3

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([195]).

Comments and critical issues

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

5.8.2. Proportion of women among grade A positions

Definition of indicator

This indicator presents the proportion of women among persons occupying the single highest-level academic positions in a given year.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women leave those careers. Overall, women are under-represented at the higher levels of the academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of academic career path as only 23.6 % of heads of higher education institutions, 31% of board members in research decision-making are women and 24.5% of board leaders are women ([196]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([197]).

As such, monitoring the number of women present at each level of academia is necessary to observe whether there is progress towards reducing vertical segregation ('the leaky pipeline'), defined as the under- or over-representation of a clearly identifiable group of workers in occupations or sectors at the top of an ordering based on 'desirable' attributes ([198]).

Computation method

Data needed

(F_{AY}) Number of women in grade A academic positions for a given year Y. **Unit: Head count**.

([195]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([196]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([197]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

([198]) Gasser C.E. and Shaffer K.S., 'Career Development of Women in Academia: Traversing the Leaky Pipeline', *The Professional Counselor*, Vol. 4, No. 4, 2014, pp. 332-352. https://mdsoar.org/bitstream/handle/11603/5410/GASSER%20&%20SHAFFER%20-%20WOMEN%20IN%20ACADEMIA.pdf?sequence=3

(M_{AY}) Number of men in grade A academic positions for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among grade A positions = $\frac{F_{AY}}{F_{AY}+M_{AY}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([199]).

Comments and critical issues

The classification of academic positions into grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

5.8.3. Proportion of grade A among academic staff, by sex

Definition of indicator

This indicator allows for a comparison of the number of women and men staff at the highest-level academic positions compared to the number of staff of the same sex across all academic positions, for a given year.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women leave those careers. Overall, women are under-represented at the higher levels of the academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of

([199]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

academic career path as only 23.6 % of heads of higher education institutions, 31% of board members in research decision-making are women and 24.5% of board leaders are ([200]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([201]).In this indicator, the low number of women grade A staff is compared to the overall number of women staff in academia, thereby correcting for the relative presence of women in academic positions overall. The advantage of such a calculation is that it moves beyond the absolute numbers of men and women in academic positions, which enhances comparability of the measure across different settings.

Computation method

Data needed

(F_{GY}) Number of female academic staff at a given grade G (G = A or T [Total]) for a given year Y. **Unit: Head count**.

(M_{GY}) Number of male academic staff at a given grade G (G = A or T [Total]) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of grade A among female academic staff = $\frac{\mathrm{F_{AY}}}{\mathrm{F_{TY}}}$

Proportion of grade A among male academic staff = $\frac{\mathrm{M_{AY}}}{\mathrm{M_{TY}}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([202]).

([200]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([201]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

([202]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Comments and critical issues

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

5.8.4. Proportion of women among academic staff, by main field of Research and Development and grade

Definition of indicator

This indicator looks at the presence of women in top academic positions across different fields of R&D, allowing for the identification of the fields in which women are more or less present for a given year.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women leave those careers. Overall, women are under-represented at the higher levels of the academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of academic career path as only 23.6 % of heads of higher education institutions, 31% of board members in research decision-making are women and 24.5% of board leaders are women ([203]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([204]). However, there may be some differences in the employment of women in top positions across different fields of Research and Development. As such, looking at the proportion of women in different positions of seniority reveals which fields of Research and Development have seen a more successful integration of women staff in top positions over time.

Computation method

Data needed

(F_{GSY}) Number of women at a given seniority grade G (G = A, B, C, D, or T [Total]), in main field of R&D S for reference year Y. **Unit: Head count**.

(M_{GSY}) Number of men at a given seniority grade G (G = A, B, C, D, or T [Total]), in main field of R&D S for reference year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among academic staff at grade G in main field S = $\frac{F_{GSY}}{F_{GSY}+M_{GSY}}$

([203]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([204]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020.

Proportion of women among academic staff at grade G in main field S = $\frac{M_{GSY}}{F_{GSY}+M_{GSY}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

The Frascati Manual ([205]) provides definitions for the six main fields of Research and Development (p.95). The following abbreviations are used:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities (H)

Unknown (U), although not a field of R&D in Frascati, has been added in the WiS questionnaire so that data can also be provided for academic staff whose field is unknown.

The breakdown by field of Research and Development is according to the field in which staff work and not according to the field of their qualification.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([206]).

Comments and critical issues

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

([205]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([206]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

5.8.5. Distribution of grade A staff across fields of Research and Development, by sex

Definition of indicator

This indicator reveals differences in the distribution of women and men grade A staff across the different fields of Research and Development for a given year, by presenting the relative proportion of grade A staff of a given sex by field.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women leave those careers. Overall, women are under-represented at the higher levels of the academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of academic career path as only 23.6 % of heads of higher education institutions, 31% of board members in research decision-making are women and 24.5% of board leaders are women ([207]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([208]).Since this indicator corrects for the total number of grade A staff for each sex, it allows for a comparison of the fields of R&D in which each sex is more or less present in the top levels.

Computation method

Data needed

(F_{ASY}) Number of grade A women in main field of R&D S for year Y. **Unit: Head count**.

(M_{ASY}) Number of grade A men in main field of R&D S for year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of S field of R&D among grade A women = $\frac{F_{ASY}}{\sum_S F_{ASY}}$

Proportion of S field of R&D among grade A men = $\frac{M_{ASY}}{\sum_S M_{ASY}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

([207]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([208]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD

The Frascati Manual ([209]) provides definitions for the six main fields of Research and Development (p.95). The following abbreviations are used:

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities (H)

Unknown (U), although not a field of R&D in Frascati, has been added in the WiS questionnaire so that data can also be provided for academic staff whose field is unknown.

The breakdown by field of Research and Development is according to the field in which staff work and not according to the field of their qualification.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([210]).

Comments and critical issues

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

5.8.6. Glass Ceiling Index

Definition of indicator

The Glass Ceiling Index (GCI) is an index comparing the proportion of women in academia (grades A, B and C) to the proportion of women in top academic positions (grade A positions; equivalent to full professorships in most countries), for a given year. The GCI is a positive function that ranges above 0 to infinity. A GCI of 1 indicates that there is no difference between women and men in the chance of being promoted. A score of less than 1 means that women are over-represented at grade A level and a GCI score of more than 1 points towards a glass ceiling effect, meaning that women are under-represented in grade A positions. In other

([209]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([210]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

words, the interpretation of the GCI is that the higher the value, the stronger the glass ceiling effect and the more difficult it is for women to move into a higher position.

Rationale

Both the Gender Statistics Database on women and men in decision-making and the Gender Equality Index of EIGE demonstrate the under-representation of women in positions of power, across a wide range of sectors in the EU. Moreover the European Parliament 'notes that despite positive changes in recent years, gender equality in science and academia has still not been achieved […] and points out the strikingly low presence of women in the highest academic and decision-making positions in scientific institutions and universities, which indicates the existence of a glass ceiling, that is, invisible barriers based on prejudices which stand in the way of women accessing positions of responsibility' ([211]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([212]).

The version of the index presented here measures the relative chance for women (as compared with men) of reaching a top academic position, correcting for the relative presence of women (as compared with men) in academic positions overall. As such, it indicates the opportunity, or lack of it, for women to move up the hierarchical ladder in their academic profession. The advantage of the GCI being a relative index is that it moves beyond the absolute numbers of men and women in possible academic positions, which enhances comparability of the measure across different settings.

Computation method

Data needed

(F_{GY}) Number of grade A, B and C (G subscript) women for a given year Y. **Unit: Head count**.

(M_{GY}) Number of grade A, B and C (G subscript) men for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

$$Glass\ Ceiling\ Index = \left(\frac{\frac{F_{AY} + F_{BY} + F_{CY}}{F_{AY} + F_{BY} + F_{CY} + M_{AY} + M_{BY} + M_{CY}}}{\frac{F_{AY}}{F_{AY} + M_{AY}}} \right)$$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definitions:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

([211]) European Parliament, Texts adopted: European Parliament resolution of 9 September 2015 on women's careers in science and universities, and glass ceilings encountered (2014/2251(INI))
https://www.europarl.europa.eu/doceo/document/TA-8-2015-0311_EN.html

([212]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([213]).

Comments and critical issues

The classification of academic positions into A, B, C grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

5.8.7. Proportion of women among grade A staff, by age group

Definition of indicator

This indicator presents the proportion of women among grade A staff across different age groups (less than 35 years, 35–44 years, 45–54 years and 55 years or more) for a given year.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women are lost. Overall, women are under-represented at the higher levels of academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of academic career path as only 23.6 % of heads of higher education institutions and 31% of board members in research decision-making are women and 24.5% of board leaders were women ([214]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([215]).

This indicator sheds light on the representation of women in grade A research positions in different age groups. There are various reasons why this may be of interest. For example, according to Eurostat, a higher proportion of women are outside of the labour force due to caring responsibilities, including for children ([216]). This may reduce their participation in the labour market during the key childbearing years of a particular country.

([213]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([214]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([215]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

([216]) In 2017, in the EU, 37.8 % of women (aged 25 to 49) who were outside of the labour force were in the position due to looking after children or incapacitated adults. For men of the same age group outside of the labour force, the rate was 3.9 %. See Eurostat, 'Inactive population not seeking employment by sex, age and main reason', data table lfsa_igar.

Computation method

Data needed

(F_{AOY}) Number of grade A women in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

(M_{AOY}) Number of grade A men in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among grade A staff in age group O = $\frac{\mathrm{F_{AOY}}}{\mathrm{F_{AOY}+M_{AOY}}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definition:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([217]).

Comments and critical issues

Given that, in some countries, the proportion of academic staff at grade A level is very small in the youngest age group (those aged under 35), this group is not commented on for these countries. The existence of a generational effect could be exemplified by the fact that the proportion of women is larger in the younger age groups. In addition, the classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics.

These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

([217]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

5.8.8. Distribution of grade A staff across age groups, by sex

Definition of indicator

This indicator presents the distribution of women and men grade A staff across age groups for a given year.

Rationale

By looking at the proportion of women present at each grade, one can track their progress in advancing through the stages of the academic career and identify the levels at which women are lost. Overall, women are under-represented at the higher levels of academic career path, with only very small improvements between 2015 and 2018. Indeed, in 2019, despite accounting for almost 54 % of all bachelor's and master's graduates in the EU, women were still severely under-represented at the higher levels of academic career path as only 23.6 % of heads of higher education institutions, 31% of board members in research decision-making are women and 24.5% of board leaders were women ([218]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([219]).

This indicator corrects for the total number of grade A staff for each sex and therefore it allows for a comparison of the presence of each sex across the different age groups.

Computation method

Data needed

(F_{AOY}) Number of grade A women in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

(M_{AOY}) Number of grade A men in age group O (<35 years, 35–44 years, 45–54 years, 55+ years) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of age group O among women grade A staff for year Y = $\frac{\mathrm{F_{AOY}}}{\sum_{\mathrm{O}} \mathrm{F_{AOY}}}$

Proportion of age group O among men grade A staff for year Y = $\frac{\mathrm{M_{AOY}}}{\sum_{\mathrm{O}} \mathrm{M_{AOY}}}$

Specifications

The grades presented in She Figures are based upon national mappings according to the following definition:

A. The single highest grade / post at which research is normally conducted within the institutional or corporate system

B. All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C

([218]) DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021. https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

([219]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020.

C. The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system

D. Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([220]).

Comments and critical issues

The classification of academic positions into A, B, C and D grades may vary across countries. This should be taken into account when comparing or aggregating statistics. These data are not always completely cross-country comparable as the seniority of grades is not yet defined in the same way across countries. Furthermore, it is not always possible to distinguish research staff from teaching staff, although the target population for 'academic staff' is researchers in higher education institutions (excluding staff involved in teaching or administration only and not at all in research).

5.8.9. Proportion of women among heads of institutions in the higher education sector (HES)

Definition of indicator

This indicator looks at the proportion of women among the heads of institutions in the higher education sector (HES) for a given year.

Rationale

The under-representation of women in leadership positions has broad implications for scientific advancement and for industries with a strong need for a technologically educated workforce. An increasing number of science institutions have been adopting in recent years a variety of measures to make improvements ([221]), such as leadership training, implicit bias training, Gender Equality Plans and the Human Resources Strategy for Researchers ([222]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([223]).

This indicator shows the proportion of women in decision-making positions as heads of institutions in the HES.

Computation method

Data needed

(F_Y) Number of women heads of institutions (in the higher education sector) for a given year Y. **Unit: Head count**.

([220]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([221]) Gvozdanović, J. and Maes, K., *Implicit bias in academia: A challenge to the meritocratic principle and to women's careers - And what to do about it*, League of European Research Universities (LERU) Advice Paper No 23, Leuven, 2018. https://www.leru.org/files/implicit-bias-in-academia-full-paper.pdf

([222]) Cameron, I., Synnott, J., Beisiegel, U., O'Carroll, C., Esposito, F., Harrap, K. A., Israel, N., Modjeska, N., Predescu, R., Prijic-Samarzija, S. and Vandevelde, K., *Shaping the future of the Human Resources Strategy for Researchers – HRS4R*, Brussels, 2015. https://cdn1.euraxess.org/sites/default/files/policy_library/experts-report-strengthened-hrs4r-9-2015_0.pdf

([223]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

(M_Y) Number of men heads of institutions (in the higher education sector) for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among heads of institution in the HES = $\frac{F_Y}{F_Y+M_Y}$

Specifications

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([224]).

5.8.10. Proportion of women among heads of universities or assimilated institutions based on capacity to deliver PhDs

Definition of indicator

This indicator looks at the proportion of women among only the heads of universities or assimilated institutions which can deliver PhDs (as opposed to the proportion of women among the heads of institutions in the higher education sector (HES) indicator, which considered all HES institutions), for a given year.

Rationale

The under-representation of women in leadership positions has broad implications for scientific advancement and for industries with a strong need for a technologically educated workforce. An increasing number of science institutions have been adopting in recent years a variety of measures to make improvements ([225]), such as leadership training, implicit bias training, Gender Equality Plans and the Human Resources Strategy for Researchers ([226]). In its Communication on achieving the European Education Area by 2025, the European Commission calls on education and training systems to consider, among other things, 'Working towards a proper gender balance in leadership positions, including in higher education institutions' ([227]).

This under-representation of women in decision-making positions is very well documented with this indicator that measures the proportion of women heads of institutions in the HES. Here, the scope is limited to universities or assimilated institutions based on capacity to deliver PhDs. These differ from general 'institutions in the higher education sector' as the HES sector 'comprises all universities, colleges of technology and other institutions providing formal tertiary education programmes, whatever their source of finance or legal status, and all research institutes, centres, experimental stations and clinics that have their R&D activities under the direct control of, or administered by, tertiary education institutions' ([228]), many of which may not offer PhD programmes.

([224]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([225]) Gvozdanović, J. and Maes, K., *Implicit bias in academia: A challenge to the meritocratic principle and to women's careers - And what to do about it*, League of European Research Universities (LERU) Advice Paper No 23, Leuven, 2018. https://www.leru.org/files/implicit-bias-in-academia-full-paper.pdf

([226]) Cameron, I., Synnott, J., Beisiegel, U., O'Carroll, C., Esposito, F., Harrap, K. A., Israel, N., Modjeska, N., Predescu, R., Prijic-Samarzija, S. and Vandevelde, K., *Shaping the future of the Human Resources Strategy for Researchers – HRS4R*, Brussels, 2015. https://cdn1.euraxess.org/sites/default/files/policy_library/experts-report-strengthened-hrs4r-9-2015_0.pdf

([227]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

([228]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Computation method

Data needed

(F_Y) Number of women heads of universities or assimilated institutions which can deliver PhDs for a given year Y. **Unit: Head count**.

(M_Y) Number of men heads of universities or assimilated institutions which can deliver PhDs for a given year Y. **Unit: Head count**.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women among heads of universities or assimilated institutions = $\frac{F_Y}{F_Y+M_Y}$

Specifications

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([229]).

5.8.11. Proportion of women on boards

Definition of indicator

This indicator presents the proportion of women members (leaders or not) of boards of national umbrella research performing / funding organisations, i.e., members of top decision-making committees that have a crucial impact on the orientation of research in a given year.

Rationale

Since research funding applications are reviewed by scientific boards, the success of women in this process depends on the boards' members that make such decisions, who are often men. It is important to include women in this 'gatekeeping' procedure in order to ensure equal access to funding ([230]); ([231]).

Furthermore, the boards of research organisations have the potential to exercise extensive influence on scientific policy, either through directing core aspects of the agenda or supporting research through an advisory and coordinating role. Given that both advisory and executive boards have considerable decision-making power, the indicator assesses the proportion of women sitting on such boards too to further investigate decision-making by women in academic careers.

Computation method

Data needed

(F_Y) Number of women on boards for a given year Y. **Unit: Head count**.

([229]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([230]) DG Research, *Mapping the maze: Getting more women to the top in research*, Office for Official Publications of the European Communities, Luxembourg, 2008. http://ec.europa.eu/research/science-society/document_library/pdf_06/mapping-the-maze-getting-more-women-to-the-top-in-research_en.pdf

([231]) Bagihole, B., *NORFACE gender equality workshop address*, Rejkjavik, Iceland, 2005.

(M_Y) Number of men on boards for a given year Y. **Unit: Head count**.

The list of boards taken into account is given in the methodological Appendix of the main She Figures publication.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Proportion of women on boards = $\frac{F_Y}{F_Y+M_Y}$

Specifications

To enhance cross-country comparability, Statistical Correspondents were asked to provide data only on the boards of umbrella, national-level research performing organisations (RPOs) and research funding organisations (RFOs). This does not include the boards / councils of individual higher education institutes. Instead, the aim is to capture the highest-level board[s] operating in the country. Umbrella, national organisations which fund industrial research are taken into account only if they also perform / fund public research.

Scientific board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to implement scientific policy by, among other things, directing the research agenda, resource allocation and management within scientific research.

Administrative / advisory board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to support the research agenda in a nonexecutive function by, among other things, administering research activities, consulting and coordinating different actors and taking a general advisory role.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([232]).

Comments and critical issues

No common definition of boards exists and the number of boards varies significantly between countries. It was requested that the metadata submitted should distinguish between boards of organisations performing research and the boards of organisations that are funding research, although both are included in the final computations.

5.8.12. Research funding success rate difference between women and men

Definition of indicator

This indicator presents research funding success-rate differences between research teams led by women and teams led by men. A positive difference means that men-led teams have a higher success rate whereas a negative difference means that women-led teams have a higher success rate.

Rationale

The European Parliament has recognised that 'whereas despite all ongoing efforts to promote gender equality and equal opportunities, women still experience unequal access to research positions, funding,

([232]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

publishing and academic awards, and are also affected by rigid criteria for promotion and recognition and lack of funding or suitable policies to support them' ([233]).

As such, this indicator looks at the differences in the success rate of teams led by men and teams led by women when applying for research funding. The calculation of a success rate rather than the use of raw numbers allows one to normalise for the total number of applications.

Computation method

Data needed

(F_{AY}) Number of female leaders of teams applying for research funding for a given year Y. **Unit: Head count**.

(F_{BY}) Number of female leaders of teams succeeding in receiving research funding for a given year Y. **Unit: Head count**.

(M_{AY}) Number of male leaders of teams applying for research funding for a given year Y. **Unit: Head count**.

(M_{BY}) Number of male leaders of teams succeeding in receiving research funding for a given year Y. **Unit: Head count**.

The list of national research funds taken into account is given in the methodological Appendix of the main She Figures publication.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Success rate difference between men and women = $\frac{\mathrm{M_{BY}}}{\mathrm{M_{AY}}} - \frac{\mathrm{F_{BY}}}{\mathrm{F_{AY}}}$

Specifications

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year([234]).

Comments and critical issues

The WiS data collection also aims to collect data on the sex of all members of research teams applying for and potentially succeeding in receiving funding. However, the collected data display many gaps. For this reason, the indicator is, at present, constrained to team leaders only.

No common definition of funds exists and the total number of funds varies significantly between the countries and over the time period being considered. However, in an attempt to harmonise the data on funds provided by Statistical Correspondents of different countries, it was requested that data should cover all publicly managed research funds (funds granted by institutions in the public sector, excluding private sector funding). Furthermore, Statistical Correspondents were asked to exclude from reporting any funds which allocate funding exclusively on a first-come, first-served basis, i.e., without other selection criteria.

([233]) European Parliament, Texts adopted: European Parliament resolution of 9 September 2015 on women's careers in science and universities, and glass ceilings encountered (2014/2251(INI))
https://www.europarl.europa.eu/doceo/document/TA-8-2015-0311_EN.html

([234]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

5.8.13. Research funding success rate difference between women and men, by field of Research and Development

Definition of indicator

This indicator presents research funding success-rate differences between women-led and men-led teams across different fields of Research and Development. A positive difference means that men-led teams have a higher success rate whereas a negative difference means that women-led teams have a higher success rate.

Rationale

The European Parliament has recognised that 'whereas despite all ongoing efforts to promote gender equality and equal opportunities, women still experience unequal access to research positions, funding, publishing and academic awards, and are also affected by rigid criteria for promotion and recognition and lack of funding or suitable policies to support them' ([235]).

As such, this indicator looks at the differences in the success rate of teams led by men and teams led by women when applying for research funding. The calculation of a success rate rather than the use of raw numbers allows one to normalise for the total number of applications.

Computation method

Data needed

(F_{ASY}) Number of female leaders of teams applying for research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

(F_{BSY}) Number of female leaders of teams succeeding in receiving research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

(M_{ASY}) Number of male leaders of teams applyingfor research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

(M_{BSY}) Number of male leaders of teams succeeding in receiving research funding for a given year Y in a given field of Research and Development S. **Unit: Head count**.

The list of national research funds taken into account is given in the methodological Appendix of the main She Figures publication.

Source of data

DG Research and Innovation – WiS – Women in Science database, with data submitted with the WiS questionnaires

Computation formula

Success rate difference between men and women for field S = $\frac{\mathrm{M_{BSY}}}{\mathrm{M_{ASY}}} - \frac{\mathrm{F_{BSY}}}{\mathrm{F_{ASY}}}$

Specifications

The Frascati Manual (OECD, 2015) provides definitions for the six main fields of Research and Development (p.95). The following abbreviations are used:

([235]) Ibid

- natural sciences (NS)
- engineering and technology (ET)
- medical sciences (MS)
- agricultural and veterinary sciences (AS)
- social sciences (SS)
- humanities (H)
- Two more fields, Unknown (U) and multi-disciplinary (MU), although not fields of R&D in Frascati, they have been added in the WiS questionnaire so that data can also be provided for applications whose field is unknown or which cover more than one specific field respectively.

Head Count (HC) is the total number of individuals contributing to intramural R&D, at the level of a statistical unit or at an aggregate level, during a specific reference period (usually a calendar year) ([236]).

5.9. Scopus™

Content-based rationale

Bibliometric indicators derived from Scopus™ are integrated into the She Figures 2024 publication. These indicators are used to identify differences between women and men regarding several indicators by country, year and Field of Research and Development (FORD). The indicators in this section focus on revealing gender disparities as they relate to author pool (the group of individuals authoring publications), author impact (the publications and citations accrued by authors) and author team composition (the gender composition of research teams contributing as authors on publications).

Additionally, the percentage of a country's research output integrating a gender dimension in its R&I content (GDRIC) is assessed using this data source.

Broad overview of the source

The indicators presented in this section were computed by Elsevier using raw bibliographic data derived from the Scopus ™ database.

Scopus™ is Elsevier's abstract and citation database of peer-reviewed literature, covering 90+ million documents. The data included in She Figures 2024 is based on documents published in over 27,000 active titles (journals, book series and conference proceedings) by over 7,000 publishers.

Its coverage is multi-lingual and global: approximately 46% of its titles are published in languages other than English (or published in both English and another language).

Its coverage is also inclusive across all major research fields, with more than 9,000 titles in the Physical Sciences, more than 7,000 in the Health Sciences, more than 5 000 in the Life Sciences and more than 12,000 in the Social Sciences (the latter including more than 4 900 Arts & Humanities related titles). Titles covered are predominantly serial publications (journals, trade journals, book series and conference material), but considerable numbers of conference papers are also covered from stand-alone proceedings volumes (a major dissemination mechanism, particularly in the Computer Sciences).

For this report, a static version of the database covering the period in scope for the analysis was aggregated by country, region, and field.

([236]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

Publications included in the computation of indicators are limited to peer-reviewed publications, that is, articles, reviews, and conference papers.

In addition to its content coverage, Scopus identifies individual authors and institutions, allowing for a robust analysis at the desired level of granularity whether it is country, state, institution, or author.

Country and regional attribution of each publication is based on the authorship by-line of each article. A publication is attributed to a country if an author indicates affiliation with an institute in that country.

A body of literature is available on the limitations and caveats in the use of 'bibliometric' data, such as the accumulation of citations over time, the skewed distribution of citations across articles and differences in publication and citation practices between fields of research, different languages and applicability to social sciences and humanities research (Bar-Ilan, 2008). In social sciences and humanities, the bibliometric indicators presented in She Figures must be interpreted with caution because a reasonable proportion of research outputs in such fields take the form of books, monographs, and non-textual media. As such, analyses of journal articles, their usage and citation, provide a less comprehensive view for social sciences and humanities than for other fields, where journal articles comprise the vast majority of research outputs.

Bibliometric approaches can only be used as a proxy to assess the productivity and impact of publishing authors. All indicators based on bibliometric approaches can only give insights into the output of authors publishing in Scopus-indexed publications. For this reason, throughout the She Figures handbook, the term 'author' is used to indicate in fact publishing authors. However, a researcher in the sense of the OECD definition is not limited to these authors, and the metrics provided in this section do not assess all researchers, including, for example, researchers in the corporate sector who may not be publishing authors.

Calculation Period

Scopus indexed full names when available for all participating authors and for all publication years dating back to 1996. Author-level indicators are calculated based on authors during the period 2018-2022. Publication-level indicators cover publications over the period 2013 to 2022 and are calculated per annum when possible as well as in aggregate for the full ten/year period.

Included authors

The authors included in the indicator calculation are defined per indicator.

Possible Issues with hypercollaboration

The terms 'hypercollaborative co-authorship' and 'hypercollaboration' have been coined to classify the growing phenomenon of articles that have hundreds or even thousands of co-authors. The rise of so-called 'Big Science' – a term used to describe research that requires major capital investment and is often, but not always, international in nature – may be one of the causes of this phenomenon.

While hypercollaborated publications may represent extreme outliers in co-authorship data and remain proportionally few, such hypercollaborative articles are included throughout the analyses. Like other collaborative articles, they are counted as single internationally co-authored articles for each country represented in them and for each country pairing.

Journal Classification

Journals may be assigned to several major and minor subject areas. Major subject areas are defined according to 27 All Science Journal Classification (ASJC) categories. These are 27 subject categories that all Scopus indexed journals are classified as. Each of the 27 subject categories is further subdivided into a total of 334 minor sub-categories. Because some journals can be classified as multi-category (i.e., more than one subject), each publication may fall into more than one subject classification. For the analyses in She Figures 2024, the ASJC classifications were mapped to the Fields of Research and Development (FORD)

([237]) classifications. A full table of the mapping of FORD classifications with the ASJC sub-categories can be found in Annex 3.

For the calculation of indicators based on FORD classifications, publications are not classified in mutually exclusive fields, with some classified in more than one field. For example, publication P may belong in both Medical Sciences (FORD 3) and Social Sciences (FORD 5). Although publication P will contribute once to the publication count of Medical Sciences and once to the publication count of Social Sciences, this publication will not be counted twice in the aggregated count of 'All' publications.

For some citation-based measures, fractional calculation has been used - distributing publication and citation counts equally across multiple journal categories; publication P would be counted as 0.5 publications for each of Medical Sciences and Social Sciences, and its citations would be shared equally between these subject areas.

Sustainable Development Goal (SDG) Classification

A methodological approach was employed to construct search queries specific to the Sustainable Development Goals (SDGs). Recognising the vast breadth of each SDG, the researchers subdivided each goal into more manageable subtopics likely to further the achievement of the respective SDG. Utilising a bottom-up approach, they built relevant sub-queries based on these subtopics which were then aggregated at the SDG level. A committee of three bibliometricians per SDG defined the subtopics, aiming for a holistic interpretation of the goals. Analysts, leveraging both previously established queries and a selection of specialist journals, constructed the search parameters. Their objective was to attain high precision, evidenced by a benchmark of at least 9 out of 10 papers aligning with the query's intent, even if it meant compromising on recall. These queries underwent rigorous validation by a senior bibliometrician to ensure the capture of most publications pertinent to the SDGs, while identifying and rectifying potential off-topic inclusions and biases.

The full description of the methodology, as well as the open source version of the SDG queries, can be found here: https://elsevier.digitalcommonsdata.com/datasets/9sxdykm8s4/4.

Gender Classification

A binary gender was inferred for author IDs using the NamSor API. The API provides a Gender Probability Score and gender classification based on three data points: country of origin, first name, and last name. These three data points for authors were generated based on information related to each author ID.

- Determination of author country of origin:

Each author's country of origin was determined based on the country of affiliation listed on the publications from their first year of publication in Scopus (i.e., articles, reviews and conference papers). In some cases, authors had published in more than one country in their first year of publication. In these cases, the country with the largest number of publications was designated as the author's country of origin. Authors with equal numbers of publications in two or more countries were excluded from the gender disambiguation analysis. The process used to determine the author country of origin is summarised here.

For each author ID:

1. Identify year of first publication in Scopus.
2. Identify all publications from the first year that the author published.
3. Identify the country affiliation indicated by the author in the first-year publications.
4. Tabulate the countries of affiliation.

([237]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

5. Tabulate the number of times each country was listed as the country of affiliation in the first year of publication.

6. Assign the author ID to the country most often indicated as the country of affiliation in the first year of publication as the country of origin.

7. If two countries appeared as the country of affiliation for an author an equal number of times in the first year of publication, then that author was excluded from the analysis.

- First name determination:

First and last name are required as input data for NamSor. Therefore, only author IDs with a first and last name were passed through the NamSor API to retrieve a Gender Probability Score. All author IDs for whom no first name data was available were not included in the analysis.

Different variants of an author's name are commonly observed across their publications. To identify the best first name to pass through NamSor for each author, all the name variants associated with each author ID were assessed. For each author ID in the Scopus snapshot, all publications on which the author ID appears in the author field were examined and a list of all distinct first names associated with the author ID was generated. Based on this list, a table was generated with a revised first name for each author ID. The process used to determine the best first name to pass through NamSor is described here.

In cases where only a single first name was associated with an author ID:

1. When the name was of zero length, the best first name was assigned as null

   e.g., author first names: []

   best first name: null

These author IDs were excluded from the analysis.

2. When the name was not of zero length, the following nonsensical characters were removed if they were leading or trailing: "-!#&"

   e.g., author first names: ["Tom&"]

   best first name: Tom

In cases where multiple first names were associated with an author ID, the longest available name following removal of nonsensical characters was selected, provided this name was not composed of a string of initials, according to the following steps for each author ID:

1. Author first names were collected into a list.

2. Author first names were initialised as the empty string ' '.

3. The list of names was looped through. Each name in the list was stripped of nonsensical leading and trailing characters (-!#&).

4. The next name in the list was considered and subjected to the same treatment. If the length of the string was longer and not composed of a string of initials, this next name was then assigned as the revised first name. A string of initials among author first names was identified by comparing the number of periods that appear in the string to the number of characters (excluding whitespace, periods and nonsensical characters). When the number of periods was equal to the number of characters, this string was identified as a string of initials.

5. This process was repeated until the end of the list of author first names is reached.

   e.g., author first names: ["Samantha", "#Sam", "Sam", "S. E."]

best first name: Samantha

e.g., author first names: [“Samuel”, “#Sam”, “Sam”, “S. E.”]

best first name: Samuel

An important limitation in using first and last names as indicators for authorship is that women often change their last name after marriage, adopting their partner’s last name. This practice can obscure the continuity of a woman's academic or professional contributions, making it challenging to track career progression, publication records, and overall impact. Such name changes might lead to the loss of visibility for women in databases and bibliometric analyses, impacting the accuracy of gender-disaggregated data.

- Gender Probability Score:

The NamSor Gender Probability Score was used to predict the gender of each author. The Gender Probability Score is the natural log of the ratio of probabilities, as determined by a Naïve-Bayes model, of the name receiving the classification of either woman or man. The Gender Probability Score is based on the best first name, last name and country of origin.

To predict the gender of each author, a table containing unique name-country combinations (represented in the Scopus snapshot) from the following three fields was created:

- best first name (based on the process described in the section “First name determination”)
- last name (based on author ID)
- country of origin (based on the process described in the section “Determination of author country of origin”)

Next the resulting name-country combinations were taken and passed through the NamSor API, which generated a classification (either woman or man) with the associated Probability Calibrated Score for each combination.

The Probability Calibrated score reflects the confidence in the gender assignment returned by NamSor. The score ranges from 0.5 to 1.0. A score of 0.5 indicates high uncertainty that the inferred gender is correct and a score of 1.0 indicates the highest level of certainty that the inferred gender is correct. Previous work with the returned data from the NamSor API indicates that a cutoff Probability Calibrated score of 0.85 would yield a fairly high F1 score for men’s names (0.95) and women’s names (0.93).

While NamSor allows the assignation of gender with high degree of certainty for most researchers included in this study, there are instances in which the probabilities (of a research being from one gender or another) are low, or that it could not be assigned because the full names (given and family names) could not be found in Namsor or because only researchers’ initials are indexed in Scopus. These cases are treated as unknowns and excluded from the computation of the indicators in this project based on the assumption that, for most indicators at least, the group of authors with a gender assigned provides a good representation of those with no gender assigned. In cases in which this assumption may impact the indicators computed in this report, the study team computed margins of error in different scenarios to account for researchers whose gender was not assigned by considering their probabilities

Cut-off date

The cut-off date for Scopus data was 1 September 2023, at which point data for the previous year (2022) can be considered complete.

5.9.1. Proportion of women among active authors, by field of R&D and seniority level, and by selected SDGs

Definition of indicator

This indicator compares the proportion of women who can be classified as "active authors" to the number of all authors who can be classified as "active" within a seniority level, country or region, SDG and field of research.

This indicator is computed in two variants, one accounting for all R&D fields and the other broken down by each R&D field (according to the FORD classification). Additionally, it is broken down by selected SDGs (in particular SDG 5 - Gender equality, SDG 8 - Decent work and economic growth and SDG 12 - Responsible consumption and production).

Rationale

The representation of women and men in the authorship team is partly dependent of the "available workforce" of authors. Assessing the proportion of women to men who can be categorised as "active authors", thereby shedding light on the level of gender balance among the authors within a selected category of country or region, field of research and seniority level can shed light onto metrics related to inclusion on authorship teams. The insights gained from this analysis will inform whether policies or interventions should focus on increasing the participation of women or men as active authors in particular countries or regions, fields and career stages.

Computation method

Data needed

$(\sum FAA_{xci})$ Number of women who are considered active authors (FAA) in a given seniority category (x), country (c) and field or all fields (s) during the period 2018-2022. **Unit: Number.**

$(\sum AAA_{xci})$ Number of all active authors (AAA) in a given seniority category (x), country (c) and field or all fields (s) during the period 2018-2022. **Unit: Number.**

Source of data

Computed using Scopus™ data and NamSor™

Computation formula

Among all authors of peer-reviewed publications considered "active" based on the criteria above, the following indicator is calculated for each seniority level (x), country (c) and FORD (i):

Proportion of women to men amongst active authors: $\frac{(\sum FAA_{xci})}{(\sum AAA_{xci})}$

Specifications

This calculation is based on **active authors** only. Active authors are defined as those that produced 10 or more papers in the last 20 years (2013-2022) and at least 1 paper in the last 5 years (2018-2022) OR those who produced 4 or more papers in last 5 years (2018-2022).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2018-2022).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2013-2017).

- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2012 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2018-2022 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2018-2022 are in that given subject.

5.9.2. Proportion of women among all authors, by field of R&D and selected SDGs, and seniority level

Definition of indicator

This indicator compares the number of women authors to the number of all authors within a seniority level, country or region, selected SDGs and field of R&D.

Rationale

Comparing the proportion of women who have authored publications may reveal the level of gender balance among the authors within a selected category of country or region, field of research and seniority level. This metric therefor provides a point of comparison with metrics related to researcher gender and inclusion of women on author teams. The insights gained from this analysis will inform whether policies or interventions should focus on increasing the participation of women as authors in particular countries or regions, fields and career stages.

Computation method

Data needed

$(\sum FA_{xci})$ Number of women authors who have authored a publication (FA) in a given seniority category (x), country (c) and field (i) during the period 2018-2022. **Unit: Number.**

$(\sum AA_{xci})$ Number of all authors who have authored a publication (AA) in a given seniority category (x), country (c) and field (i) during the period 2018-2022. **Unit: Number.**

Source of data

Computed using Scopus™ data and NamSor™

Computation formula

Among all authors of peer-reviewed publications based on the criteria above, the following indicator is calculated for each seniority level (x), country (c) and FORD (i):

Proportion of women to men amongst active authors: $\frac{(\sum FA_{xci})}{(\sum AA_{xci})}$

Specifications

This calculation is based on **all authors**. Authors are defined as those that produced at least one publication in the last 5 years (2018-2022).

This indicator is broken down by each R&D field (according to the FORD classification), according to the methodology described before.

Additionally, it provides results, broken down, by selected SDGs.

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2018-2022).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2012-2016).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2011 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2018-2022 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2018-2022 are in that given subject.

5.9.3. Ratio of average number of publications by women to those by men in all fields of R&D and by field of R&D, per seniority level

Definition of indicator

This indicator compares the average number of publications by women who can be classified as "active authors" to the average number of publications by men who can be classified as "active authors" within a seniority level, country or region and field of research. A value above 1 indicates that, among active authors, women publish more than men on average; a value below 1 indicates that, among active authors, men publish more than women on average.

Rationale

An author's publication output is sometimes used to assess the productivity of a researcher. Comparing the average publication count of women and men in particular countries or regions, fields and career stages may reveal differences in overall output within those categories. The insights gained from this indicator will inform whether policies based on publication output (such as promotional policies) should be benchmarked against norms within a gender category.

Computation method

Data needed

$(\sum FAA_{xci})$ Number of women who are considered active authors (FAA) in a given seniority category (x), country (c) and field (i) during the period 2018-2022. **Unit: Number.**

$(\sum MAA_{xci})$ Number of men who are considered active authors (MAA) in a given seniority category (x), country (c) and field (i) during the period 2018-2022. **Unit: Number.**

P Number of publications by an author. **Unit: Number.**

Source of data

Computed using Scopus™ data and NamSor™

Computation formula

Among all authors of peer-reviewed publications from a country, the following metrics are computed for each seniority level:

Average number of publications per woman: $\frac{1}{(\sum FAA_{xci})}\sum_{j=1}^{FAA_{xci}}(P)_j$

Average number of publications per man: $\frac{1}{(\sum MAA_{xci})}\sum_{j=1}^{MAA_{xci}}(P)_j$

Ratio of average number of publications by women to those by men : $\frac{\frac{1}{(\sum FAA_{xci})}\sum_{j=1}^{FAA_{xci}}(P)_j}{\frac{1}{(\sum MAA_{xci})}\sum_{j=1}^{MAA_{xci}}(P)_j}$

Specifications

This calculation is based on the publication output of **active authors** only. Active authors are defined as those that produced 10 or more papers in the last 20 years (2012-2021) and at least 1 paper in the last 5 years (2018-2022) OR those who produced 4 or more papers in last 5 years (2018-2022).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2018-2022).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2012-2016).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2011 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2018-2022 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2018-2022 are in that given subject.

5.9.4. Ratio of average FWCI of publications by women to that of men in all fields of R&D and by field of R&D, per seniority level

Definition of indicator

This indicator compares the average citation impact (based on field-weighted citation impact, FWCI) of publications by women who can be classified as "active authors" to the citation impact of publications by men who can be classified as "active authors" within a seniority level, country or region and field of research. A value above 1 indicates that, among active authors, women's publications have a citation impact that is, on average, higher than that of men's; a value below 1 indicates that, among active authors, men's publications have a citation impact that is, on average, higher than that of women's.

Rationale

Comparing the citation impact of men and women may reveal differences in citation behaviours. The insights gained from this analysis inform whether policies based on publication impact (such as promotional policies) should be benchmarked against norms within a gender category.

Computation method

Data needed

$(\sum FAA_{xci})$ Number of women who are considered active authors (FAA) in a given seniority category (x), country (c) and field (i) during the period 2018-2022. **Unit: Number.**

$(\sum FAA_{xci})$ Number of men who are considered active authors (MAA) in a given seniority category (x), country (c) and field (i) during the period 2018-2022. **Unit: Number.**

$\overline{FWCI}$ Mean FWCI of publications by an author. **Unit: Number.**

$$\overline{FWCI} = \frac{1}{N}\sum_{j=1}^{N}\frac{C_j}{E_j}$$

with

C_j = citations received by publication *i*

E_j = expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_j.

Source of data

Computed using Scopus data ™ and NamSor ™

Computation formula

Among all authors of peer-reviewed publications from a country, the following metrics are computed for each seniority level:

Average FWCI for publications by women: $\frac{1}{\sum FAA_{xci}}\sum_{j=1}^{FAA_{xci}}(\overline{FWCI})_j$

Average FWCI for publications by men: $\frac{1}{\sum MAA_{xci}}\sum_{j=1}^{MAA_{ci}}(\overline{FWCI})_j$

Ratio of average FWCI of publications by women to that of men : $\frac{\frac{1}{\sum FAA_{xci}}\sum_{j=1}^{FAA_{xci}}(\overline{FWCI})_j}{\frac{1}{\sum MAA_{xci}}\sum_{j=1}^{MAA_{ci}}(\overline{FWCI})_j}$

Specifications

This calculation is based on the publication output of **active authors** only. Active authors are defined as those that produced 10 or more papers in the last 20 years (2002-2021) and at least 1 paper in the last 5 years (2018-2022) OR those who produced 4 or more papers in last 5 years (2018-2022).

To estimate career stage, the following definitions are applied:

Seniority level is estimated via the time elapsed since an author's first publication in a journal indexed in Scopus and has three categories:

- **<5**: authors whose first paper in Scopus was published in the last 5 years (2018-2022).
- **5 to 10**: authors whose first paper in Scopus was published more than 5 years ago and up to 10 years ago (2012-2016).
- **>10**: authors whose first paper in Scopus is published more than 10 years ago (2011 or prior).

Country designation is attributed to authors based on their publication output. An author counts towards a country's metrics if at least 30% of his/her publications during the period 2018-2022 list the country in the affiliation details.

Subject designation is attributed to authors based on their publication output. An author counts towards a FORD subject's metrics if at least 30% of his/her publications during the period 2018-2022 are in that given subject.

Field-weighted citation impact (FWCI) is an indicator of citation impact of a publication based on the actual number of citations received by an article compared to the expected number of citations for articles of the same document type (article, review or conference proceeding paper), publication year and subject field. When an article is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies overall and over time and document types.

In general, the FWCI for a publication is defined as:

$$FWCI = \frac{C_j}{E_j}$$

where

C_j: citations received by publication i

E_j: expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_i.

5.9.5. Average proportion of women among authors on publications in all fields of R&D, by field of R&D and by selected SDGs

Definition of indicator

This indicator is the average proportion of women among authors on publications from a given country or region and field of research. A value near 50%, indicates that, on average, women and men are represented at equal proportions on teams; a value above 50%, indicates that, on average, women are more highly represented than men on teams; a value below 50% indicates that, on average, men are more highly represented than women on teams.

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in research teams is valuable for ensuring that research outcomes reflect the experience of both men and women. This indicator looks at the contribution of women and men to research teams across countries or regions and fields of research and development (FORD).

Computation method

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum P_{cyi})$ The number of publications in a given country (c), year (y) and field (i). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (c), year (y) and field (i), the formula for average proportion of women among authors on publications (APW) is:

$$(APW)_{CYi} = \frac{1}{\sum P_{cyi}} \sum_{j=1}^{P_{CYi}} \frac{\sum F}{\sum F + \sum M}$$

Specifications

Country attributions of each publication is based on the authorship by-line of each article. A publication is attributed to a country if an author indicates affiliation with an institute in that country.

Subject designation of each publication is based on the source title of each article. A publication is attributed to a subject if the source title has been mapped to a subject in that field.

5.9.6. Compound annual growth rate (CAGR) of average proportion of women among authors on publications, by field of R&D

Definition of indicator

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout, the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the average proportion of women among authors on publications.

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in research teams is valuable for ensuring that research outcomes reflect the experience of both men and women. This indicator looks at the growth rates of the scientific contribution of women on authorship teams across different countries and FORD based on authorship.

Computation method

Data needed

$(APW)_{cyi}$ Average proportion of women among authors on publications in a given country (c), year (y) and field (i). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus data ™ and NamSor ™

Computation formula

CAGR for average proportion of women among authors = $(\mathrm{APW_{Cei}}/\mathrm{APW_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{APW_{Csi}}$ denotes the average proportion of women among authors in the start year;

$\mathrm{APW_{Cei}}$ denotes the average proportion of women among authors in the end year.

5.9.7. Average proportion of women among authors on publications resulting from international collaborations in all fields of R&D, and by field of R&D

Definition of indicator

This indicator is the average proportion of women among authors on publications resulting from international collaboration. A value near 50% indicates that, on average, women and men are represented at equal proportions on international authorship teams; a value above 50% indicates that, on average, women are more highly represented than men on international authorship teams; a value below 50%, indicates that, on average, men are more highly represented than women on international authorship teams.

Rationale

Women are less likely than men to collaborate internationally on research papers ([238]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([239]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men.

This indicator looks at the average proportion of women among authorship teams resulting from internationally co-authored publications across different countries and fields of research and development (FORD).

Computation method

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PI_{cyi})$ The number of publications resulting from international collaboration in a given country (*c*), year (*y*) and field (*i*). **Unit: Number.**

([238]) Elsevier, *Gender in the Global Research Landscape*, 2017.
https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([239]) Jappelli, T., Nappi, C.A., Torrini, R., '*Gender effects in research evaluation*', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Source of data

Computed using Scopus data ™ and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and field (*i*), the formula for average proportion of women among authors on internationally collaborated publications (AIPW) is:

$$(AIPW)_{cyi} = \frac{1}{\sum PI_{cyi}} \sum_{i=1}^{PI_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (*c*), year (*y*) and FORD (*i*).

Specifications

International collaboration is defined as multi-authored research outputs, where at least one author is from an institution inside the country of interest and at least one author is from an institution outside the country of interest (or EU for EU-27).

5.9.8. Compound annual growth rate (CAGR) of average proportion of women among authors on publications resulting from international collaborations, by field of R&D

Definition of indicator

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout, the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the average proportion of women among authors on publications resulting from international collaboration across different countries and fields of research and development (FORD).

Rationale

Women publish fewer research papers on average than men and women are less likely than men to collaborate internationally on research papers ([240]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([241]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator looks at the change in the average proportion of women among authors on publications resulting from international collaboration.

([240]) Elsevier, *Gender in the Global Research Landscape*, 2017.
https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([241]) Jappelli, T., Nappi, C.A., Torrini, R., 'Gender effects in research evaluation', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Computation method

Data needed

$(AIPW)_{cyi}$ Average proportion of women among authors on publications resulting from international collaboration in a given country (*c*), year (*y*) and field (*i*). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

CAGR for average proportion of women among authors = $(\mathrm{AIPW_{Cei}}/\mathrm{AIPW_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{AIPW_{Csi}}$ denotes the average proportion of women authors on publications resulting from international collaboration in a given country (*c*) and field (*i*) in the start year;

$\mathrm{AIPW_{CeS}}$ denotes the average proportion of women authors on publications resulting from international collaboration in a given country (*c*) and field (*i*) in the end year.

5.9.9. Average proportion of women among authors on publications resulting from national collaboration in all fields of R&D

Definition of indicator

This indicator is the average proportion of women among authors on publications resulting from national collaboration. A value near 50% indicates that, on average, women and men are represented at equal proportions on national authorship teams; a value above 50% indicates that, on average, women are more highly represented than men on national authorship teams; a value below 50% indicates that, on average, men are more highly represented than women on national authorship teams.

Rationale

Women are less likely than men to collaborate internationally on research papers ([242]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([243]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to national collaboration can provide insight into whether there are barriers to authors collaborating outside of their own institution, which may be a first step in a researcher's path towards international collaboration.

([242]) Elsevier, *Gender in the Global Research Landscape*, 2017.
https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([243]) Jappelli, T., Nappi, C.A., Torrini, R., 'Gender effects in research evaluation', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

This indicator looks at the average proportion of women among authorship teams resulting from nationally co-authored publications across different countries and fields of research and development (FORD).

Computation method

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PN_{cyi})$ The number of publications resulting from national collaboration in a given country (*c*), year (*y*) and field (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and field (*i*), the formula for average proportion of women among authors on internationally collaborated publications (ANPW) is:

$$(ANPW)_{cyi} = \frac{1}{\sum PN_{cyi}} \sum_{j=1}^{PN_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (*c*), year (*y*) and FORD (*i*).

Specifications

National collaboration is defined as multi-authored research outputs, where authors are affiliated with more than one institution within the same country.

5.9.10. Average proportion of women among authors on publications resulting from intra-EU-27+ collaborations in all fields of R&D

Definition of indicator

This indicator is the average proportion of women among authors on publications resulting from intra-EU-27+ collaboration. A value near 50% indicates that, on average, women and men are represented at equal proportions on EU-27+ authorship teams; a value above 50% indicates that, on average, women are more highly represented than men on EU-27+ authorship teams; a value below 50% indicates that, on average, men are more highly represented than women on EU-27+ authorship teams.

Rationale

Women are less likely than men to collaborate internationally on research papers ([244]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([245]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention

([244]) Elsevier, *Gender in the Global Research Landscape*, 2017.
https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([245]) Jappelli, T., Nappi, C.A., Torrini, R., 'Gender effects in research evaluation', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

and impact as men. Statistics related to intra-EU-27+ collaboration can provide insight into whether there are barriers to authors collaborating outside of their own institution.

This indicator looks at the average proportion of women among authorship teams resulting from intra-EU-27+ co-authored publications across different countries and fields of research and development (FORD).

Computation method

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PIE_{cyi})$ The number of publications resulting from intra-EU27+ collaboration in a given country (*c*), year (*y*) and field (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and field (*i*), the formula for average proportion of women among authors on internationally collaborated publications (AIEPW) is:

$$(AIEPW)_{cyi} = \frac{1}{\sum PIE_{cyi}} \sum_{j=1}^{PIE_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (*c*), year (*y*) and FORD (*i*).

Specifications

Intra-EU-27+ collaboration is defined as multi-authored research outputs, where authors are affiliated with institutions in more than one of the 27 EU countries but all authors are based within the 27 countries.

5.9.11. Ratio of fractional FWCI for women to men authors on publications in all fields of R&D and by field of R&D

Definition of indicator

This indicator compares the fractional citation impact (based on field-weighted citation impact, FWCI) of publications by women to that of men on a publication by publication basis. Fractional calculation refers to distributing publication and citation counts equally across multiple authors; publication P with two authors would be counted as 0.5 publications for each author, and its citations would be shared equally between these authors. A value above 1 indicates that women are more represented on highly cited publications; a value below 1 indicates that men are more represented on highly cited publications. The ratio is based on the fractional FWCI for all women and men on a publication.

Rationale

Women publish fewer research papers on average than men ([246]) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles ([247]) and this may affect the impact of their publications. Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([248]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men.

This indicator looks at the contribution of women and men authors to the mean FWCI in a given country and subject.

Computation method

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

$(\sum M)$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum A)$ Number of gendered authors in authorship byline. **Unit: Number.**

$(\sum P_{cyi})$ The number of publications in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(FWCI)$ FWCI for given publication

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), field (*i*) and year (*y*), the formula for ratio of FWCI for women to men based on fractional authorship (RFWtMFA) is:

$$fractional\ FWCI\ for\ women\ \ authors = \frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum F / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum F / \sum A)}$$

$$fractional\ FWCI\ for\ men\ \ authors = \frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum M / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum M / \sum A)}$$

$$(RFWtMFA) = \frac{\frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum F / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum F / \sum A)}}{\frac{\sum_{j=1}^{P_{cyi}}\left((FWCI) * (\sum M / \sum A)\right)}{\sum_{j=1}^{P_{cyi}}(\sum M / \sum A)}}$$

([246]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([247]) Larivière, V., Ni, C., Gingras, Y., Cronin, B. and Sugimoto, C. R., 'Global gender disparities in science', *Nature*, No. 504, 2013, pp. 211–213.

([248]) Jappelli, T., Nappi, C.A., Torrini, R., 'Gender effects in research evaluation', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Specifications

FWCI is an indicator of citation impact of a publication based on the actual number of citations received by an article compared to the expected number of citations for articles of the same document type (article, review or conference proceeding paper), publication year and subject field. When an article is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies overall and over time and document types.

In general, the FWCI for a publication is defined as:

$$FWCI = \frac{C_j}{E_j}$$

where

C_j: citations received by publication i

E_j: expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_i.

5.9.12. Compound annual growth rate (CAGR) of ratio of fractional FWCI for women to men

Definition of indicator

This indicator presents the compound annual growth rate of the ratio of FWCI for women to men based on fractional authorship, meaning the average yearly percentage increase/decrease, year on year.

Rationale

Women publish fewer research papers on average than men ([249]) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles ([250]) and this may affect the impact of their publications. Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([251]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator assesses whether there has been any change in.

([249]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([250]) Larivière, V., Ni, C., Gingras, Y., Cronin, B. and Sugimoto, C. R., 'Global gender disparities in science', *Nature*, No. 504, 2013, pp. 211–213.

([251]) Jappelli, T., Nappi, C.A., Torrini, R., 'Gender effects in research evaluation', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Computation method

Data needed

$(RFWtMFA)_{cyi}$ Ratio of FWCI for women to men based on fractional authorship in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

CAGR for ratio of FWCI for women to men based on fractional authorship = $((RFWtMFA)_{cei}/(RFWtMFA)_{csi})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$RFWtMFA_{\mathrm{Csi}}$ denotes the ratio of FWCI for women to men based on fractional authorship in a given country (*c*) and FORD (*i*) in the start year;

$RFWtMFA_{\mathrm{Cei}}$ denotes the ratio of FWCI for women to men based on fractional authorship in a given country (*c*) and field (*i*) in the end year.

5.9.13. Proportion of women with corresponding authorship in all fields of R&D, by field of R&D and by selected SDGs

Definition of indicator

This indicator is the percentage of publications in which a woman is the corresponding author. It is based on peer-reviewed scientific publications (articles, reviews, conference papers).

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women.

This indicator looks at the contribution of women as the corresponding author to research across countries and fields of research and development (FORD).

Computation method

Data needed

$(\sum PCF_{cyi})$ The number of publications with a woman as the corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(\sum PCA_{cyi})$ The number of publications with a corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and field (*i*), the formula for the indicator, Proportion of Women Corresponding authorship (*PC*) is:

$$(PC)_{cyi} = \frac{\sum PCF_{cyi}}{\sum PCA_{cyi}}$$

5.9.14. Compound annual growth rate (CAGR) of proportion of women with corresponding authorships

Definition of indicator

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout, the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the proportion of women authors out of all corresponding authors.

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. This indicator looks at the growth rates of the scientific contribution of women as corresponding author across different countries and fields of research and development (FORD).

Computation method

Data needed

$(PC)_{cyi}$ Proportion of women authors to all corresponding authors in a start and an end year (*y*) in a given country (*c*) and FORD (*i*). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

CAGR for ratio of corresponding authorships of women to men = $(\mathrm{PC_{Cei}}/\mathrm{PC_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

PC_{Csi} denotes the proportion of women authors to all of corresponding authors in a given country (*c*) and FORD (*i*) in the start year;

PC_{Cei} denotes the proportion of women authors to all of corresponding authors in a given country (*c*) and FORD (*i*) in the end year.

5.9.15. Proportion of women with corresponding authorships in international collaborations, in all fields of R&D and by field of R&D

Definition of indicator

This indicator is the proportion of publications resulting from international collaboration in which a woman is the corresponding author.

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. Women are less likely than men to lead internationally collaborated research ([252]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([253]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men.

This indicator looks at the contribution of women as the corresponding author on publications resulting from international collaboration, to research across countries and fields of research and development (FORD).

Computation method

Data needed

$(\sum PICF_{cyi})$ The number of publications resulting from international collaboration with a woman as corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(\sum PICA_{cyi})$ The number of publications resulting from international collaboration with a corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and FORD (*i*), the formula for the indicator, proportion of women acting as corresponding authors (*PCI*) is:

$$(PCI)_{cyi} = \frac{\sum PICF_{cyi}}{\sum PICA_{cyi}}$$

([252]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([253]) Jappelli, T., Nappi, C.A., Torrini, R., 'Gender effects in research evaluation', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Specifications

International collaboration is defined as multi-authored research outputs, where at least one author is from an institution inside the country of interest and at least one author is from an institution outside the country of interest (or the EU, for EU 27 calculations).

5.9.16. Proportion of women with corresponding authorships in national collaborations in all fields of R&D and by field of R&D

Definition of indicator

This indicator is the proportion of publications resulting from national collaboration in which a woman is the corresponding author.

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. Women are less likely than men to lead internationally collaborated research ([254]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([255]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to national collaboration can provide insight into whether there are barriers to authors collaborating outside of their own institution.

This indicator looks at the contribution of women as corresponding authors on publications resulting from national collaboration, to research across countries and fields of research and development (FORD).

Computation method

Data needed

$(\sum PNCF_{cyi})$ The number of publications resulting from national collaboration with a woman as corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(\sum PNCA_{cyi})$ The number of publications resulting from national collaboration with a corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and FORD (*i*), the formula for the indicator Proportion of women among corresponding authorships resulting from national collaborations (*PCN*) is: $(PCN)_{cyi} = \frac{\sum PNCF_{cyi}}{\sum PNCA_{cyi}}$

([254]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([255]) Jappelli, T., Nappi, C.A., Torrini, R., '*Gender effects in research evaluation*', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Specifications

National collaboration is defined as multi-authored research outputs, where authors are affiliated with more than one institution within the same country.

5.9.17. Proportion of women with corresponding authorships in all fields of R&D in intra-EU-27+ collaborations

Definition of indicator

This indicator is the proportion of publications resulting from intra-EU-27+ collaborations in which a woman is the corresponding author.

Rationale

Representation of diverse viewpoints can impact how research questions are formulated and answered. Therefore, gender parity in the contributions of both women and men in research output is valuable for ensuring that research outcomes reflect the experience of both men and women. Women are less likely than men to lead internationally collaborated research ([256]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([257]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at the same rates and gather similar attention and impact as men. Statistics related to intra-EU27+ collaboration can provide insight into whether there are barriers to authors collaborating outside of their own country.

This indicator looks at the contribution of women as corresponding authors on publications resulting from intra-EU27+ collaboration, to research across countries and fields of science.

Computation method

Data needed

$(\sum PIECF_{cyi})$ The number of publications resulting from Intra-EU27+ collaboration with a woman as corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(\sum PIECA_{cyi})$ The number of publications resulting from Intra-EU27+ collaboration with a corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and FORD (*i*), the formula for the indicator, Proportion of women among corresponding authorships, intra-EU-27+ collaborations (*PCIE*) is: $(PCIE)_{cyi} = \frac{\sum PIECF_{cyi}}{\sum PIECA_{cyi}}$

Specifications

Intra-EU27+ collaboration is defined as multi-authored research outputs, where authors are affiliated with institutions in more than one of the 27 EU countries but all authors are based within the 27 countries.

([256]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([257]) Jappelli, T., Nappi, C.A., Torrini, R., '*Gender effects in research evaluation*', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

5.9.18. Ratio of average FWCI for publications with women as corresponding authors to average FWCI for publications with men as corresponding authors, in all fields of R&D, and by field of R&D

Definition of indicator

This indicator compares the average citation impact (based on field-weighted citation impact, FWCI) of publications with women as corresponding authors to that of men. A value above 1 indicates that publications with women as corresponding authors are more highly cited; a value below 1 indicates that publications with men as corresponding authors are more highly cited.

Rationale

Women publish fewer research papers on average than men ([258]) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles ([259]) and this may affect the impact of their publications. Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([260]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men.

This indicator looks at the citations received by publications with women and men corresponding authors in a given country and subject.

Computation method

Data needed

$(\sum PCF_{cyi})$ Number of publications with a woman as corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(\sum PCM_{cyi})$ Number of publications with a man as corresponding author in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

$(FWCI)$ FWCI for a given publication

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

For a given country (*c*), year (*y*) and FORD (*i*), the formula for the ratio of FWCI for women to men based on corresponding authorship (*RFWtMCA*) is:

$$FWCI\ for\ publications\ with\ women\ as\ corresponding\ authors = \frac{1}{\sum PCF_{cyi}} \sum_{j=1}^{PCF_{cyi}} (FWCI)$$

$$FWCI\ for\ publications\ with\ men\ as\ corresponding\ authors = \frac{1}{\sum PCM_{cyi}} \sum_{j=1}^{PCM_{cyi}} (FWCI)$$

([258]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([259]) Larivière, V., Ni, C., Gingras, Y., Cronin, B. and Sugimoto, C. R., '*Global gender disparities in science*', *Nature*, No. 504, 2013, pp. 211–213.

([260]) Jappelli, T., Nappi, C.A., Torrini, R., '*Gender effects in research evaluation*', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

$$(RFWtMCA) = \frac{\frac{1}{\sum PCF_{cyi}} \sum_{j=1}^{PCF_{cyi}} (FWCI)}{\frac{1}{\sum PCM_{cyi}} \sum_{j=1}^{PCM_{cyi}} (FWCI)}$$

Specifications

FWCI is an indicator of citation impact of a publication based on the actual number of citations received by an article compared to the expected number of citations for articles of the same document type (article, review or conference proceeding paper), publication year and subject field. When an article is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies overall and over time and document types.

In general, the Field-Weighted Citation Impact (FWCI) for a publication is defined as:

$$FWCI = \frac{C_j}{E_j}$$

Where

C_j: citations received by publication i

E_j: expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than one discipline, the harmonic mean is used to calculate E_i.

5.9.19. Compound annual growth rate (CAGR) of ratio of average FWCI for publications with women as corresponding authors to average FWCI for publications with men as corresponding authors

Definition of indicator

This indicator presents the compound annual growth rate of the ratio of FWCI for women to men based on corresponding authorship, meaning the average yearly percentage increase/decrease, year-on-year.

Rationale

Women publish fewer research papers on average than men ([261]) and women in a leading role of authorship receive fewer citations than in cases when a man was in one of these roles ([262]) and this may affect the impact of their publications. Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([263]), disadvantaging women in grant competitions with their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator assesses whether there has been any change in the citations received by publications with women and men corresponding authors in a given country and subject.

([261]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([262]) Larivière, V., Ni, C., Gingras, Y., Cronin, B. and Sugimoto, C. R., '*Global gender disparities in science*', *Nature*, No. 504, 2013, pp. 211–213.

([263]) Jappelli, T., Nappi, C.A., Torrini, R., '*Gender effects in research evaluation*', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

Computation method

Data needed

$(RFWtMFA)_{cyi}$ Ratio of FWCI for women to men based on fractional authorship in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus ™ data and NamSor ™.

Computation formula

CAGR for ratio of FWCI for women to men based on fractional authorship = $(RFWtMCA_{\mathrm{Cei}}/RFWtMCA_{\mathrm{Csi}})^{1/\mathrm{N}} - 1$

where

s refers to the start year

e refers to the end year

$RFWtMCA_{\mathrm{Csi}}$ denotes the ratio of FWCI for women to men based on corresponding authorship in a given country (*c*) and FORD (*i*) in the start year

$RFWtMCA_{\mathrm{Cei}}$ denotes the ratio of FWCI for women to men based on corresponding authorship in a given country (*c*) and FORD (*i*) in the end year.

5.9.20. Average proportion of women among authors on publications that list among the author affiliations, both a corporate entity and any other entity, in all fields of R&D and by field of R&D

Definition of indicator

This indicator is the average proportion of women among authors on publications resulting from collaboration with a corporate entity. It is based on peer-reviewed scientific publications (articles, reviews, conference papers).

Rationale

This indicator mirrors indicator 3.2.2 “Public-private co-publications” in the Innovation Union Scoreboard ([264]). It is directly relevant to the EU’s Innovation Union Scoreboard and Open Innovation objectives, and holds relevance for Digital Economy and Society policy objectives & Open Innovation.

Computation method

Data needed

$(\sum F)$ Number of women authors in authorship byline. **Unit: Number.**

([264]) The scoreboard can be found here: https://research-and-innovation.ec.europa.eu/statistics/performance-indicators/european-innovation-scoreboard_en

$(\sum M))$ Number of men authors in authorship byline. **Unit: Number.**

$(\sum PC_{cys})$ The number of publications resulting from corporate collaboration in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number.**

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula a given country (*c*), year (*y*) and FORD (*i*), the formula for average proportion of women among authors on corporate collaborated publications (*ACPW)* is:

$$(ACPW)_{cyi} = \frac{1}{\sum PC_{cyi}} \sum_{j=1}^{PI_{cyi}} \frac{\sum F}{\sum F + \sum M}$$

The indicator is calculated for each country (*c*), year (*y*) and FORD (*i*).

Specifications

Corporate collaboration is defined as publications that list among the author affiliations both a corporate entity and any other entity (academic, government, or medical).

5.9.21. Proportion of a country's publications with a gender dimension in its R&I content (GDRIC) in all fields of R&D and by field of R&D

Definition of indicator

The indicator shows the proportion of peer-reviewed publications that integrate gender-sensitive analysis and the impact of these publications, broken down by field and country.

Rationale

The European Commission seeks to promote the integration of the methods of sex- and/or gender-based analysis into research design and process as a way of preventing bias in research, promoting better quality of outcomes in R&I and achieving cross-cutting benefits. To assess whether the introduction of new policies has resulted in changes in the research landscape, this indicator will reveal changes over time in how much research:

- Explicitly addresses gender-related issues, recognising the diverse experiences and needs of different gender groups;
- Considers the sex characteristics and implications for all sexes, ensuring that findings are relevant and applicable.

The above categories are by no means meant to be mutually exclusive. This assessment includes studies on non-human species, which can be models to study human conditions.

Bibliometric analyses are used to identify the body of research that integrates a 'gender dimension'. To strengthen the integration of sex and gender analysis in research and innovation, the European Commission established the Gendered Innovations expert group that produced reports on incorporating the gender dimension across various research fields, defining research integrating the 'gender dimension' as research that integrates sex and gender analysis into their content. 'Sex' refers to basic biological characteristics of

females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge ([265]) ([266]).

The strengthening of the integration of the gender dimension into R&I content is one of the gender equality priorities set for Horizon Europe, the EU Framework Programme for Research and Innovation 2021-2027. Under Horizon Europe, the integration of a gender dimension into research and innovation content is a requirement by default, and evaluated under the excellence criterion, unless the topic description explicitly specifies otherwise.

Computation method

Data needed

$\mathbf{P}_{cyi}$ Number of publications in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number**

$\mathbf{P}GDRIC_{cyi}$ Number of publications integrating GDRIC in a given country (*c*), year (*y*) and FORD (*i*). **Unit: Number**

Source of data

Computed using Scopus™ data

Computation formula

For a given country (*c*), year (*y*) and FORD (*i*), the formula for this indicator is:

Percent of a country's publications integrating GDRIC$_{cyi}$ = $\frac{\mathbf{P}GDRIC_{cyi}}{\mathbf{P}_{cyi}}$

The aggregation over fields or countries is implicitly carried out by extending the range of fields or countries respectively over which the sums in the numerator and denominator extend. Even if a paper is assigned to more than one field covered in the indicator, the algorithm counts it only once.

Specifications

The bibliometric analysis is based on the following strategy:

1. Target research that mentions at least two sexes or genders (researchers often use 'gender' when in fact they mean 'sex') in the context of the abstract or title. This approach ensures identification of research in which the researchers have made some effort to compare characteristics/behaviour of women and men by including more than single sex or gender in the research rationale or output and, therefore, the findings can be influenced by sex or gender-based variables.

2. Include research on non-human species in the search. Whilst discussions of the methods of analysis of gender dimension in research have so far been focused on human context, it has been recognised that sex, in particular, can play an important role in controlling the 'lifepath' of non-human species (for example plants, animals or cells) that make up the natural ecosystems in which humans coexist and whose wellbeing they influence. Therefore, and especially in the context of societal and environmental challenges, such as those linked to the effects of climate change, it is important to also promote methods of sex/gender analysis in research more widely.

3. Exclude research which applies to only a single-sex or gender as a recognition that while some research may apply to only single-sex out of necessity (for example – pregnancy or prostate research), a large body of research in animals contains persistent historical male bias, thereby

([265]) European Commission, *Gendered Innovations 2: How Inclusive Analysis contributes to Research and Innovation*, 2020. https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en

([266]) European Commission, *Gendered Innovations – How gender analysis contributes to research: report of the expert group 'Innovation through gender'*, 2013. https://op.europa.eu/en/publication-detail/-/publication/d15a85d6-cd2d-4fbc-b998-42e53a73a449

excluding females and biasing results ([267]). This approach avoids including this latter body of animal research in which gender dimension has not been considered in the research design.

For the categories listed above, that is, to assess research that accounts for both sexes in study design as a proxy for GDRIC, publications were identified that mention at least two sexes or genders in the context of the abstract or title. This approach has been chosen as a means to identify research in which the researchers have made some effort to compare characteristics/behaviour of women/females and men/males by including more than a single-sex or gender in the research rationale or output. Research on non-human species is included in the search.

Creation of datasets for gender dimension in R&I content

The publications included in the assessment were retrieved using any of the following eight queries:

The

i	(({men} OR {man} OR boy* OR {male} or mascul*) AND ({women} OR {woman} OR girl* OR {female} OR femin*))
ii	{sexual dimorphism}
iii	(sexual* AND dimorph*)
iv	“gender dimension”
v	“gender difference”
vi	{sex ratio}
vii	“sex difference”
viii	“sexual reproduction”

assessment of the validity of each query used to create the publication set was done with input from gender experts selected by Portia Inc. These experts validated a set of 100 randomly selected publications retrieved by each query (eight queries in total) to generate a false positive rate. The false positive rate for each of the eight queries was less than 1%. Therefore, all eight queries were used in union to generate the publication set.

5.9.22. Distribution of publications by sex composition of the authors team

Definition of indicator

This indicator analyses the gender composition of the authors’ team for each publication (e.g., by teams consisting of women only). It shows the proportions of each type of composition out of all teams.

Rationale

Each publication can have one or multiple authors (working collaboratively). Determining the gender of each author allows the identification of mutually exclusive sets of publications, i.e., those referred to as working-alone women or men), those developed by teams of the same gender and those referred to as mixed-gender teams.

The indicator sheds light on the propensity of women and men to work alone or in same-gender teams versus working in mixed-gender teams as well as how such collaboration patterns vary between countries and evolve over time.

([267]) Berry, A.K. and Zucker, I., ‘Sex bias in neuroscience and biomedical research’, *Neuroscience & Biobehavioural Reviews*, Vol. 35, No. 3, 2011, pp. 565-572.

Computation method

Data needed

(WI_{cy}) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from women authors working alone. **Unit: Count**.

(MI_{cy}) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from men authors working alone. **Unit: Count**.

($allWI_{cy}$) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from teams of women authors. **Unit: Count**.

($allMI_{cy}$) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from teams of men authors. **Unit: Count**.

(pWI_{cy}) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from predominantly women teams. **Unit: Count**.

(pMI_{cy}) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from predominantly men teams. **Unit: Count**.

(pTI_{cy}) Number of publications in Scopus attributed to the country (*c*) and year (*y*) from sex-balanced teams. **Unit: Count**.

The ratios were computed using **a four-year period** (2018-2022).

Computation formula

Once the genders of its authors have been identified (see the general introduction to this section), it is possible to compute the following variables for each *y* and *c*:

- number of publications from women/men authors working alone: WI_{cy} or MI_{cy}
- number of publications from teams with authors of the same gender: $allWI_{cy}$ or $allMI_{cy}$
- number of publications from teams which consist predominantly of women: pWI_{cy}. These are teams with more than 60% women.
- number of publications from teams which consist predominantly of men: pMI_{cy}. These are teams with more than 60% men.
- number of publications from gender-balanced teams: bTI_{cy}. These are teams with between 40% and 60% women.

All the above situations are mutually exclusive, so that their sum corresponds to the total number of publications for each year (*y*) and country (*c*):

$$WI_{cy} + MI_{cy} + allWI_{cy} + allMI_{cy} + pWI_{cy} + pMI_{cy} + bTI_{cy} = TOT_{cy}$$

All the indicators were referred to TOT_{cy}; in particular:

- proportion of publications from women/men authors working alone: WI_{cy}/TOT_{cy} or MI_{cy}/TOT_{cy}
- proportion of publications from teams with authors of the same gender: $allWI_{cy}/TOT_{cy}$ or $allMI_{cy}/TOT_{cy}$
- proportion of publications from teams consisting predominantly of women: pWI_{cy}/TOT_{cy}. These are teams with more than 60% women.

- proportion of publications from teams consisting predominantly of men: pMI_{cy}/TOT_{cy}. These are teams with more than 60% men.

- proportion of publications from gender-balanced teams: $bTI_{cy/}/TOT_{cy}$. These are teams with between 40% and 60% women.

5.9.23. Ratio of internationally mobile women compared to men

Definition of indicator

This indicator represents the proportion of women and men who have been mobile, based on the affiliation history detected in published papers. A ratio above 1.0 would be indicative of women on average being more mobile than men and a ratio below 1.0 would be indicative of men on average being more mobile than women.

Rationale

Internationally mobile researchers tend to have a higher citation impact than those who are not. If women are less mobile than men, it may have implications for the impact of their scholarly output. Impact (measured by output and citations) may be used for assessment purposes and therefore may have an influence on funding decisions, employment options and – in general – on the existence of equal possibilities and opportunities. This indicator would serve a useful complement to the existing indicators on sex differences in international mobility of researchers during their PhDs and in post-PhD career stages.

Computation method

This indicator uses Scopus author profile data to derive a history of active author affiliations recorded in published papers. The study team defines as “mobile” those authors having published with affiliations in at least two countries. Then, the proportion of women author profiles out of all mobile author profiles is computed. To compute this indicator, we can follow these steps, along with the corresponding mathematical notation that serves as an exemplification of the indicator, capturing the essence of the concept:

Step 1: Define the sets of active authors.

- Let A_W be the set of active women authors.

- Let A_M be the set of active men authors.

Step 2: Identify mobile authors.

- For the sake of exemplification, define a function M(a) that returns 1 if an author 'a' has published with affiliations in at least two countries, and 0 otherwise.

Step 3: Calculate the number of mobile women and men authors.

- N_W = sum(M(a_w) for a_w in A_W)

- N_M = sum(M(a_m) for a_m in A_M)

Step 4: Compute the proportion of women and men authors among all mobile authors.

- P_W = N_W / (N_W + N_M)

- P_M = N_M / (N_W + N_M)

Step 5: Calculate the ratio of mobile women authors to mobile men authors.

- R = P_W / P_M

As mentioned in the description, the computed ratio R represents the sex differences in international mobility among active authors. If R > 1.0, it indicates that women authors are, on average, more mobile than men authors. If R < 1.0, it suggests that men authors are more mobile than women authors. If R is equal to 1.0, the mobility of women and men authors is balanced.

The provided mathematical notation exemplifies the computation of the indicator, but the actual implementation may be subject to variations depending on the data structure, data distribution and computational considerations.

5.9.24. Relative activity index for research contributing to goal SDG 5

Definition of indicator

This indicator measures the percentage of publications from each country that contribute to advancement of the SDG 5 goal, its targets and indicators relative to the percentage of publications that contribute to advancement of SDG 5 globally.

Rationale

This indicator assesses the extent to which individual countries contribute to gender-focused research in comparison to global contributions. By analysing the disparity between national and global research outputs on SDG 5, which promotes gender equality, the indicator highlights which countries are leading (or lagging) in this area. This not only offers insights into the academic commitment to gender issues but also points areas that need attention. The methodology used to determine the SDGs is outlined above and the related full description, as well as the open source version of the SDG queries, can be found at the following link: https://elsevier.digitalcommonsdata.com/datasets/9sxdykm8s4/4.

There are numerous studies examining the Sustainable Development Goals (SDGs), including SDG 5. Outcomes from these studies can differ based on the methodology employed. Different research designs and analytical techniques can yield varied conclusions. The methodology used here, which has been previously applied in other European Union projects, strives for rigor, consistency and transparency. It offers a robust framework for assessing individual country contributions to gender-focused research, though, like all methodologies, it has its inherent limitations.

Computation method

This indicator is based on Elsevier's SDG mapping, for the year range (2018-2022) and document types in scope. To compute this indicator, the following steps are performed, along with the corresponding mathematical notation that serves as an exemplification of the indicator, capturing the essence of the concept:

Step 1: Define the sets of publications,

P_i^{SDG5} $publications\ contributing\ to\ SDG5\ from\ Country\ i$

P_i $all\ publications\ of\ country\ i$

P^{SDG5} $all\ publications\ contributing\ to\ SDG5\ on\ a\ global\ scale$

P $overall\ publications\ on\ a\ global\ scale$

Step 2: Calculate the relative activity index for research contributing to SDG 5. $RAI_i^{SDG5} = \frac{\frac{P_i^{SDG5}}{P_i}}{\frac{P^{SDG5}}{P}}$ $relative\ activity\ index\ of\ country\ i\ on\ SDG5$.

The computed relative activity index (RAI) represents the ratio of a country's percentage of SDG 5-contributing publications to the global percentage of SDG 5-contributing publications. If RAI > 1.0, it indicates that the specific country is contributing more to SDG 5-focused research compared to the global average. If RAI < 1.0, it suggests that the country is contributing less to SDG 5-focused research compared to the global average. If RAI is equal to 1.0, the country's contribution is in line with the global average.

The provided mathematical notation exemplifies the computation of the indicator. However, the actual implementation may be subject to variations depending on the data structure, data distribution and

computational considerations. Furthermore, different methodologies may yield different outcomes, and this approach aims to maintain rigor, consistency and transparency in assessing individual country contributions to SDG5-focused research.

5.9.25. Compound Annual Growth Rate (CAGR) of women to men ratio of corresponding authorships in international collaborations

Definition of indicator

Compound annual growth rate (CAGR) is defined as the year-over-year constant growth rate over a specified period of time (2012-2021). Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series. Throughout, the term CAGR is also referred to as '(yearly) growth rate.'

The indicator is calculated for the average proportion of women corresponding authors on publications resulting from international collaboration across different countries and fields of research and development (FORD).

Rationale

Women publish fewer research papers on average than men and women are less likely than men to collaborate internationally on research papers ([268]) and this may affect the impact of their publications (as measured by citations). Funding agencies emphasise the above-mentioned dimensions in the evaluation of research proposals, and so there may be a gender gap in research evaluation ([269]), disadvantaging women in grant competitions compared to their men counterparts. This may lead to a vicious circle as with less funding women may not be able to publish at same rates and gather similar attention and impact as men. This indicator looks at the change in the average proportion of women among authors on publications resulting from international collaboration.

Computation method

Data needed

$(AIPW)_{cyi}$ Average proportion of women among corresponding authors on publications resulting from international collaboration in a given country (*c*), year (*y*) and field (*i*). **Unit: Unitless**.

N Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number**.

Source of data

Computed using Scopus ™ data and NamSor ™

Computation formula

CAGR for average proportion of women among authors = $(\mathrm{AIPW_{Cei}}/\mathrm{AIPW_{Csi}})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

([268]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([269]) Jappelli, T., Nappi, C.A., Torrini, R., '*Gender effects in research evaluation*', *Research Policy*, Vol. 46, No. 5, 2017, pp. 911-924.

$\mathrm{AIPW}_{\mathrm{Csi}}$ denotes the average proportion of women corresponding authors on publications resulting from international collaboration in a given country (*c*) and field (*i*) in the start year;

$\mathrm{AIPW}_{\mathrm{CeS}}$ denotes the average proportion of women corresponding authors on publications resulting from international collaboration in a given country (*c*) and field (*i*) in the end year.

5.10. Official Portal for EU open data (data.europe.eu)

Content-based rationale

The measurement of the extent to which research incorporates a gender dimension is not limited to research results published in peer-reviewed publications. Using data available on data.europe.eu, the extent to which a gender dimension in R&I content is incorporated within projects funded by the European Union Framework Programme for Research and Innovation for 2014-2020 (Horizon 2020) and 2021-2027 (Horizon Europe) can be explored. This is particularly relevant in light of strengthened requirements under Horizon Europe to include a gender dimension by default in all projects, unless the topic description explicitly specifies otherwise. Additionally, given the importance of understanding how gender interacts with other characteristics, an indicator is included measuring the proportion of Horizon 2020 and Horizon Europe projects which are also integrating intersectional aspects in their content, i.e., they are seeking to improve understanding of how gender intersects with other characteristics (such as age, disability status, and ethnicity) to affect experiences of disadvantage and discrimination.

Broad overview of the source

The indicators presented in this section were computed by Elsevier using Horizon Europe and Horizon 2020 data, retrieved from the official portal for European data (https://data.europa.eu/data/datasets/cordish2020projects?locale=en), as well as Horizon Europe data, retrieved from the same portal (https://data.europa.eu/data/datasets/cordis-eu-research-projects-under-horizon-europe-2021-2027?locale=en).

The Portal contains all project deliverables and other publications linked to Horizon 2020 and Horizon Europe projects. This dataset contains projects and organisations funded by the European Union under the Horizon 2020 Framework Programme for Research and Innovation from 2014 to 2020 as well as the Horizon Europe Framework Programme from 2021 onwards. Various files available from the Portal have been used, linked by the project ID. However, only files containing report summaries, projects, publications, titles, final results or project deliverables were included in this process. Since a limited number of files were checked, the calculated values may underestimate the actual integration of the gender dimension in R&I content in Horizon 2020 or Horizon Europe projects.

Although Horizon Europe builds on its predecessor Horizon 2020, it introduced more stringent and comprehensive policies regarding the integration of a gender dimension in R&I content. While Horizon 2020 encouraged the integration of sex- and/or gender-based analysis into research and innovation, Horizon Europe mandates it as a requirement for all R&I projects, reflecting a stronger commitment to gender equality. This shift underscores the recognition that the consideration of a gender perspective can enhance the societal relevance and impact of the conducted research. Another factor to take into account is that Horizon 2020 is already closed, therefore its project number will not change and it may be considered complete. Horizon Europe, on the other hand, is in its middle period and many projects have not begun or not yet been reported on. To illustrate this further, Horizon Europe included approximately 9 000 projects at the time of data extraction, all of which were in their early years, while Horizon 2020 had more than 30 000 projects, most of which were completed by the time of the analysis. Based on these differences, both indicators are not directly comparable.

Cut-off date

The cut-off date for Horizon 2020 and Horizon Europe data was 1 November 2023. As noted above, this means that the dataset for Horizon Europe is not complete as the Framework Programme is ongoing.

5.10.1. Proportion (%) of a country's Horizon 2020 and Horizon Europe projects integrating a gender dimension in their R&I content (GDRIC)

Definition of indicator

This indicator calculates the proportion of Horizon 2020 and Horizon Europe projects that integrate gender or sex-sensitive analysis, broken down by country.

Rationale

The European Commission seeks to promote the integration of the methods of sex/gender analysis into research design and process as a way of preventing bias in research, promoting better quality of outcomes in R&I and achieving cross-cutting benefits. To assess whether the introduction of new policies has resulted in changes in the research landscape, this indicator will reveal changes over time in how much projects stemming from Horizon 2020, European Union's Framework Programme for Research and Innovation (2014-2020), are integrating sex or gender as part of the project content. Integration of a gender dimension into R&I content is one of the gender equality provisions set for Horizon Europe, the EU Framework Programme for Research and Innovation 2021-2027, and therefore this integration is now a requirement by default. It is evaluated under the excellence criterion, unless the topic description specifies otherwise.

To ensure consistency with the indicator on the percent of a country's publication integrating a gender dimension in research content, the same bibliometric approach was used. Bibliometric analyses are used to identify Horizon 2020 and Horizon Europe projects that integrate a 'gender dimension.' To strengthen the integration of sex and gender analysis in research and innovation, the European Commission established the Gendered Innovations expert group that produced reports on incorporating the gender dimension across various research fields, defining research integrating the 'gender dimension' as research that integrates sex and gender analysis into their content. 'Sex' refers to basic biological characteristics of females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge ([270]) ([271]).

Computation method

Data needed

$\mathbf{HP}GDP_{cy}$ Number of Horizon 2020- / Horizon Europe-funded projects, tagged as human or animal as the research subject, integrating a gender dimension in a given country (*c*) and year (*y*). **Unit: Number**

$\mathbf{HP}_{CY}$ Number of Horizon 2020- / Horizon Europe-funded projects, tagged as human or animal as the research subject, in a given country (*c*) and year (*y*). **Unit: Number**

Source of data

Computed using Horizon 2020 and Horizon Europe data, retrieved from: https://data.europa.eu/data/datasets/cordis-eu-research-projects-under-horizon-europe-2021-2027?locale=en and for Horizon Europe projects from https://data.europa.eu/data/datasets/cordish2020projects?locale=en for Horizon 2020 projects. Please see section above on the broad overview of the sources.

Computation formula

For a given country (*c*) and year (*y*), the formula for this indicator is:

Proportion of a country's projects integrating $GD_{cy} = \frac{\mathbf{HP}GDP_{cy}}{\mathbf{HP}_{cy}}$

([270]) European Commission, *Gendered Innovations 2: How Inclusive Analysis contributes to Research and Innovation*, 2020. https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en

([271]) European Commission, *Gendered Innovations – How gender analysis contributes to research: report of the expert group 'Innovation through gender'*, 2013. https://op.europa.eu/en/publication-detail/-/publication/d15a85d6-cd2d-4fbc-b998-42e53a73a449

Specifications

The bibliometric analysis is based on the following strategy:

For each project, tag the project based on available text fields (abstracts, titles, objectives, results etc):

- Research subject: classify as human, animal, or neither.

- Include research on non-human species in the search. Whilst discussions of the methods of analysis of gender dimension in research have been focused so far on human context, it has been recognised that sex, in particular, can play an important role in controlling the 'lifepath' of non-human species (for example, plants, animals or cells) that make up the natural ecosystems in which humans coexist and whose wellbeing they influence. Therefore, and especially in the context of societal and environmental challenges, such as those linked to the effects of climate change, it is important to also promote methods of sex/gender analysis in research more widely.

- Exclude research which applies to only a single-sex or gender as a recognition that while some research may apply to only single-sex out of necessity (for example – pregnancy or prostate research), a large body of research in animals contains persistent historical bias, thereby excluding both sexes and biasing results ([272]). This approach avoids including this latter body of animal research in which sex and gender dimension have not been considered in the research design.

This approach has been chosen as a means to identify research in which the researchers have made some effort to compare characteristics/behaviour of women/females and men/males by including more than a single-sex or gender in the research rationale or output. Research on non-human species is included in the search.

A key limitation of this approach for Horizon Europe lies in the timeline and the availability of relevant data. The Horizon Europe programme commenced in 2021, with the first calls for project proposals launched in that year, and projects funded under the 2021 Work Programmes starting in 2022. The analysis covers projects from 2021 to 2023, limiting the number of substantial outputs available for assessment. It is also important to note that, at the European Commission level, the effective integration of the gender dimension into R&I content is assessed upon project completion. Consequently, the findings for Horizon Europe should be interpreted with caution.

Creation of datasets for gender dimension in R&I content

The publications included in the assessment were retrieved using any of the following eight queries:

i	(({men} OR {man} OR boy* OR {male} or mascul*) AND ({women} OR {woman} OR girl* OR {female} OR femin*))
ii	{sexual dimorphism}
iii	(sexual* AND dimorph*)
iv	"gender dimension"
v	"gender difference"
vi	{sex ratio}
vii	"sex difference"
viii	"sexual reproduction"

([272]) Berry, A.K. and Zucker, I., 'Sex bias in neuroscience and biomedical research', *Neuroscience & Biobehavioural Reviews*, Vol. 35, No. 3, 2011, pp. 565-572.

5.10.2. Proportion (%) of a country's Horizon 2020 and Horizon Europe projects integrating intersectional aspects

Definition of indicator

This indicator calculates the proportion of Horizon 2020 and Horizon Europe projects that integrate gender or sex-sensitive analysis combined with intersectional aspects, broken down by country.

Rationale

The Gender Equality Strategy 2020-2025 states that "the intersectionality of gender with other grounds of discrimination will be addressed across EU policies." ([273]) Action 5 of the ERA Policy Agenda 2022-2024 explicitly underlines that "there is a particular need to (…) open gender equality policies to inclusiveness and intersections with other diversity categories and potential grounds for discrimination, such as ethnicity, disability or sexual orientation ([274]). Consequently, one of the four outcomes of ERA Action 5 is the design of "a policy approach to inclusive gender equality, that addresses gender mainstreaming and opening to intersectionality with other diversity dimensions to advance the new ERA".

Therefore, this indicator is relevant for the EU's wider objective to improve understanding of how gender intersects with other characteristics (such as age, disability status and ethnicity) to affect experiences of disadvantage and discrimination ([275]). The Gendered Innovations 2 report developed and highlighted a methodology for intersectional research, which was applied for this indicator.

Computation method

Data needed

$\mathbf{HP}GDI_{CY}$ Number of Horizon projects, integrating a gender dimension plus an intersectional aspect in a given country (*c*) and year (*y*). **Unit: Number**

$\mathbf{HP}_{CY}$ Number of Horizon projects, in a given country (*c*) and year (*y*). **Unit: Number**

Source of data

Computed using Horizon data, retrieved from CORDIS (https://cordis.europa.eu/)

Computation formula

For a given country (*c*) and year (*y*), the formula for this indicator is:

Percent of a country's projects integrating GD_{cy} = $\frac{\mathbf{HP}GDI_{cy}}{\mathrm{HP}_{cy}}$

Specifications

The bibliometric analysis is based on the same strategy as indicator 'Proportion (%) of a country's Horizon projects integrating a gender dimension in their R&I content (GDRIC)'.

([273]) European Commission, *Gender Equality Strategy 2020-2025*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

([274]) European Commission, Directorate-General for Research and Innovation. *European Research Area Policy Agenda – Overview of actions for the period 2022-2024* [2021]. Available at: https://research-and-innovation.ec.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

([275]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final*, 2020. https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

The resulting Horizon 2020 and Horizon Europe projects were again queried with a search query related to intersectionality. This query utilised a shortlist of keywords from the Gendered innovations 2 report ([276]) which were refined and augmented with additional keywords. The final list of keywords used in addition to the GDRIC query were:

("ethnic*" OR "race" OR "racis*" OR "indigen*" OR "LGBT*" OR "religio*" OR "sexual orientation" OR "belief" OR (("vulnerable" OR "marginali?ed" OR "disadvantaged" OR "underprivileged") AND ("group" OR "population" OR "communit*")) OR "social class" OR "social origin" OR "disabilit*" OR "conflict?affected" OR "violence?affected" OR "migrat* status" OR "intersect*"

The use of “age” as an intersectional factor was removed from the query due to the inclusion of statistics for medicine studies – mentioning the age of patients in abstracts for stratification of groups – which let to too many false positive results.

5.11. EPO Worldwide Patent Statistical Database (PATSAT)

Content-based rationale

Women have been shown to lag behind men in terms of the size (as measured by the number of peer-reviewed scientific publications) and impact (as measured by citations to their publications) of their scientific production, as well as their propensity to partner on an international scale (as measured by the proportion of papers co-authored by researchers located in at least two countries) ([277]) ([278]). Because of the emphasis placed by funding agencies on the above dimensions in the evaluation of research proposals, women could be disadvantaged in grant competitions relative to their men counterparts. In grant competitions focusing more heavily on applied research, the number of patent applications in which a researcher is listed as an inventor might also prove to be a decisive factor in the funding decision. Thus, techno metric indicators derived from PATSTAT are integrated in the She Figures publication to monitor gaps in the contribution of women and men to the production of inventions by country, year and technological fields.

Broad overview of the source

The indicators presented in this section were computed using raw bibliographic data derived from the European Patent Office (EPO) Worldwide Patent Statistical Database (PATSTAT). PATSTAT covers patent data from over 150 offices worldwide, including EPO, the United States Patent and Trademark Office (USPTO) and the Japan Patent Office (JPO). The USPTO covers the United States, the EPO covers Europe, the JPO covers Japan, and so forth. For the She Figures publication, the statistics are based on the EPO within PATSTAT, as the European market is one of the largest in the world and certainly the most relevant in the context of the She Figures publication, since it covers all countries associated with the European Research Area.

Note that statistics on inventorships can be produced by measuring issued patents or patent applications when working with EPO data. On a conceptual level, if the goal is to get a sense of the inventive/innovative capacity of a given entity (e.g. women in a given country) rather than of ‘marketable/innovative outputs’, as in this study, then applications are more appropriate. Furthermore, in cases where trends in the inventiveness of entities are to be investigated, also as in this study, the capacity to produce timely data is important. In this regard, issued patents have the disadvantage of running behind and becoming visible only years after the innovative activity has taken place. Thus, from a methodological standpoint, applications are still preferable. Consequently, EPO patent applications (kind codes: A1 and A2) were retained in computing this set of indicators.

([276]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020)0628 final*, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

([277]) DG for Internal Policies, *Precarious Employment in Europe Part 1: Patterns, Trends and Policy Strategies*, Brussels, 2016. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/587285/IPOL_STU(2016)587285_EN.pdf

([278]) DG Research and Innovation, *She Figures 2018*, Publications Office of the European Union, Luxembourg, 2019. https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en

Definitions and preliminary data treatment

A European patent application is characterised by one or more applicants, inventors and classified according to the different areas of technology to which they pertain.

The applicant is the natural and/or legal person and/or entity that filed the application. The inventor is the real creator of the invention. In EU this term is defined by each Member State's legislation. In general, in order for one to be considered an inventor, it is acknowledged that a certain level of contribution to the development of the creative elements of an invention (technical creativity) must be met. Inventors are always private individuals and are always entitled to be designated on the patent, regardless of who files the application. Joint inventors or co-inventors exist when a patentable invention is the result of the inventive work of more than one inventor, even if they did not contribute in equal parts.

Concerning the technological classification system, this is based on the International Patent Classification (IPC) system used in over 100 countries to identify the content of patents in a uniform manner. It was created under the Strasbourg Agreement (1971).([279]) The classification is updated on a regular basis (the 1st of January of each year) by a Committee of Experts, consisting of representatives of the Contracting States of that Agreement with observers from other organisations, such as the European Patent Office. The first level of the IPC hierarchy permits to identify eight sections, in particular:

A. Human Necessities
B. Performing Operations, Transporting
C. Chemistry, Metallurgy
D. Textiles, Paper
E. Fixed Constructions
F. Mechanical Engineering, Lighting, Heating, Weapons and Blasting
G. Physics
H. Electricity.

PATSTAT data are organised in a relational database, consisting of different tables that can be merged by means of specific keys. In this case, all the applications were considered with a first filing year from 2005 to 2018, attributing to each of these a binary variable (0=No/1=Yes) that specifies the IPC section to which it belongs. In a second step, the application was attributed to the country of the first applicant (despite such information not always being filled).

As a last step, all the inventors referred to each application were extracted, to which a specific strategy had already attributed the sex.

The algorithm is based on a standard methodology offered by GendRE API, a package developed by NamSor™ (a European designer of name recognition software committed to promoting diversity and equal opportunity); in fact, this API enables the extraction of the sex from personal names, surnames and countries.

Cut-off date

The cut-off date for PATSTAT data was 1 September 2023.

([279]) Summary of the Strasbourg Agreement concerning the International Patent Classification (1971). Available at: https://www.wipo.int/treaties/en/classification/strasbourg/summary_strasbourg.html

5.11.1. Proportion of women among inventors for all classes and by IPC class

Definition of indicator

This indicator is the proportion of women among inventors (based on inventorships). The absolute number of inventorships used in computing this indicator is based on fractionalised counts of patent applications across their corresponding inventors: for example, if a patent application involves 10 inventors, each inventor is attributed an equal fraction of the inventorships (i.e. 1/10 of the invention).

Rationale

Women still lag behind men in terms of their measurable scholarly output ([280]) ([281]) ([282]). Given the increasing reliance on bibliometric statistics (i.e. statistical analyses of written publications such as books or articles) for research evaluation purposes in research assessment exercises and grant competitions, the lower scientific output of women could lead to reduced chances of being funded (or the receipt of lower funding amounts), which could in turn decrease their scientific output, thereby creating a vicious circle. In grant competitions focusing more heavily on applied research, the number of patent applications on which a researcher is listed as an inventor might also prove to be a decisive factor in the funding decision. This indicator looks at the size of the technological output of women across different countries and fields of technology.

Computation method

Data needed

(WI_{cyi}) Sum of fractionalised inventorships for women in a given country (*c*), year (*y*) and section (*i*, based on the International Patent Classification [IPC]). **Unit: Total of fractionalised counts**.

(N_{cyi}) Total number of fractionalised inventorships in a given country (*c*), year (*y*) and IPC section (*i*). **Unit: Total of fractionalised counts**.

Source of data

Computed using PATSTAT data

Computation formula

Proportion of Women among inventors, for a given country (*c*), year (*y*) and IPC section (*i*) $= \frac{WI_{cyi}}{N_{cyi}}$

Specifications

Once having identified the sex of the inventors of each application (following the approach described in the general part of this section), it is possible to obtain a dataset that constitutes the base of the computation of this indicator that will be exploited for each year, country and by crossing this information with the IPC classifications.

All EPO patent applications are classified based on the **International Patent Classification (IPC)** of the World Intellectual Property Organization (WIPO) in PATSTAT. This hierarchical classification is divided into eight sections (Level 1), which are further divided into classes (Level 2), subclasses (Level 3), main groups (Level 4) and subgroups (lower level). This classification is not mutually exclusive (i.e. each patent application is classified into one or more sections, classes, subclasses, main groups and subgroups). Thus, a given patent application can contribute to the scores of more than one of the eight IPC sections for which this

([280]) Elsevier, *Gender in the Global Research Landscape*, 2017. https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

([281]) DG Research and Innovation, *She Figures 2018*, Publications Office of the European Union, Luxembourg, 2019. https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en

([282]) European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities*, 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

indicator has been computed, in addition to the total for all EPO patent applications (subscript I in the above formula):

A. Human Necessities
B. Performing Operations, Transporting
C. Chemistry, Metallurgy
D. Textiles, Paper
E. Fixed Constructions
F. Mechanical Engineering, Lighting, Heating, Weapons and Blasting
G. Physics
H. Electricity

(T) Total across all sections including unclassified patent applications (unique/distinct count of patent applications across sections).

Comments and critical issues

Although the data covers all the patent applications registered in the considered years, it is necessary to observe that for some of these it was not possible to correctly derive the sex of all inventors. This can then be associated with a **confidence interval (CI),** for a given country, year and IPC classification. The CI takes into account the total number of inventorships (including those with an unclassified sex of the inventors), N_{cyi} and considers those with classified sex of all inventors as a random sample.

The 95% CI is determined as follows:

1. Create 1,000 resamples in which N applications (N being the number of applications in each group), resampled with replacement.

2. In each of these resamples, the gender of inventors is assigned based on their probabilities.

3. Then, the proportion of inventors that are women is computed for each application, as described above.

4. For each resample, compute the average proportion of women inventors.

5. Take the percentiles 2.5 and 97.5 from the empirical distribution generated in (3)

The steps above are followed for three scenarios, one being the base case, and two additional scenarios that consider the possibility of systematic errors on the probabilities of gender assignment: the first one consists of reducing the probabilities of authors being women, while the second increases these probabilities. The more extreme values for percentiles 2.5 and 97.5 in each scenario are used as the boundaries of the margin of error.

Potential errors in coverage mainly refer to:

- Bias in the number of documents over time: there are applications for which the name of the inventor is missing. Despite this phenomenon, the available and correct sample for the prior She Figures reports was sufficiently large to produce accurate statistics over the entire time frame covered in the publication.

- Bias in favour of some countries: more applications could be attributed to some countries and fewer to others. However, since the indicators are ratios of variables referring to the same given country, it is expected that such bias does not affect the cross-country comparability of them. A similar picture emerges from PATSTAT data.

- Bias in favour of disciplines: it was observed that the proportion of EPO patent applications for which the affiliation country and full given name is available for all inventors on them was high and relatively similar across IPC classes. It was therefore concluded that none of the IPC classes contributed significantly more or less than the others to the women proportion computed at higher aggregation levels (e.g. for IPC Sections or for all EPO patent applications). A similar picture emerges from PATSTAT data.

5.11.2. Compound annual growth rate (CAGR) of the proportion of the four-year percentage of women inventors

Definition of indicator

This indicator presents the compound annual growth rate of the proportion of women among inventors, meaning the average yearly percentage increase/decrease in the proportion, moving from one period to the next (using four-year moving periods, e.g. 2012–2015, 2013–2016 and so on, until 2018-2021), year on year.

Rationale

Prior She Figures reports have shown that women still lag behind men in terms of measurable scholarly output and among inventors. Given the increasing reliance on bibliometric statistics (i.e. statistical analyses of written publications such as books or articles) for research evaluation purposes in research assessment exercises and grant competitions, the lower scientific output of women could lead to reduced chances of being funded or the receipt of lower funding amounts, which could in turn decrease their scientific output, thereby creating a vicious circle. This indicator looks at the size of the scientific output (in terms of patent applications) of women compared to men across different countries and IPC sections.

Computation method

Data needed

(F_s) Estimated proportion in the start period. **Unit: Unitless.**

(F_e) Estimated proportion in the end period. **Unit: Unitless.**

(N) Number of years in the reference period (i.e. last year of end period – last year of start period). **Unit: Year.**

Source of data

Computed using PATSTAT data

Computation formula

CAGR for the proportion of women inventors = $(F_e/F_s)^{1/N} - 1$

5.11.3. Distribution of patent application by sex composition of the inventors' team

Definition of indicators

The indicators analyse the sex composition of the inventors' team for each patent application (e.g. by teams consisting of women only) and are the proportions of each type of composition and ratios of such proportions over each other.

Rationale

A patent is a legal title granting its holder the right, in specific countries and for a certain period, to prevent third parties from exploiting an invention for commercial purposes without authorisation. Any legal entity (one or more individuals and/or firms) could register a patent; when doing such action, it is mandatory to specify also the related inventors (which could be different from the entity that applies for the patent).

Thus, each patent application can have one named inventor (a lone/individual inventor) or multiple inventors (working collaboratively as part of a team). The determination of the sex of each named inventor allows for the identification of mutually exclusive sets of applications, i.e. those referred to a working-alone women (or men), those developed by teams of the same sex and those referred to mixed-sex teams.

The indicators shed light on the propensity of the women and men to work alone or in same-sex teams versus working in mixed teams as well as on how such collaboration patterns vary between countries and evolve over time.

Computation method

Data needed

(WI_{cy}) Number of applications attributed to the country *(c*) and year (*y*) from women inventors working alone. **Unit: Count**.

(MI_{cy}) Number of applications attributed to the country (*c*) and year (*y*) from men inventors working alone. **Unit: Count**.

($allWI_{cy}$) Number of applications attributed to the country (*c*) and year (*y*) from teams of women inventors. **Unit: Count**.

($allMI_{cy}$) Number of applications attributed to the country (*c*) and year (*y*) from teams of men inventors. **Unit: Count**.

(pWI_{cy}) Number of applications attributed to the country (*c*) and year (*y*) from predominantly women teams. **Unit: Count**.

(pMI_{cy}) Number of applications attributed to the country (*c*) and year (*y*) from predominantly men teams. **Unit: Count**.

(pTI_{cy}) Number of applications attributed to the country (*c*) and year (*y*) from sex-balanced teams. **Unit: Count**.

The ratios were computed using **a four-year period** (2017–2020). This way, the samples used were larger, providing more robust estimates.

Source of data

Computed using PATSTAT database

Computation formula

Once the sexes of its inventors have been identified (see the general introduction to this section), it is possible to compute the following variables for each *y* and *c*:

- number of applications from women/men inventors working alone: WI_{cy} or MI_{cy}
- number of applications from teams with inventors of the same sex: $allWI_{cy}$ or $allMI_{cy}$
- number of applications from teams which consist predominantly of women: pWI_{cy}. These are teams with more than 60% women.
- number of applications from teams which consist predominantly of men: pMI_{cy}. These are teams with more than 60% men.
- number of applications from sex-balanced teams: bTI_{cy}. These are teams with between 40% and 60% women.

All the above situations are mutually exclusive, so that their sum corresponds to the total number of applications for each year (*y*) and country (*c*):

$$WI_{cy} + MI_{cy} + allWI_{cy} + allMI_{cy} + pWI_{cy} + pMI_{cy} + bTI_{cy} = TOT_{cy}$$

The indicators able to synthesise the phenomenon consider comparisons between 'similar' variables or with the total number of applications. The IPC section was ignored when estimating this indicator to avoid issues due to small samples for country/year combinations.

All the indicators were referred to TOT_{cy}; in particular:

- proportion of applications from women/men inventors working alone: WI_{cy}/TOT_{cy} or MI_{cy}/TOT_{cy}
- proportion of applications from teams with inventors of the same sex: $allWI_{cy}/TOT_{cy}$ or $allMI_{cy}/TOT_{cy}$
- proportion of applications from teams which consist predominantly of women: pWI_{cy}/TOT_{cy}. These are teams with more than 60% women.
- proportion of applications from teams consisting predominantly of men: pMI_{cy}/TOT_{cy}. These are teams with more than 60% men.
- proportion of applications from sex-balanced teams: $bTI_{cy}//TOT_{cy}$. These are teams with between 40% and 60% women.

Specifications

An application is characterised by a reference year (*y*) and a list of applicants. The latter may be legal entities and may involve different persons than the inventors.

In case of multiple applicants, the order in which they were specified in the application is maintained also in the collected data. To associate an application to a specific country (*c*), it is considered the country of residence of the first (main) applicant.

Comments and critical issues

Potential errors in coverage mainly refer to:

- Bias in the number of documents over time: there are applications for which the name of the inventor is missing. Despite this phenomenon, the available and correct sample for prior She Figures reports was sufficiently large to produce accurate statistics over the entire time frame covered in the publication. A similar picture emerges from the PATSTAT data.
- Bias in favour of some countries: more applications could be attributed to some countries and fewer to others. However, since the indicators are ratios of variables referring to the same given country it is expected that such bias does not affect the cross-country comparability of them.
- Bias in favour of disciplines: it was observed that the proportion of EPO patent applications for which the affiliation country and full given name is available for all inventors on them was high and relatively similar across IPC classes. It was therefore concluded that none of the IPC classes contributed significantly more or less than the others to the women-to-men ratios computed at higher aggregation levels (e.g. for IPC Sections or for all EPO patent applications). A similar picture emerges from the PATSTAT data.

5.11.4. Compound annual growth rate (CAGR) of the four-year proportions of patent applications, by sex composition of the inventors' team

Definition of indicators

This indicator presents the compound annual growth rate of the types of working teams, representing the average yearly percentage of increase/decrease in the related proportions. To obtain more robust estimation, this indicator is built considering a four-year moving period, e.g. 2016–2019, 2017–2020.

Rationale

A patent is a legal title granting its holder the right, in specific countries and for a certain period, to prevent third parties from exploiting an invention for commercial purposes without authorisation. Any legal entity (one or more individuals and/or firms) could register a patent; when doing such action, it is mandatory to specify also the related inventors (which could be different from the entity that applies for the patent).

Thus, each patent application can have one named inventor (a lone/individual inventor) or multiple inventors (working collaboratively as part of a team). The determination of the sex of each named inventor permits to identify mutually exclusive sets of applications, i.e. those referred to a working-alone women (or men), those developed by teams of the same sex and those referred to mixed-sex teams.

The indicators shed light on the propensity of the two sexes to work alone or in same-sex teams versus working in mixed-sex teams as well as on how such collaboration patterns vary between countries and evolve over time.

Computation method

Data needed

(F) Estimated proportions of each type of working team in a start and an end year. **Unit: Unitless.**

(N) Number of years in the reference period (calculated by subtracting the defined start year from the defined end year). **Unit: Number.**

Source of data

Computed using PATSTAT data

Computation formula

CAGR for each type of working team = $(\mathrm{F_e}/\mathrm{F_s})^{1/\mathrm{N}} - 1$

where:

s refers to the start year;

e refers to the end year;

$\mathrm{F_s}$ denotes the estimated proportion of each type of working team in the start year;

$\mathrm{F_e}$ denotes the estimated proportion of each type of working team in the end year.

The ratios were computed using **four-year moving periods** (e.g. 2016–2019, 2017–2020). This way, the samples used were larger, providing estimates that are more robust.

5.12. She Figures Index

5.12.1. Introduction

The She Figures Index aims to serve as a tool to measure the extent to which EU Member States have made progress towards gender equality in the European Research Area (ERA).

Building upon the broad commitment of the European Commission regarding gender equality in the Gender Equality Strategy 2020-2025, the Commission's Communication on a new ERA ([283]) highlighted that progress towards gender equality remains slow and insufficient at EU level with different rates of progress across EU Member States. Recognising this imperative need for accelerated change, Horizon Europe, the European Union's Framework Programme for Research and Innovation for 2021-2027, introduced strengthened provisions aimed at fostering inclusive and gender-equal research environments. Firstly, certain legal entities are now required to have a Gender Equality Plan (GEP) in place (applicable for EU Member States and non-EU countries associated to Horizon Europe). Secondly, the integration of a gender dimension into R&I content is a mandatory requirement, evaluated as part of the excellence criterion. Finally, there is the European Commission's commitment to increasing gender balance throughout the programme, including a target of 50% for women's representation in Horizon Europe-related boards, expert groups and evaluation committees.

The She Figures Index is designed to reflect the current policy context regarding gender equality in R&I through selecting appropriate and representative indicators. Its development aligns with the framework of the ERA Policy Agenda 2022-2024, and particularly Action 5. 'Promote gender equality and foster inclusiveness, taking note of the Ljubljana declaration' ([284]). This Action takes a comprehensive approach to advance gender equality. Key components include the development of inclusive Gender Equality Plans, a strategy at EU level to counteract gender-based violence (including sexual harassment), a policy approach to inclusive gender equality and gender mainstreaming, and principles for the integration and evaluation of the gender perspective in R&I content in cooperation with national Research Funding Organisations (RFOs).

Moreover, as indicated in the European Strategy for Universities ([285]), the European Commission, in close cooperation with the stakeholders and the Member States, will develop a European framework for diversity and inclusion, including on gender gaps, identifying challenges and solutions for universities, and the needed support of public authorities and **address the underrepresentation of women in STEM fields** (…)". The She Figures Index contributes to the Strategy by considering these objectives, especially gender gaps and the underrepresentation of women in STEM fields.

Additionally, the She Figures Index aims to align with the objectives outlined in the New European Innovation Agenda ([286]) regarding attracting and retaining talent, especially promoting the representation of women in innovation start-ups and scale-ups, as well as addressing the gender funding gap.

The development of the pilot She Figures Index marks an important milestone in marking progress towards gender equality within the EU R&I system. Serving as an essential benchmarking tool, this index lays the groundwork for monitoring and evaluating progress in gender equality efforts.

This summary provides an overview of the development of the She Figures Index starting with its foundational conceptual framework and initial measurement framework. It then outlines the metrics that have been developed, followed by an assessment of the methodological approaches considered, such as weighting and aggregation. These are used to assess the correlation structure and refine the measurement framework

([283]) COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A New European Innovation Agenda [2022] COM/2022/332. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0332

([284]) European Commission, Directorate-General for Research and Innovation. European Research Area Policy Agenda – Overview of actions for the period 2022-2024 [2021]. Available at: https://research-and-innovation.ec.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

([285]) COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS on a European strategy for universities COM/2022/16 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A16%3AFIN

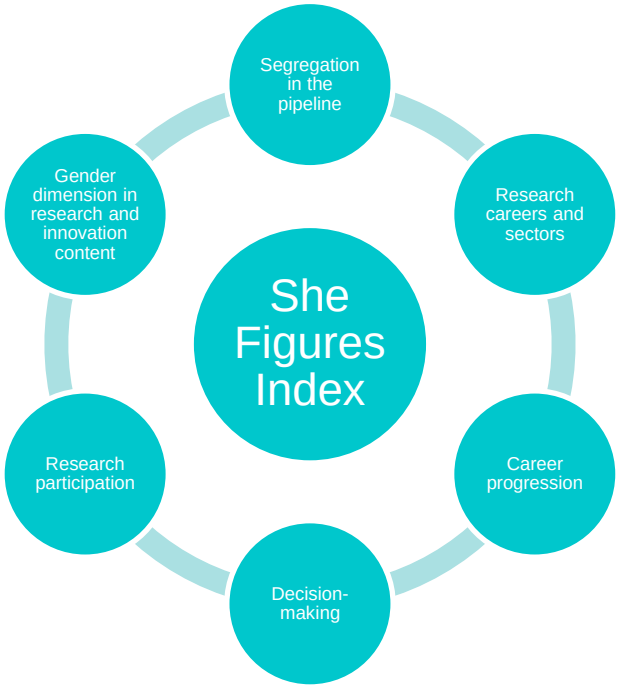
([286]) COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A New European Innovation Agenda [2022] COM/2022/332. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0332

before conducting a multi-model analysis. The section concludes by summarising the final model and presenting the results from rigorous robustness checks. Additional information can be found in Annex 4.

5.12.2. Conceptual framework development

Since 2012 and the ERA communication, gender equality and mainstreaming were prioritised with 3 objectives: gender equality in careers at all levels, gender balance in decision making and integration of the gender dimension into the content of R&I. The Index was designed to reflect these aspects and examines six key dimensions: segregation in the pipeline; research careers and sectors; career progression; representation in decision-making positions; research participation; and the gender dimension in R&I content (GDRIC). These dimensions form the conceptual framework for the She Figures Index, as shown in the following figure.

Figure 1 Conceptual framework for the She Figures Index



5.12.3. Measurement framework development

The conceptual framework served as a basis for selecting indicators from the 2021 edition of the She Figures report, which included the latest indicators available during the methodology's development. Selection criteria prioritised indicators with both conceptual relevance and methodological rigor. For instance, indicators were required to demonstrate consistent availability over time and minimise missing values. This approach was guided by the OECD and European Commission's Joint Research Centre Handbook on composite indicator development ([287]).

The process of short-listing indicators was conducted iteratively, involving extensive checks of missing values, correlation structures and Principal Components Analysis (PCA) ([288]). In total, 19 indicators were short-listed (see following table), and 14 indicators were ultimately included (see Table 5 Final list of indicators for the She Figures Index). Further details can be found in Annex 4.

([287]) OECD/European Union/EC-JRC, *Handbook on Constructing Composite Indicators: Methodology and User Guide*, OECD Publishing, Paris, 2008, https://doi.org/10.1787/9789264043466-en

([288]) PCA is a useful tool in that it can assess whether the selected indicators measure the same underlying construct. If that is the case, it is possible to verify that all indicators can be represented by a single component, using the Kaiser criterion of eigenvalues above 1 as a guide, ensuring that all rotated factors loadings are typically above 0.5, and that fit is adequate by checking that KMO values are normally above 0.5. Further information is available in Annex 4.

Table 3 Initial shortlisted indicators

Segregation in the pipeline	Research careers and sectors	Career progression	Representation in decision-making positions	Research participation	Gender dimension in Research and Innovation Content (GDRIC)
Proportion (%) of women among doctoral graduates, by broad field of study (Arts and Humanities)	Proportion of women researchers in the higher education sector	Proportion (%) of women among academic staff, by grade and total (Grade A)	Proportion (%) of women among heads of institutions in the Higher Education Sector (HES)	Average proportion (%) of women among authors on publications in all fields of R&D	Proportion (%) of a country's publications with a sex or gender dimension in their research content
Proportion (%) of women among doctoral graduates, by broad field of study (Business and Law)	Proportion of women researchers in the government sector	Proportion (%) of women among academic staff, by grade and total (Grade B)	Proportion (%) of women on boards, leaders	Distribution of patent application by sex composition of the inventors' team (%)	Proportion of country's Horizon* projects integrating a sex or gender dimension (GDRIC)
Proportion (%) of women among doctoral graduates, by broad field of study (Information and Communication Technologies)	Proportion of women researchers in the business enterprise sector	Proportion (%) of women among academic staff, by grade and total (Grade C)	Proportion (%) of women on boards, members and leaders	Research funding success rate differences between women and men	Proportion of country's Horizon* projects incorporating intersectional aspects
Proportion (%) of women among doctoral graduates, by broad field of study (Engineering, Manufacturing and Construction)					

* In this context, "Horizon" refers to both Horizon 2020 and Horizon Europe. These indicators were calculated separated for Horizon 2020 and Horizon Europe (see section 5.10).

5.12.4. Choice of metrics

The short-listed indicators require transformation into metrics, i.e. a score ranging from 0 to 1, where 0 represents the lowest score and 1 the highest score. The indicators fall into one of two categories, depending on the unit of analysis.

- Indicators focusing on share: measure the proportion of women and men in specific aspects relevant to gender equality in R&I. The unit of analysis here is the individual.

- Indicators focusing on level: measure the level of research publications or projects that incorporate a sex or gender dimension. The unit of analysis for these indicators is either projects or publications.

Therefore, careful consideration is needed to operationalise all selected indicators into the appropriate metrics, thus enabling their aggregation into a robust She Figures Index score.

Choice of metrics for shares

For indicators focusing on the shares of women and men in different aspects of R&I, two primary considerations were evaluated: (1) the choice of metric (linear or curvilinear metric) and (2) threshold selection (whether or not to cap at the parity point). In practice, this means that there are four types of metrics to consider, as illustrated below (Figure 2):

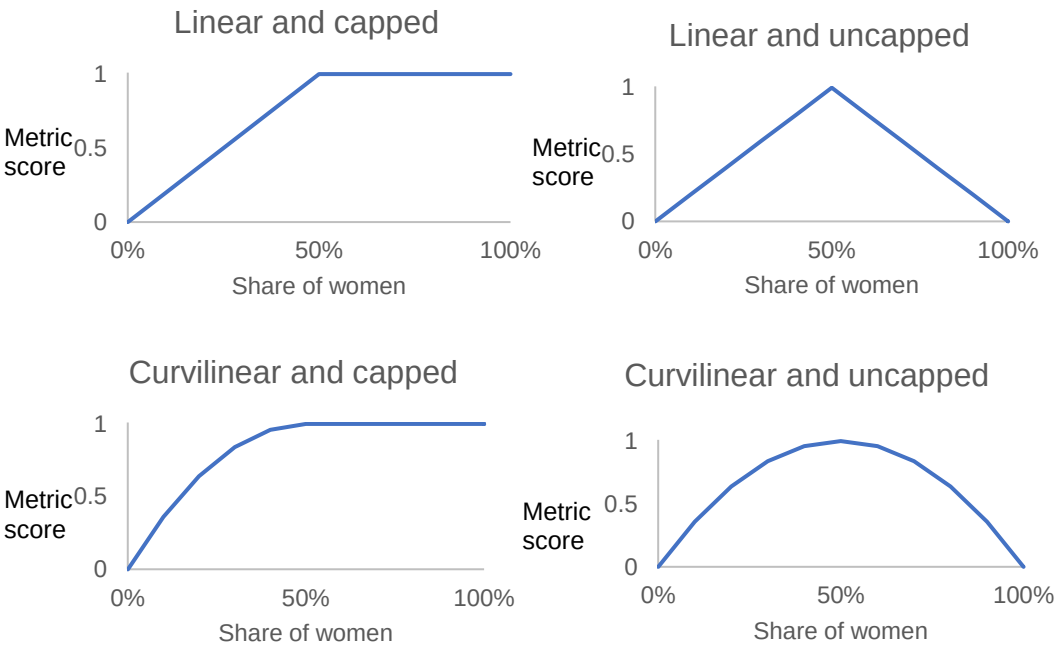
- Linear and capped: $X_W/0.5$ if $X_W \leq 0.5$, 1 otherwise

- Linear and uncapped: $\min(X_W, X_M)/0.5$

- Curvilinear and capped: $Blau\ (X_W, X_M)$ if $X_W \leq 0.5$, 1 otherwise
- Curvilinear and uncapped: $Blau\ (X_W, X_M)$

where X_W and X_M represent the share of women and men respectively for a given indicator, and where the Blau index is calculated as $Blau\ (X_W, X_M) = 2 \times (1 - (X_W^2 + X_M^2))$.

Figure 2 Metrics curvi/linear and un/capped



The following points should be considered in relation to how it is possible to interpret the different metrics presented above:

- Each metric measures gender equality in relation to exact parity, i.e. the maximum score of 1 is obtained where there are 50% of women and 50% of men.
- Linear metrics measure progress towards the goal of parity equally, regardless of whether women's share of representation is low or closer to parity, i.e. a score increase will be the same where women's share increases from 0% to 10% and when it increases from 40% to 50%.
- The benefit of using curvilinear metrics, in contrast, is that it rewards progress at lower levels of representation more highly than when representation is near parity. In fact, where women's representation reaches 40%, the associated score is 0.96 when computed using a curvilinear metric. This may therefore be interpreted as sensitive to 'critical mass', that is where progress beyond 30% leads to smaller improvements in score, and with minimal progress in the score beyond 40% representation. Another benefit of using a curvilinear metric is that it puts less emphasis on the parity point (exactly 50% representation) and is instead more apt at measuring 'gender balance' (the interval 40%-60%).
- Capping is associated with a women's empowerment approach, i.e. reflecting policy priorities that focus on improving the situation of women.
- Uncapping is associated with a gender equality approach, i.e. reflecting policy priorities that seek to ensure that both women and men are equally represented.

All four types of metrics (linear capped, linear uncapped, curvilinear capped, and curvilinear uncapped) underwent rigorous testing as part of a multi-modeling process.

Following extensive discussions with Members of the Steering Group for She Figures ([289]), a consensus was reached in favor of using a curvilinear uncapped metric. This choice was informed by the fact that it offers a more nuanced representation of gender balance as it assigns higher scores within the interval of 40% to 60% rather than focus narrowly on parity (i.e. highest score at the exact parity point, corresponding to 50%). In addition, it allows for a comprehensive evaluation of gender imbalances in both directions since it does not only highlight scenarios where women are underrepresented but also identifies situations where men may be disadvantaged.

Choice of metrics for levels

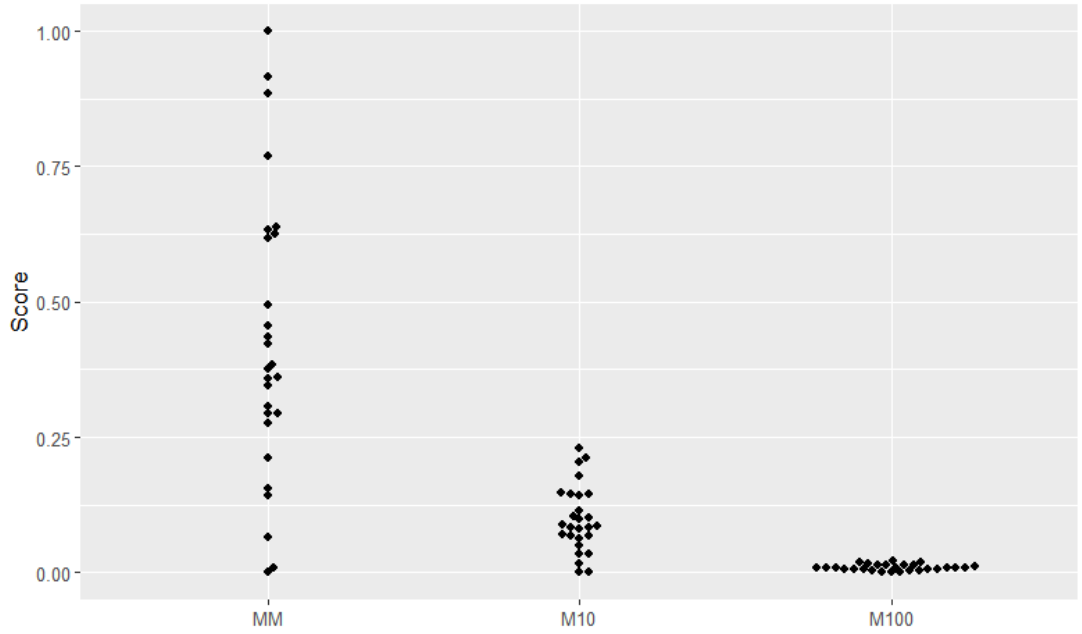
For indicators focusing on levels, a different approach was required. To create metrics ranging from 0 to 1, different variations of the Min-Max approach were considered. In its basic form, the Min-Max approach uses the observed minimum and maximum values in the dataset to define the interval boundaries, thereby rescaling the data to a range of 0 to 1.

Several options for metrics were examined. First, taking the usual Min-Max approach, calculated on the basis of the observed minimum and maximum (MM); a Min-Max approach using the ceiling value of 10%, i.e. considering the range from the observed minimum to 10% (M10); and a Min-Max approach using the ceiling value of 100%, i.e. considering the full potential range of values from the observed minimum to the theoretical maximum of 100% (M100). The range of respective scores for these three approaches are illustrated in **Error! Reference source not found.** for the indicator measuring the extent to which the gender dimension h as been included in publications.

The first approach (MM) means that the scores can take on the full range of values between 0 and 1, and can be interpreted as the observed level of the extent to which the gender dimension has been used. The second (M10) and third (M100) approaches have more limited range of scores, and represent a more normative interpretation where the ceiling values used could be interpreted as targets or objectives. However, the arbitrary nature of these thresholds is problematic, in that the target value has yet to be determined, or even if a target is desirable or feasible. Despite these notes of caution, all three options were included in the multi-modelling assessment.

([289]) Members of the Steering group comprise of representatives from different services of the European Commission, (international) statistical organisations and national statistical correspondents involved in She Figures 2024 that expressed interest in being members of the Steering Group.

Figure 3 Beeswarm plots of metrics for the gender dimension in publications by method (MM, M10, M100)



As described further in the next sections, the selection of the metric for indicators focusing on levels was based on testing these alternatives, leading to the selection of the Min-Max in its basic form.

5.12.5. Assessment of methodological alternatives (aggregation, weights)

The two main aspects characterising the assessment of methodological alternatives are aggregation and weights.

Aggregation

Various aggregation methods exist to combine individual indicators across dimensions. These methods differ primarily in the extent to which they allow for compensation among indicators. The arithmetic mean provides full compensation across all indicators, while the geometric and harmonic means increasingly impose penalties for disparities between indicators. The following summary table outlines these different levels of compensation, together with some illustrations (Table 4). The table provides three examples, each with three indicators.

In example 1, all three indicators are equal, with a value of 50, representing a situation where there is full equality. In example 2, there is moderate inequality, as the three indicators range from 40 to 60, though the average is the same as example 1. In example 3, there is greater inequality, as despite maintaining an average of 50, the range increases from 10 to 90. Using an arithmetic mean allows for full compensation between indicators, and does not distinguish between the disparities in indicators within the three examples. Using a geometric mean introduces a penalty, providing a mean of 50 when indicators are equal but increasingly pulling down this score as inequality between indicators increases. This penalisation is even more blatant when a harmonic mean is used. It is important to note that it is not necessary to apply the same aggregation methods at all levels (indicators, dimensions), and indeed, allowing for full compensation at levels closer to the indicators themselves might be appropriate.

Table 4 Summary table of aggregation methods and associated compensations

	Indicator 1	Indicator 2	Indicator 3	Arithmetic	Geometric	Harmonic
				$\frac{\sum_{i=1}^{n} x_i}{n}$	$\left(\prod_{i=1}^{n} x_i\right)^{1/n}$	$\left(\frac{\sum_{i=1}^{n} x_i^{-1}}{n}\right)^{-1}$
				Full compensation	Medium compensation	Limited compensation
Example 1	50	50	50	50	50	50
Example 2	40	50	60	50	49.3	48.6
Example 3	10	50	90	50	35.6	22.9

The choice of indicators was discussed with Steering Group members, who expressed a preference for simpler aggregation methods, i.e. arithmetic mean, provided that it yielded scores comparable to those from more complex methods. This preference was assessed as part of the multi-modelling approach, which provided empirical evidence supporting the use of the arithmetic mean.

Weighting

Another consideration is the weighting scheme adopted to compute the scores of the She Figures Index. Two weighting options were evaluated:

- The most straightforward approach assigns **equal weights to all indicators and sub-pillars**. This is a choice that is easily communicable and interpretable. However, due to the correlation structure of the indicators, certain indicators exert greater influence on the overall results of the Index than others.

- Another approach is to consult with relevant stakeholders and derive a set of weights that is informed by the **importance given to each pillar by the community of practice** (see Annex 4 for further details). Generally, assessing the relative importance at the indicator level is not possible due to the level of details involved. This is another easily communicable and interpretable method, however it is highly subjective and dependent on the representativeness of the community from which these weights are derived.

After discussion with the Steering Group members, these options were retained and it was decided that other approaches would not be further considered. This decision was influenced by two primary factors. First, the dataset encompassing data from 27 Member States lacked sufficient statistical power to derive meaningful PCA-based weights. Second, equalising standard deviations across indicators means losing information about the actual variability of the indicators across the different countries, seen as valuable information in its own right.

5.12.6. Assessment of the correlation structure and refinement of the measurement framework

To refine the measurement framework, initial tests were performed using a limited choice of methods to assess the correlation structure. These tests were conducted using a symmetric linear metric for shares indicators, equal weights and the arithmetic mean. Detailed results of these tests can be found in Annex 4.

The assessment of the correlation structure is an iterative process, examining different solutions and seeking refinements to the measurement framework. The procedure used is an adaptation of the audit methodology developed by the JRC's Competence Centre on Composite Indicators and Scoreboards (JRC-COIN) ([290]) and relies on the following checks:

([290]) For more information, see: https://knowledge4policy.ec.europa.eu/composite-indicators/about_en

- Examining the correlations between indicators with the same dimension with ideally moderate to strong positive correlations coefficients (i.e. typically 0.5 to 0.8). This confirms that these indicators represent a single underlying construct, building upon the PCA analysis.

- Ensuring that there are no strong negative correlations (i.e. below -0.4) between indicators in a given dimension, and indicators in another dimension. This ensures that there is no trade-off between dimensions, which would hinder the interpretability of the scores of the Index.

- Examining the correlations between indicators and the scores for their respective dimensions, to ensure that they are sufficiently contributing (i.e. above 0.5, and ideally no less than 0.4).

- Examining the correlations between indicators, dimensions and Index scores, to ensure that they contribute substantially to the final Index score (i.e. above 0.5, and ideally no less than 0.4).

- Screen for very high correlations (i.e. 0.9 and above) among indicators, which may signal that two indicators are so strongly correlated that do in fact represent only a single underlying construct.

These tests revealed the following five issues:

- A moderate negative correlation between the share of women and men as doctoral graduates in Arts and Humanities and the share of women and men in Grades A and B positions.

- Low contribution of the indicator measuring the share of women and men in Grade C positions to the score of the career progression dimension.

- Marginal contribution of the indicator measuring the shares of women and men as doctoral graduates in Business and Law and as Member and Leaders of boards to the Index score.

- No contribution of the two indicators measuring the extent to which the sex and gender dimension has been included in publication and Horizon projects; and the indicator measuring the share of women and men in Arts and Humanities to the Index score.

- No contribution of the gender dimension in R&I content to the Index score in the case where M100 or M10 is used (M100 and M10 are the Min-Max with 100% and 10% as upper bound, respectively), though there is a contribution when the MM method is used (MM is the Min-Max in its basic form, i.e. with observed minimum and maximum). Consequently, this suggested that the most suitable method for the metrics capturing the gender dimension in R&I content is the MM method.

In the subsequent iteration, the correlation structure was reassessed after excluding indicators measuring the share of women and men as doctoral students in Arts and Humanities, as well as in Business and Law. Both indicators are related to the same underlying construct, with no or low contributions to the Index score. Further, the removal of these indicators mitigates issues related to negative correlations with other indicators.

Following this adjustment, another iteration was conducted wherein the indicator measuring the share of women and men in Grade C positions was removed. The results confirmed that, empirically, there is a stronger correlation structure after this indicator is omitted.

In the final iteration, indicators related to the gender and intersectional dimensions in Horizon projects were removed. This decision was guided by a broader interpretation of indicators based on publications, rather than Horizon projects alone, as these projects represent only a small proportion of R&I content in Europe. The final iteration provides a suitable correlation structure, with most indicators contributing strongly to all dimensions and to the Index scores.

These iterative refinements provided a solid basis for the She Figures Index measurement framework. The She Figures Index relies on a final set of 14 indicators (see table below), distributed across six dimensions relevant to gender equality in the context of R&I in Europe.

Table 5 Final list of indicators for the She Figures Index

Segregation in the pipeline	Research careers and sectors	Career progression	Representation in decision-making positions	Research participation	Gender dimension in Research and Innovation Content (GDRIC)
Proportion (%) of women among doctoral graduates, by broad field of study (Information and Communication Technologies)	Proportion of women researchers in the higher education sector	Proportion (%) of women among academic staff, by grade and total (Grade A)	Proportion (%) of women among heads of institutions in the Higher Education Sector (HES)	Average proportion (%) of women among authors on publications in all fields of R&D	Proportion (%) of a country's publications with a sex or gender dimension in their research content
Proportion (%) of women among doctoral graduates, by broad field of study (Engineering, Manufacturing and Construction)	Proportion of women researchers in the government sector	Proportion (%) of women among academic staff, by grade and total (Grade B)	Proportion (%) of women on boards, leaders	Distribution of patent application by sex composition of the inventors' team (%)	
	Proportion of women researchers in the business enterprise sector		Proportion (%) of women on boards, members and leaders	Research funding success rate differences between women and men	

5.12.7. Multi-modelling results

To construct the She Figures Index, a multi-modelling approach was considered, using the data from the latest She Figures report at the time of writing. The principle of multi-modelling is that when faced with uncertainty, computing all possible versions can guide the choice of the final methodology. This approach not only strengthens the methodological robustness of the Index but also increases its acceptability, as it indicates that the results are not determined by a single methodological choice in building the Index.

The multi-modelling analysis examined the effects of aggregation methods on the scores and ranks of the She Figures Index. Findings indicated that aggregation methods which do not permit full compensation disproportionately penalise Member States with lower performance compared to their counterparts. The option of using the arithmetic mean as a method of aggregation is thus deemed preferable.

The multi-modelling analysis has also shown that the indicator measuring the gender dimension in R&I content not only relies on a different metric, but also introduces a lot of variability in the She Figures Index scores and ranks. For this reason, using expert weights that amplify these fluctuations is not advisable, and equal weights are recommended. Full details regarding the multi-modelling analysis are provided in Annex 4.

5.12.8. Robustness checks

To check the robustness of the Index, assessments of the correlation structure and of the Index scores and ranks by dimension were performed.

Assessment of the correlation structure

The assessments of the correlation structure for the 2021 and 2024 editions of the She Figures Index (as found in Annex 4) show that the structure of the She Figures Index is adequate, as there are no strong negative correlations and indicators contribute to their respective dimensions and the Index overall.

Moderate to high positive correlations exist:

- Among indicators within the same dimensions.
- Between indicators and their respective dimensions.
- Between indicators and the aggregate Index score.
- Across different dimensions.

- Between dimension scores and overall index scores.

There are no strong negative correlations.

Assessment of robustness of the scores and ranks by dimension

A robustness check was performed to gauge the contribution of each dimension. This is done by subsequently removing each dimension in the computation of the She Figures Index and assessing the impact in terms of scores and ranks (see Annex 4 for details). This assessment shows that there is relative homogeneity in scores and ranks when the first five dimensions are removed from the calculation.

However, notable variations emerge when excluding the gender dimension in R&I content; its removal significantly impacts final scores and ranks. This insight suggests that countries can make most progress by addressing gender equality in that area.

6. Quality plan: verification and validation of data

In preparing the present study, data quality was viewed as a multi-faceted concept. The quality framework deployed covered four different dimensions to be considered in selecting indicators: relevance, accuracy, cross-country comparability and availability (Table 6). Each indicator was to be evaluated by grading it for each dimension and by an overall assessment.

The relevance of an indicator was determined by a qualitative assessment of the value contributed by that indicator in terms of its policy relevance. An indicator had to be policy relevant by addressing key policy issues related to gender equality in Research and Innovation (R&I) in the EU Member States and Associated Countries.

The accuracy of an indicator is the degree to which the indicator correctly estimates or describes the quantities or characteristics it is designed to measure. The data collection method was considered sound if the data correctly estimated or described the quantities or characteristics that it was designed to measure. Thus, accuracy based on the data collection method refers to the closeness between the values provided and the (unknown) true value.

The evaluation of the accuracy of data collection methods was significant in the present study, given that the data used had to be collected not only from high-quality databases of national statistical offices and international organisations but also from other databases, via web-scraping, as well as directly from Statistical Correspondents (i.e., indirectly from the data sources that they used). The accuracy of data collection methods in the present study can be evaluated as being very good, good and acceptable.

Cross-country comparability refers to whether an indicator is comparable across countries. Its assessment required consideration of the methods of data collection and concepts adopted in the countries concerned. For example, an indicator was comparable if the same question was asked in all countries in the same way and by the same means. It was desirable to have the highest degree of comparability across countries. For data collected through the Statistical Correspondents, guidelines were prepared to maximise cross-country comparability. Metadata were also collected with the WiS questionnaire from every participating country to allow an assessment of comparability. Additionally, much attention has been paid to ensuring data quality by regular consultation with Statistical Correspondents throughout the process of data collection.

The concept of availability relates to the accessibility of a given indicator in various countries and for a given time frame. For the present study, it was desirable to have data from as many EU Member States and Associated Countries as possible. In addition, the more reference years an indicator was available for, the better its availability was considered.

Table 6 Dimensions of the data quality framework

	Depends on	Addressed by
RELEVANCE	Relevance of selected indicators in the current models and measuring systems of R&I : are they up-to-date from a policy perspective?	Steering group members, who discussed the relevance of existing and newly proposed indicators in two steering group meetings Mapping of state of the art with respect to R&I indicators Identification of new indicators to introduce in She Figures 2024
ACCURACY	Data collection method used	Rely whenever possible on existing official classifications and manuals for data collection (e.g. Frascati Manual, etc.), international standards, etc. Guidelines and aiming to have Statistical Correspondents adhere as much as possible to quality standards of data collection Validity/coherence checks after data gathering and computation of confidence intervals (for certain indicators)
COMPARABILITY	Alignment between countries in reporting system, classifications used, etc., by data source	Rely whenever possible on existing official classifications and manuals for data collection (e.g. Frascati Manual, etc.), international standards, etc. Guidelines addressed to the have Statistical Correspondents aiming at eliminating the room for misinterpretation of the data requested Close collaboration of the project team with the Statistical Correspondents in order to clarify any aspects that caused difficulties to any of them Metadata sheets (to systematically register potential deviations from the defined classifications and standards)
AVAILABILITY	Capacity and resources of governments to collect the required information Availability of additional data sources next to official statistical data sources	Steering group members, who discussed the availability of existing and newly proposed indicators in two steering group meetings; in what concerns new indicators, the group members relied on assessments of data availability prepared by the project team Information and feedback gathered while producing earlier editions of She Figures Flagging system (to systematically register missing data)

At the same time, ensuring maximum quality and reliability of the compiled data requires an a posteriori verification and validation of the data received.

6.1. Coherence checks

For data broken down in categories, totals are also available. Categories are defined regarding:

- Sex, which refers to the given sex at birth within the data used by She Figures
- Age groups (see later in this section ‘Additional data considerations’)
- Institutional sectors (see main grouping as defined by the Frascati Manual)
- NACE activities (see NACE Rev. 2.0 categories under Section 5.5.11)

- Fields of Research and Development ([291])
- Education levels and fields of education ([292]) ([293])
- Grades (see definition of grades in Annex 4)
- ISCO-08 categories ([294])
- R&D personnel categories
- Countries
- Years

Having available both the broken-down data and the totals allowed for coherence checks by comparing provided totals with the sum of provided data by categories. For example, a check which was performed on most of the tables was that the sum of the values for women (w) and men (m) corresponds to the reported totals (t), hence the verification was done by applying formula t = w + m. A similar data verification procedure was followed to assess whether reported totals corresponded to the sums of breakdowns at the level of the above-mentioned categorisations. Table 7 below shows the details of these coherence checks of WiS data.

Table 7 List of coherence checks on the WiS data

Data	Verification Description	Verification Formula
Tables T1 (HC) and T2 (FTE): RESEARCHERS and ACADEMIC STAFF BY SEX, GRADE AND MAIN FIELD OF RESEARCH AND DEVELOPMENT AND AGE GROUP (FOR LATEST YEAR ONLY)	Check Total sexes	{Men} + {Women} = Total
	Check Total age groups	{<35} + {35–44} + {45–54} + {55+} = Total
	Check Total fields of Research and Development	{NS} + {ET} + {MS} + {AS} + {SS} + {H} + {Unknown} = Total per grade
	Check Total grades	{A} + {B} + {C} + {D} = ALL grades
Tables T3 (Team leaders) and T4 (Team members): APPLICANTS AND BENEFICIARIES OF PUBLICLY MANAGED RESEARCH FUNDS BY SEX, MAIN FIELD OF RESEARCH AND DEVELOPMENT, AMOUNT APPLIED FOR AND AMOUNT RECEIVED	Check Total sexes	{Men} + {Women} = Total
	Check Total fields of Research and Development	{NS} + {ET} + {MS} + {AS} + {SS} + {H} + {MU} + {Unknown} = Total per fund
	Check Comparison applicants and beneficiaries	# applicants > # beneficiaries
Tables T5 (scientific boards) and T6 (administrative / advisory boards): PRESIDENTS / LEADERS AND MEMBERS OF BOARDS BY SEX,	Check Total sexes	{Men} + {Women} = Total
	Check Total fields of Research and Development	{NS} + {ET} + {MS} + {AS} + {SS} + {H} + {Unknown} = Total per board

([291]) OECD, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, 2015. http://dx.doi.org/10.1787/9789264239012-en

([292]) UNESCO, *International standard classification of education*, 2012. http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

([293]) UNESCO, *ISCED Fields of Education and Training 2013*, 2014. http://uis.unesco.org/sites/default/files/documents/isced-fields-of-education-and-training-2013-en.pdf

([294]) International Labour Organization, *International classification of occupations (ISCO-08)*, 2012. http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_172572.pdf

Data	Verification Description	Verification Formula
FIELD OF RESEARCH AND DEVELOPMENT AND POSITION	Check coherence of president and member counts with the respective data in EIGE's 'Research funding organisations: presidents and members of the highest decision-making body' ([295])	
Tables T7 (institutions) and T8 (universities): HEADS OF INSTITUTIONS IN THE HES BY SEX & HEADS OF UNIVERSITIES OR ASSIMILATED INSTITUTIONS (BASED ON CAPACITY TO DELIVER PhDS) BY SEX	Check Total sexes	{Men} + {Women} = Total
	Check Number of institutes	# institutes T7 ≥ # institutes T8

Note that a final visual scan of the formatted tables/charts appearing in the main publication was performed in the end to detect any inconsistencies that were possibly overlooked in previous validation steps.

6.2. Additional data considerations

Age groups

Data referring to the labour force refer to all persons aged 15+ living in private households and include the employed and the unemployed. Data referring to research personnel, to the working conditions of researchers in the HES (or to a target population used as a proxy) and to human resources in science and technology (HRST) refer to the age group 25–64.

Small numbers

For some countries with small populations, raw data relating to small numbers of people have been compiled. The percentages and indicators have not always been included (mostly growth rates) and this is identified in the footnotes to the indicators. The reader is therefore asked to bear this in mind when interpreting the most data at their most disaggregated level.

EU estimates

EU totals estimated by DG Research and Innovation (as noted in the footnotes) are based upon existing data for the reference year in combination with the next available year if the reference year is unavailable, in the following sequence (n-1, n+1, n-2, n+2 etc.).

The aggregates were estimated by DG Research and Innovation only when at least 60 % of the EU population on a given indicator was available. *These estimates are intended only as an indication for the reader.*

Rounding error

In some cases, the row or column totals do not match the sum of the data. This may be due to rounding error.

Decimal places

All the data in the figures have been calculated at the precision levels of one or two decimals. However, the values have been rounded in the figures (typically to one decimal place) and in text (to two significant figures) to support comprehension.

([295]) EIGE, Gender Statistics Database, Research funding organisations: presidents and members of the highest decision-making body

Cut-off date

The cut-off date for data downloaded from Eurostat's dissemination database (Eurostat) was the 29th of November 2023. Due to the large variety of data sources and variability in data availability, some other cut-off dates were used in order to gather all the required data. Details on the cut-off dates are included in She Figures 2024 (main report).

7. Annexes

Annex 1: Definitions of key terms

In its understanding of 'gender' and 'gender equality', She Figures builds upon the definitions developed by UN Women (Concepts and definitions) ([296]):

Gender refers to 'the social attributes and opportunities associated with being men and women and the relationships between women and men and girls and boys, as well as the relations between women and those between men. These attributes, opportunities and relationships are socially constructed and are learned through socialisation processes. They are context/time-specific and changeable. Gender determines what is expected, allowed and valued in a woman or a man in a given context ...'. **Non-binary** is an umbrella term for gender identities that fall outside the gender binary of categories of women or men. This includes individuals whose gender identity is neither exclusively woman nor man, a combination of woman and man or between or beyond genders (Trans Equality Network Ireland).

Equality between women and men (gender equality) refers to the 'equal rights, responsibilities and opportunities of women and men and girls and boys. Equality does not mean that women and men will become the same but that women's and men's rights, responsibilities and opportunities will not depend on whether they are born male or female. Gender equality implies that the interests, needs and priorities of both women and men are taken into consideration, recognizing the diversity of different groups of women and men ...'.

Whilst sex and gender are often used interchangeably, they are not the same. In general, the She Figures project understands sex to be a biological category, whilst gender relates to historical, cultural and social realities. For example, when data are broken down to show the individual data for women and men, these are understood to be *sex-disaggregated* data, and not *gender-disaggregated* data.

Within the fields of education, research and innovation, there are a range of additional terms useful for measuring gender equality (which is understood to be a multi-dimensional concept). Many of these relate to the notion of 'segregation'. Definitions of each of these are given below, based on those discussed by the European Commission's Expert Group on Gender and Employment ([297]):

Gender segregation in the labour market refers to the gendered division of labour in employment. It is a broad term, describing the tendency for women and men to work in different occupations, sectors, fields, etc. It is often associated with potentially negative effects, including narrowed choice for women and men, perpetuation of gender stereotypes, vertical segregation (see below) and finally, the under-valuing of skills and abilities linked to women's work (affecting their pay). Since the 1960s, a range of additional terms have emerged to understand gender segregation more fully, including horizontal, vertical, sectoral and occupational segregation.

Horizontal segregation relates to the concentration of women and men around different sectors (sectoral segregation) and occupations (occupational segregation). It can occur within both education (e.g. over-/under-representation of one sex in particular subjects) and employment (e.g. over-/under-representation of one sex in particular professions, industries, etc.). Unlike vertical segregation, these occupations and sectors

([296]) As found here: https://www.un.org/womenwatch/osagi/conceptsandefinitions.htm

([297]) European Commission's Expert Group on Gender and Employment (EGGE), *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*, Publications Office of the European Union, Luxembourg, 2009. http://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=364&furtherPubs=yes

are not ordered by a particular criterion. However, the issue of horizontal segregation may in turn lead to greater vertical segregation. For example, the under-valuing of capacities associated with 'women's work' may limit women's prospects for career advancement.

Vertical segregation refers to the concentration of either men or women in 'top' posts or positions of responsibility. Such roles are often associated with 'desirable' features, including greater pay, prestige and security. In the context of R&I, the over-representation of men amongst heads of universities is an example of such segregation. Below is the list of positions used in producing the She Figures publication:

Grade A: The single highest grade / post at which research is normally conducted within the institutional or corporate system.

Grade B: All researchers working in positions which are not as senior as the top position (A) but definitely more senior than the newly qualified PhD holders (C); i.e. below A and above C.

Grade C: The first grade/post into which a newly qualified PhD (ISCED 8) graduate would normally be recruited within the institutional or corporate system.

Grade D: Either postgraduate students not yet holding a PhD (ISCED 8) degree who are engaged as researchers (on the payroll) or researchers working in posts that do not normally require a PhD.

A table with the qualifications for each grade by country is presented in the Appendix 2 of the main publication of She Figures.

She Figures 2015 introduced new definitions of 'boards' as part of the Women in Science questionnaire, based on consultation with the European Commission and the Statistical Correspondents. These distinguish more clearly between the functions of different boards, by focusing on 'scientific boards' and 'administrative/advisory boards':

Scientific board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to implement scientific policy by, among other things, directing the research agenda, resource allocation and management within scientific research.

Administrative / advisory board of research organisation: A publicly or privately managed and financed group of elected or appointed experts that exists to support the research agenda in a non-executive function by, among other things, administering research activities, consulting and coordinating different actors and taking a general advisory role.

Where boards fall into both categories, this was indicated by Statistical Correspondents. She Figures includes only research boards of umbrella, national-level research performing organisations (RPOs) and research funding organisations (RFOs), as opposed to all research organisations operating in a particular country. Whilst data were collected separately for the two types of boards, this indicator remained combined in one indicator in the publication.

Gender dimension in research and innovation is a concept that 'implies analysing and taking into account the possible differences between men and women (biological characteristics as well as the social and cultural features), boys and girls, or males and females, in the R&I content of the project' ([298]). The European Commission established the Gendered Innovations expert group that produced reports on incorporating the gender dimension across various research fields, defining research integrating the 'gender dimension' as research that integrates sex and gender analysis into their content. 'Sex' refers to basic biological characteristics of females and males and 'gender' refers to cultural attitudes and behaviours that shape 'feminine' and 'masculine' behaviours, products, technologies, environments, and knowledge ([299]) ([300]).

([298]) European Commission, European Research Executive Agency. *Gender in EU research and innovation*, 2023.

([299]) European Commission, *Gendered Innovations 2: How Inclusive Analysis contributes to Research and Innovation*, 2020. https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en

([300]) European Commission, *Gendered Innovations – How gender analysis contributes to research: report of the expert group 'Innovation through gender'*, 2013. https://op.europa.eu/en/publication-detail/-/publication/d15a85d6-cd2d-4fbc-b998-42e53a73a449

Annex 2: Index list of indicators

Annex 3: Correspondence of ASJC sub-categories with Fields of Research and Development (FORD)

The Fields of Research and Development (FORD) classifications were mapped with the ASJC sub-categories from Scopus [TM] according to the following mapping.

The classifications were mapped at category level (ASJC category to target category).

Table 8 Dimensions of the data quality framework

ASJC code	ASJC name	FORD code	FORD Name
1100	General Agricultural and Biological Sciences	4.1	Agriculture, forestry, and fishery
1101	Agricultural and Biological Sciences (miscellaneous)	4.5	Other agricultural sciences
1102	Agronomy and Crop Science	4.1	Agriculture, forestry, and fishery
1103	Animal Science and Zoology	1.6	Biological sciences
1103	Animal Science and Zoology	4.2	Biological sciences | Animal and dairy science
1104	Aquatic Science	1.6	Biological Sciences
1104	Aquatic Science	4.1	Agriculture, forestry, and fishery
1105	Ecology, Evolution, Behaviour and Systematics	1.6	Biological sciences
1105	Ecology, Evolution, Behaviour and Systematics	1.5	Earth and related environmental sciences
1106	Food Science	4.2	Animal and dairy science
1107	Forestry	4.1	Agriculture, forestry, and fishery
1108	Horticulture	4.1	Agriculture, forestry, and fishery
1109	Insect Science	1.6	Biological sciences
1110	Plant Science	1.6	Biological sciences
1111	Soil Science	4.1	Agriculture, forestry, and fishery
1200	General Arts and Humanities	6.5	Other humanities
1201	Arts and Humanities (miscellaneous)	6.5	Other humanities
1202	History	6.1	History and archaeology
1203	Language and Linguistics	6.2	Languages and literature
1204	Archaeology (arts and humanities)	6.1	History and archaeology
1205	Classics	6.2	Languages and literature
1206	Conservation	6.5	Other humanities

ASJC code	ASJC name	FORD code	FORD Name
1207	History and Philosophy of Science	6.5	Other humanities
1208	Literature and Literary Theory	6.2	Languages and literature
1209	Museology	6.5	Other humanities
1210	Music	6.4	Arts (arts, history of arts, performing arts, music)
1211	Philosophy	6.3	Philosophy, ethics and religion
1212	Religious Studies	6.3	Philosophy, ethics and religion
1213	Visual Arts and Performing Arts	6.4	Arts (arts, history of arts, performing arts, music)
1300	General Biochemistry, Genetics and Molecular Biology	1.6	Biological sciences
1301	Biochemistry, Genetics and Molecular Biology (miscellaneous)	1.6	Biological sciences
1302	Aging	1.6	Biological sciences
1302	Aging	3.1	Basic medicine
1303	Biochemistry	1.6	Biological sciences
1304	Biophysics	1.6	Biological sciences
1305	Biotechnology	1.6	Biological sciences
1306	Cancer Research	1.6	Biological science
1306	Cancer Research	3.1	Basic medicine
1307	Cell Biology	1.6	Biological sciences
1308	Clinical Biochemistry	3.2	Clinical medicine
1309	Developmental Biology	1.6	Biological sciences
1310	Endocrinology	1.6	Biological sciences
1311	Genetics	1.6	Biological sciences
1312	Molecular Biology	1.6	Biological sciences
1313	Molecular Medicine	1.6	Biological science
1313	Molecular Medicine	3.1	Basic medicine
1314	Physiology	1.6	Biological science
1314	Physiology	3.1	Basic medicine
1315	Structural Biology	1.6	Biological sciences

ASJC code	ASJC name	FORD code	FORD Name
1400	General Business, Management and Accounting	5.2	Economics and business
1401	Business, Management and Accounting (miscellaneous)	5.2	Economics and business
1402	Accounting	5.2	Economics and business
1403	Business and International Management	5.2	Economics and business
1404	Management Information Systems	5.2	Economics and business
1405	Management of Technology and Innovation	5.2	Economics and business
1406	Marketing	5.2	Economics and business
1407	Organizational Behaviour and Human Resource Management	5.2	Economics and business
1408	Strategy and Management	5.2	Economics and business
1409	Tourism, Leisure and Hospitality Management	5.2	Economics and business
1410	Industrial Relations	5.2	Economics and business
1500	General Chemical Engineering	2.4	Chemical engineering
1501	Chemical Engineering (miscellaneous)	2.4	Chemical engineering
1502	Bioengineering	2.11	Other engineering and technologies
1503	Catalysis	1.4	Chemical Sciences
1503	Catalysis	2.4	Chemical engineering
1504	Chemical Health and Safety	2.4	Chemical engineering
1505	Colloid and Surface Chemistry	1.4	Chemical Sciences
1505	Colloid and Surface Chemistry	2.4	Chemical sciences | Chemical engineering
1506	Filtration and Separation	1.4	Chemical Sciences
1506	Filtration and Separation	2.4	Chemical sciences | Chemical engineering
1507	Fluid Flow and Transfer Processes	1.3	Physical sciences
1507	Fluid Flow and Transfer Processes	1.4	Chemical Sciences
1507	Fluid Flow and Transfer Processes	2.4	Chemical engineering
1508	Process Chemistry and Technology	2.4	Chemical engineering

ASJC code	ASJC name	FORD code	FORD Name
1600	General Chemistry	1.4	Chemical sciences
1601	Chemistry (miscellaneous)	1.4	Chemical sciences
1602	Analytical Chemistry	1.4	Chemical sciences
1603	Electrochemistry	1.4	Chemical sciences
1604	Inorganic Chemistry	1.4	Chemical sciences
1605	Organic Chemistry	1.4	Chemical sciences
1606	Physical and Theoretical Chemistry	1.4	Chemical sciences
1607	Spectroscopy	1.4	Chemical sciences
1700	General Computer Science	1.2	Computer and information sciences
1701	Computer Science (miscellaneous)	1.2	Computer and information sciences
1702	Artificial Intelligence	1.2	Computer and information sciences
1703	Computational Theory and Mathematics	1.2	Computer and information sciences
1704	Computer Graphics and Computer -Aided Design	1.2	Computer and information sciences
1705	Computer Networks and Communications	1.2	Computer and information sciences
1706	Computer Science Applications	1.2	Computer and information sciences
1707	Computer Vision and Pattern Recognition	1.2	Computer and information sciences
1708	Hardware and Architecture	1.2	Computer and information sciences
1709	Human-Computer Interaction	1.2	Computer and information sciences
1710	Information Systems	1.2	Computer and information sciences
1711	Signal Processing	1.2	Computer and information sciences
1712	Software	1.2	Computer and information sciences
1800	General Decision Sciences	5.9	Other social sciences
1801	Decision Sciences (miscellaneous)	5.9	Other social sciences
1802	Information Systems and Management	1.2	Computer and information sciences
1803	Management Science and Operations Research	5.9	Other social sciences
1804	Statistics, Probability and Uncertainty	1.1	Mathematics

ASJC code	ASJC name	FORD code	FORD Name
1900	General Earth and Planetary Sciences	1.5	Earth and related environmental sciences
1901	Earth and Planetary Sciences (miscellaneous)	1.5	Earth and related environmental sciences
1902	Atmospheric Science	1.5	Earth and related environmental sciences
1903	Computers in Earth Sciences	1.5	Earth and related environmental sciences
1904	Earth-Surface Processes	1.5	Earth and related environmental sciences
1905	Economic Geology	1.5	Earth and related environmental sciences
1906	Geochemistry and Petrology	1.5	Earth and related environmental sciences
1907	Geology	1.5	Earth and related environmental sciences
1908	Geophysics	1.5	Earth and related environmental sciences
1909	Geotechnical Engineering and Engineering Geology	1.5	Earth and related environmental sciences
1910	Oceanography	1.5	Earth and related environmental sciences
1911	Palaeontology	1.5	Earth and related environmental sciences
1912	Space and Planetary Science	1.5	Earth and related environmental sciences
1913	Stratigraphy	1.5	Earth and related environmental sciences
2000	General Economics, Econometrics and Finance	5.2	Economics and business
2001	Economics, Econometrics and Finance (miscellaneous)	5.2	Economics and business
2002	Economics and Econometrics	5.2	Economics and business
2003	Finance	5.2	Economics and business
2100	General Energy	1.7	Other natural sciences
2101	Energy (miscellaneous)	1.7	Other natural sciences
2102	Energy Engineering and Power Technology	1.7	Other natural sciences
2103	Fuel Technology	1.7	Other natural sciences

ASJC code	ASJC name	FORD code	FORD Name
2103	Fuel Technology	2.7	Environmental engineering
2104	Nuclear Energy and Engineering	2.11	Other engineering and technologies
2105	Renewable Energy, Sustainability and the Environment	1.5	Earth and related environmental sciences
2105	Renewable Energy, Sustainability and the Environment	1.7	Other natural sciences
2105	Renewable Energy, Sustainability and the Environment	2.7	Environmental engineering
2200	General Engineering	2.11	Other engineering and technologies
2201	Engineering (miscellaneous)	2.11	Other engineering and technologies
2202	Aerospace Engineering	2.3	Mechanical Engineering
2203	Automotive Engineering	2.3	Mechanical Engineering
2204	Biomedical Engineering	2.6	Medical engineering
2204	Biomedical Engineering	3.4	Medical biotechnology
2205	Civil and Structural Engineering	2.1	Civil engineering
2206	Computational Mechanics	1.3	Physical sciences
2207	Control and Systems Engineering	2.11	Other engineering and technologies
2208	Electrical and Electronic Engineering	2.2	Electrical engineering, electronic engineering, information engineering
2209	Industrial and Manufacturing Engineering	2.3	Mechanical Engineering
2209	Industrial and Manufacturing Engineering	2.4	Chemical engineering
2209	Industrial and Manufacturing Engineering	2.11	Other engineering and technologies
2210	Mechanical Engineering	2.3	Mechanical Engineering
2211	Mechanics of Materials	2.5	Materials engineering
2212	Ocean Engineering	2.7	Environmental engineering
2213	Safety, Risk, Reliability and Quality	2.11	Other engineering and technologies
2214	Media Technology	2.11	Other engineering and technologies
2215	Building and Construction	2.1	Civil engineering
2216	Architecture	2.1	Civil engineering
2300	General Environmental Science	1.5	Earth and related environmental sciences

ASJC code	ASJC name	FORD code	FORD Name
2301	Environmental Science (miscellaneous)	1.5	Earth and related environmental sciences
2302	Ecological Modelling	1.5	Earth and related environmental sciences
2303	Ecology	1.5	Earth and related environmental sciences
2304	Environmental Chemistry	1.5	Earth and related environmental sciences
2304	Environmental Chemistry	1.4	Chemical sciences
2305	Environmental Engineering	1.5	Earth and related environmental sciences
2306	Global and Planetary Change	1.5	Earth and related environmental sciences
2307	Health, Toxicology and Mutagenesis	1.6	Biological sciences
2307	Health, Toxicology and Mutagenesis	3.1	Biological sciences | Basic medicine
2308	Management, Monitoring, Policy and Law	5.5	Law
2308	Management, Monitoring, Policy and Law	5.6	Political Sciences
2309	Nature and Landscape Conservation	1.5	Earth and related environmental sciences
2310	Pollution	1.5	Earth and related environmental sciences
2311	Waste Management and Disposal	1.5	Earth and related environmental sciences
2312	Water Science and Technology	1.5	Earth and related environmental sciences
2400	General Immunology and Microbiology	1.6	Biological sciences
2400	General Immunology and Microbiology	3.1	Basic medicine
2401	Immunology and Microbiology (miscellaneous)	1.6	Biological sciences
2401	Immunology and Microbiology (miscellaneous)	3.1	Basic medicine
2402	Applied Microbiology and Biotechnology	1.6	Biological sciences
2402	Applied Microbiology and Biotechnology	2.9	Industrial biotechnology

ASJC code	ASJC name	FORD code	FORD Name
2402	Applied Microbiology and Biotechnology	3.1	Basic Medicine
2403	Immunology	1.6	Biological sciences
2403	Immunology	3.1	Basic medicine
2404	Microbiology	1.6	Biological sciences
2404	Microbiology	3.1	Basic medicine
2405	Parasitology	1.6	Biological sciences
2405	Parasitology	3.1	Basic medicine
2406	Virology	1.6	Biological sciences
2406	Virology	3.1	Basic medicine
2500	General Materials Science	1.7	Other natural sciences
2500	General Materials Science	2.5	Materials engineering
2501	Materials Science (miscellaneous)	1.7	Other natural sciences
2501	Materials Science (miscellaneous)	2.5	Materials engineering
2502	Biomaterials	1.7	Other natural sciences
2502	Biomaterials	2.5	Materials engineering
2502	Biomaterials	2.6	Medical engineering
2503	Ceramics and Composites	1.4	Chemical sciences
2503	Ceramics and Composites	2.5	Materials engineering
2504	Electronic, Optical and Magnetic Materials	1.4	Chemical sciences
2504	Electronic, Optical and Magnetic Materials	2.5	Materials engineering
2505	Materials Chemistry	1.4	Chemical sciences
2505	Materials Chemistry	2.5	Materials engineering
2506	Metals and Alloys	1.4	Chemical sciences
2506	Metals and Alloys	2.5	Materials engineering
2507	Polymers and Plastics	1.4	Chemical sciences
2507	Polymers and Plastics	2.5	Materials engineering
2508	Surfaces, Coatings and Films	1.4	Chemical sciences
2508	Surfaces, Coatings and Films	2.5	Materials engineering

ASJC code	ASJC name	FORD code	FORD Name
2600	General Mathematics	1.1	Mathematics
2601	Mathematics (miscellaneous)	1.1	Mathematics
2602	Algebra and Number Theory	1.1	Mathematics
2603	Analysis	1.1	Mathematics
2604	Applied Mathematics	1.1	Mathematics
2605	Computational Mathematics	1.1	Mathematics
2606	Control and Optimization	1.1	Mathematics
2607	Discrete Mathematics and Combinatorics	1.1	Mathematics
2608	Geometry and Topology	1.1	Mathematics
2609	Logic	1.1	Mathematics
2610	Mathematical Physics	1.1	Mathematics
2611	Modelling and Simulation	1.1	Mathematics
2612	Numerical Analysis	1.1	Mathematics
2613	Statistics and Probability	1.1	Mathematics
2614	Theoretical Computer Science	1.1	Mathematics
2614	Theoretical Computer Science	1.2	Computer and information sciences
2700	General Medicine	3.1	Basic medicine
2701	Medicine (miscellaneous)	3.1	Basic medicine
2702	Anatomy	3.1	Basic medicine
2703	Anesthesiology and Pain Medicine	3.1	Basic medicine
2704	Biochemistry (medical)	3.2	Clinical medicine
2705	Cardiology and Cardiovascular Medicine	3.2	Clinical medicine
2706	Critical Care and Intensive Care Medicine	3.2	Clinical medicine
2707	Complementary and Alternative Medicine	3.2	Clinical medicine
2708	Dermatology	3.2	Clinical medicine
2709	Drug Guides	3.2	Clinical medicine
2710	Embryology	3.2	Clinical medicine
2711	Emergency Medicine	3.2	Clinical medicine

ASJC code	ASJC name	FORD code	FORD Name
2712	Endocrinology, Diabetes and Metabolism	3.2	Clinical medicine
2713	Epidemiology	3.2	Clinical medicine
2714	Family Practice	3.3	Health sciences
2715	Gastroenterology	3.2	Clinical medicine
2716	Genetics (clinical)	3.1	Basic medicine
2717	Geriatrics and Gerontology	3.2	Clinical medicine
2718	Health Informatics	3.3	Health sciences
2719	Health Policy	3.3	Health sciences
2720	Haematology	3.2	Clinical medicine
2721	Hepatology	3.2	Clinical medicine
2722	Histology	3.2	Clinical medicine
2723	Immunology and Allergy	3.2	Clinical medicine
2724	Internal Medicine	3.2	Clinical medicine
2725	Infectious Diseases	3.2	Clinical medicine
2726	Microbiology (medical)	3.2	Clinical medicine
2727	Nephrology	3.2	Clinical medicine
2728	Neurology (clinical)	3.2	Clinical medicine
2729	Obstetrics and Gynaecology	3.2	Clinical medicine
2730	Oncology	3.2	Clinical medicine
2731	Ophthalmology	3.2	Clinical medicine
2732	Orthopaedics and Sports Medicine	3.2	Clinical medicine
2733	Otorhinolaryngology	3.2	Clinical medicine
2734	Pathology and Forensic Medicine	3.1	Basic medicine
2735	Paediatrics, Perinatology and Child Health	3.2	Clinical medicine
2736	Pharmacology (medical)	3.2	Clinical medicine
2737	Physiology (medical)	3.2	Clinical medicine
2738	Psychiatry and Mental Health	3.2	Clinical medicine
2739	Public Health, Environmental and Occupational Health	3.3	Health sciences

ASJC code	ASJC name	FORD code	FORD Name
2740	Pulmonary and Respiratory Medicine	3.2	Clinical medicine
2741	Radiology, Nuclear Medicine and Imaging	3.2	Clinical medicine
2742	Rehabilitation	3.3	Health sciences
2743	Reproductive Medicine	3.2	Clinical medicine
2743	Reproductive Medicine	3.3	Health sciences
2744	Reviews and References (medical)	3.2	Clinical medicine
2744	Reviews and References (medical)	3.3	Health sciences
2745	Rheumatology	3.2	Clinical medicine
2746	Surgery	3.2	Clinical medicine
2747	Transplantation	3.2	Clinical medicine
2748	Urology	3.2	Clinical medicine
2800	General Neuroscience	3.1	Basic medicine
2801	Neuroscience (miscellaneous)	3.1	Basic medicine
2802	Behavioural Neuroscience	3.1	Basic medicine
2802	Behavioural Neuroscience	5.1	Psychology and cognitive sciences
2803	Biological Psychiatry	3.2	Clinical medicine
2804	Cellular and Molecular Neuroscience	3.1	Basic medicine
2805	Cognitive Neuroscience	3.1	Basic medicine
2805	Cognitive Neuroscience	5.1	Psychology and cognitive sciences
2806	Developmental Neuroscience	3.1	Basic medicine
2806	Developmental Neuroscience	5.1	Psychology and cognitive sciences
2807	Endocrine and Autonomic Systems	3.2	Clinical medicine
2808	Neurology	3.2	Clinical medicine
2809	Sensory Systems	3.1	Basic medicine
2809	Sensory Systems	3.4	Medical biotechnology
2900	General Nursing	3.3	Health sciences
2901	Nursing (miscellaneous)	3.3	Health sciences
2902	Advanced and Specialized Nursing	3.3	Health sciences

ASJC code	ASJC name	FORD code	FORD Name
2903	Assessment and Diagnosis	3.3	Health sciences
2904	Care Planning	3.3	Health sciences
2905	Community and Home Care	3.3	Health sciences
2906	Critical Care Nursing	3.3	Health sciences
2907	Emergency Nursing	3.3	Health sciences
2908	Fundamentals and Skills	3.3	Health sciences
2909	Gerontology	3.2	Clinical medicine
2909	Gerontology	3.3	Health sciences
2910	Issues, Ethics and Legal Aspects	3.3	Health sciences
2911	Leadership and Management	3.3	Health sciences
2912	LPN and LVN	3.3	Health sciences
2913	Maternity and Midwifery	3.3	Health sciences
2913	Maternity and Midwifery	5.1	Psychology and cognitive sciences
2914	Medical and Surgical Nursing	3.2	Clinical medicine
2914	Medical and Surgical Nursing	3.3	Health sciences
2915	Nurse Assisting	3.3	Health sciences
2916	Nutrition and Dietetics	3.3	Health sciences
2917	Oncology (nursing)	3.3	Health sciences
2918	Pathophysiology	3.2	Clinical medicine
2919	Paediatrics	3.2	Clinical medicine
2919	Paediatrics	3.3	Health sciences
2920	Pharmacology (nursing)	3.3	Health sciences
2921	Psychiatric Mental Health	3.3	Health sciences
2922	Research and Theory	3.3	Health sciences
2923	Review and Exam Preparation	3.3	Health sciences
3000	General Pharmacology, Toxicology and Pharmaceutics	3.1	Basic medicine
3001	Pharmacology, Toxicology and Pharmaceutics (miscellaneous)	3.1	Basic medicine
3002	Drug Discovery	3.1	Basic medicine
3003	Pharmaceutical Science	3.1	Basic medicine

ASJC code	ASJC name	FORD code	FORD Name
3004	Pharmacology	3.1	Basic medicine
3005	Toxicology	3.1	Basic medicine
3100	General Physics and Astronomy	1.3	Physical sciences
3101	Physics and Astronomy (miscellaneous)	1.3	Physical sciences
3102	Acoustics and Ultrasonics	1.3	Physical sciences
3103	Astronomy and Astrophysics	1.3	Physical sciences
3104	Condensed Matter Physics	1.3	Physical sciences
3105	Instrumentation	1.3	Physical sciences
3106	Nuclear and High Energy Physics	1.3	Physical sciences
3107	Atomic and Molecular Physics, and Optics	1.3	Physical sciences
3108	Radiation	1.3	Physical sciences
3109	Statistical and Nonlinear Physics	1.3	Physical sciences
3110	Surfaces and Interfaces	1.3	Physical sciences
3200	General Psychology	5.1	Psychology and cognitive sciences
3201	Psychology (miscellaneous)	5.1	Psychology and cognitive sciences
3202	Applied Psychology	5.1	Psychology and cognitive sciences
3203	Clinical Psychology	5.1	Psychology and cognitive sciences
3204	Developmental and Educational Psychology	5.1	Psychology and cognitive sciences
3205	Experimental and Cognitive Psychology	5.1	Psychology and cognitive sciences
3206	Neuropsychology and Physiological Psychology	5.1	Psychology and cognitive sciences
3207	Social Psychology	5.1	Psychology and cognitive sciences
3300	General Social Sciences	5.9	Other social sciences
3301	Social Sciences (miscellaneous)	5.9	Other social sciences
3302	Archaeology	6.1	History and archaeology
3303	Development	5.1	Psychology and cognitive sciences
3304	Education	5.3	Education
3305	Geography, Planning and Development	5.7	Social and economic geography

ASJC code	ASJC name	FORD code	FORD Name
3306	Health (social science)	3.3	Health sciences
3306	Health (social science)	5.9	Other social sciences
3307	Human Factors and Ergonomics	5.1	Psychology and cognitive sciences
3308	Law	5.5	Law
3309	Library and Information Sciences	5.9	Other social sciences
3310	Linguistics and Language	6.2	Languages and literature
3311	Safety Research	2.11	Other engineering and technologies
3312	Sociology and Political Science	5.4	Sociology
3312	Sociology and Political Science	5.6	Political Sciences
3313	Transportation	2.11	Other engineering and technologies
3314	Anthropology	5.4	Sociology
3315	Communication	5.8	Media and communications
3316	Cultural Studies	5.4	Sociology
3317	Demography	5.4	Sociology
3318	Gender Studies	5.4	Sociology
3319	Life-span and Life-course Studies	5.4	Sociology
3320	Political Science and International Relations	5.6	Political Sciences
3321	Public Administration	5.6	Political Sciences
3322	Urban Studies	5.4	Sociology
3322	Urban Studies	5.7	Social and economic geography
3400	General Veterinary	4.3	Veterinary science
3401	Veterinary (miscellaneous)	4.3	Veterinary science
3402	Equine	4.3	Veterinary science
3403	Food Animals	4.3	Veterinary science
3404	Small Animals	4.3	Veterinary science
3500	General Dentistry	3.2	Clinical medicine
3501	Dentistry (miscellaneous)	3.2	Clinical medicine
3502	Dental Assisting	3.2	Clinical medicine
3503	Dental Hygiene	3.2	Clinical medicine

ASJC code	ASJC name	FORD code	FORD Name
3504	Oral Surgery	3.2	Clinical medicine
3505	Orthodontics	3.2	Clinical medicine
3506	Periodontics	3.2	Clinical medicine
3600	General Health Professions	3.3	Health sciences
3601	Health Professions (miscellaneous)	3.3	Health sciences
3602	Chiropractics	3.2	Clinical medicine
3602	Chiropractics	3.3	Health sciences
3603	Complementary and Manual Therapy	3.3	Health sciences
3604	Emergency Medical Services	3.2	Clinical medicine
3604	Emergency Medical Services	3.3	Health sciences
3605	Health Information Management	3.3	Health sciences
3606	Medical Assisting and Transcription	3.3	Health sciences
3607	Medical Laboratory Technology	2.6	Medical engineering
3608	Medical Terminology	3.3	Health sciences
3609	Occupational Therapy	3.3	Health sciences
3610	Optometry	3.2	Clinical medicine
3611	Pharmacy	3.3	Health sciences
3612	Physical Therapy, Sports Therapy and Rehabilitation	3.2	Clinical medicine
3612	Physical Therapy, Sports Therapy and Rehabilitation	3.3	Health sciences
3613	Podiatry	3.2	Clinical medicine
3614	Radiological and Ultrasound Technology	2.6	Medical engineering
3615	Respiratory Care	3.2	Clinical medicine
3616	Speech and Hearing	3.2	Clinical Medicine
3616	Speech and Hearing	5.1	Psychology and cognitive sciences

Annex 4: She Figures Index – supplementary material

This Annex includes further detail on the methodology of the She Figures Index. Specifically, it expands on the measurement framework development, the results of the multi-modelling process and results of robustness checks.

7.1.1. Measurement framework development

This section provides details of the process of selecting indicators for each of the six dimensions identified in the conceptual framework ([301]). The process of short-listing indicators was conducted iteratively, involving extensive checks of missing values, correlation structures and Principal Components Analysis (PCA) ([302]). Some key terms that support interpretation of a PCA are as follows:

- **Eigenvalues** can be thought of as scores that indicate how much of the data's overall variation is explained by each principal component. Higher eigenvalues mean that a component captures more of the data's variation. A minimum threshold often used as a guide is the value of 1, the so-called Kaiser criterion, to decide how many components to use.

- **Factor loadings** show how much each original indicator contributes to each principal component. Higher loadings mean a stronger association of that indicator with the component. Typically, a minimum value of 0.5 is considered adequate.

- The **cumulative variance** shows the total amount of variation captured by the principal components together. For example, if the first two components capture 80% of the cumulative variance, it means they together explain 80% of the variation in the data. A minimum threshold of 70% is usually used.

- The **Kaiser-Meyer-Olkin** (KMO) measure is a statistical test to assess if the indicators in the dataset are sufficiently correlated to make a PCA meaningful. KMO values range from 0 to 1, with higher values indicating that the dataset is suitable.

Segregation in the pipeline

Indicators to measure horizontal segregation, i.e. the under- or over-representation of women and men in different subject areas, were selected in relation to doctoral graduates rather than researchers or grade level due to the large of number of missing values for these indicators. The She Figures report provides a breakdown in ten subject areas, and the correlations between women's representation in each of these subject areas was examined (Table 9 **Correlation matrix – share of women as doctoral graduates by subject areas, 2018**).

Indicator selection is guided by conceptual relevance and coherence in the correlation structure, marked by positive correlation. Table 9 **Correlation matrix – share of women as doctoral graduates by subject areas, 2018** shows a moderate positive correlation (r = 0.39) between the share of women in Information and Communication Technologies (ICT) and in Engineering, manufacturing and construction. These areas also represent conceptual relevant subject areas to include in the She Figures Index given that there are typically subject areas where women are under-represented, despite the relevance of these dimensions for technological and digital growth in Europe. In 2018, in the EU-27, women represented 22.37% of doctoral graduates in ICT, and 29.43% in Engineering, manufacturing and construction (She Figures, 2021). A further strong positive correlation exists between the share of women graduates in Arts and humanities and in Business and law (r = 0.61), although there is overall more of a gender balance in these dimensions where in 2018, in the EU-27, women represented 55.46% of doctoral graduate in Arts and humanities, and 44.76% in Business and law (She Figures, 2021). The share of women graduates in Arts and humanities is also moderately positively correlated with the share of women in Engineering, manufacturing and construction (r = 0.37). Accordingly, all four indicators were short-listed.

([301]) These are: segregation in the pipeline, research careers and sectors, career progression, representation in decision-making positions, research participation and GDRIC.

([302]) PCA is a statistical tool used to identify the main structure of a datasets. It aims to identify the main dimensions by analysing the correlation structure, and thus help identify what the principal components are. These principal components capture most of the important information

Table 9 Correlation matrix – share of women as doctoral graduates by subject areas, 2018

	Subject area:	1	2	3	4	5	6	7	8	9	10
1	Education	1.00									
2	Arts and humanities	-0.42	1.00								
3	Social sciences, journalism and information	0.13	0.32	1.00							
4	Business, administration and law	-0.21	0.61	0.11	1.00						
5	Natural sciences, mathematics and statistics	0.34	-0.10	0.12	-0.07	1.00					
6	Information and Communication Technologies	-0.04	0.25	-0.01	0.08	0.05	1.00				
7	Engineering, manufacturing and construction	0.17	0.37	0.01	0.25	0.10	0.39	1.00			
8	Agriculture, forestry, fisheries and veterinary	-0.02	-0.08	-0.03	0.22	0.09	-0.08	-0.36	1.00		
9	Health and welfare	0.15	0.07	0.11	-0.18	0.17	-0.04	0.26	-0.32	1.00	
10	Services	0.19	-0.19	0.05	-0.57	-0.23	0.08	0.13	0.16	0.29	1.00

The results of the PCA in the dimension ‘segregation in the pipeline’ suggest that there are two underlying constructs (Table 10). An examination of the rotated solution for the two components solution shows that a differentiation between related subject areas, with ‘Arts and humanities’ and ‘Business and law’ combined into one component, and ‘ICT’ and ‘Engineering, manufacturing, and construction’ in the other component (Table 11). This suggests that, if all four indicators were to be retained in the She Figures Index, it may be necessary to create two sub-dimensions. To ensure that it is valid to aggregate all four indicators into a single construct, rotated factor loadings are produced for the one component solution (table 12). These results suggest that a single dimension is possible, though with lower loadings for some indicators, and that it might therefore be more suitable to only retain some of the indicators.

Table 10 Principal Components – Eigenvalues and cumulative proportion of variance for segregation in the pipeline indicators

	Eigenvalue	Cumulative proportion of variance
Component 1	2.00	0.50
Component 2	1.06	0.76
Component 3	0.58	0.91
Component 4	0.36	1.00

Notes: Indicators: Share of women in Arts and humanities; Share of women in Business and law; Share of women in ICT; Share of women in Engineering, manufacturing and construction

Table 11 Principal Components – Rotated factor loadings for segregation in the pipeline indicators (two components solution)

	Component 1	Component 2	KMO
Share of women in Arts and humanities	0.65		0.58
Share of women in Business and law	0.73		0.55
Share of women in ICT		0.78	0.59
Share of women in Engineering, manufacturing and construction		0.61	0.69
Overall KMO			0.59

Notes: Values below |0.3| are blanked

Table 12 Principal Components – Rotated factor loadings for segregation in the pipeline indicators (one components solution)

	Component 1	KMO
Share of women in Arts and humanities	0.59	0.58
Share of women in Business and law	0.52	0.55
Share of women in ICT	0.37	0.59
Share of women in Engineering, manufacturing and construction	0.49	0.69
Overall KMO		0.59

Notes: Values below |0.3| are blanked

Research careers and sectors

The results of the PCA in the dimension of research careers and sectors confirm that the three indicators correspond to a single underlying construct, accounting for 73% of the variation (Table 13) and with a good fit as indicated by a KMO value of 0.69 (Table 14). The rotated factors loadings are well above the threshold of 0.5, suggesting an appropriate structure.

Table 13 Principal Components – Eigenvalues and cumulative proportion of variance for research careers and sectors

	Eigenvalue	Cumulative proportion of variance
Component 1	2.19	0.73
Component 2	0.50	0.90
Component 3	0.31	1.00

Notes: Indicators: Share of women researchers in the higher education sector; Share of women researchers in the government sector; Share of women researchers in the business enterprise sector.

Table 14 Principal Components – Rotated factor loadings for research careers and sectors

	Component 1	KMO
Share of women researchers in the higher education sector	0.60	0.64
Share of women researchers in the government sector	0.57	0.70
Share of women researchers in the business enterprise sector	0.55	0.75
Overall KMO		0.69

Notes: Values below |0.3| are blanked

Career progression

Indicators measuring gender representation in Grade A, Grade B and Grade C were considered for the dimension of career progression, that is all grades for which a doctoral qualification is expected. However, an in-depth assessment was performed to decide whether to use the set of all post-doctoral grades (A, B and C) or, alternatively, only Grades A and B. This assessment was motivated, methodologically, by the presence of missing values in the indicator measuring women's representation in Grade C positions, and importantly, conceptually, by whether the dimensions sought to measure career progression (Grades A and B represented the higher hierarchical levels in research organisations, i.e. typically representing Full or Associate Professors) or representation at different grades more generally.

The PCA analysis confirms that for the career progression dimension, the use of indicators measuring the share of women in Grade A and Grade B position represent a single underlying construct which account for 94% of the variance (Table 15). The rotated factor loadings are both over the threshold of 0.5, and there is an adequate KMO value of 0.5 (Table 16).

Table 15 Principal Components – Eigenvalues and cumulative proportion of variance for career progression (Grades A and B)

	Eigenvalue	Cumulative proportion of variance
Component 1	1.88	0.94
Component 2	0.12	1.00

Notes: Indicators: Share of women among academic staff in Grade A positions; Share of women among academic staff in Grade B positions

Table 16 Principal Components – Rotated factor loadings for career progression (Grades A and B)

	Component 1	KMO
Share of women among academic staff in Grade A positions	0.71	0.50
Share of women among academic staff in Grade B positions	0.71	0.50
Overall KMO		0.50

Notes: Values below |0.3| are blanked

A PCA analysis was also carried out with the addition of the indicator measuring the share of women academic staff in Grade C positions, although conceptually, it was acknowledged that this would shift the interpretation of the dimension from career progression to representation. The results suggested that a one-dimension solution was most appropriate as the three indicators could be seen as representing a single underlying construct, accounting for 73% of the overall variance (Table 17). However, while the KMO was acceptable at 0.59, the rotated factor loadings exceeded the thresholds of 0.5 only for the indicators measuring the share of women academic staff in Grades A and B (Table 18). This provided further evidence that retaining the indicator measuring the share of women academic staff in Grade C positions is not desirable.

Table 17 Principal Components – Eigenvalues and cumulative proportion of variance for career progression (Grades A, B and C)

	Eigenvalue	Cumulative proportion of variance
Component 1	2.18	0.73
Component 2	0.70	0.96
Component 3	0.12	1.00

Notes: Indicators: Share of women among academic staff in Grade A positions; Share of women among academic staff in Grade B positions; Share of women among academic staff in Grade C positions

Table 18 Principal Components – Rotated factor loadings for career progression (Grades A, B and C)

	Component 1	KMO
Share of women among academic staff in Grade A positions	0.62	0.56
Share of women among academic staff in Grade B positions	0.64	0.55
Share of women among academic staff in Grade C positions	0.45	0.84
Overall KMO		0.59

Notes: Values below |0.3| are blanked

Representation in decision-making positions

The PCA analysis confirmed that the indicators used to measure representation in decision-making positions represented a single underlying construct, accounting for 0.64 of variance (Table 19). The KMO suggested an adequate fit, with an overall value of 0.5 (Table 19 **Principal Components – Eigenvalues and cumulative proportion of variance for representation in decision-making positions**

	Eigenvalue	Cumulative proportion of variance
Component 1	1.91	0.64
Component 2	0.83	0.91
Component 3	0.26	1.00

Notes: Indicators: Share of women among heads of institutions in the Higher Education Sector (HES); Share of women on boards, as leaders; Share of women on boards, as members and leaders

Table 20). However, while indicators measuring the share of women as either leaders, or as members and leaders of boards demonstrated rotated factor loadings above 0.5, the indicator measuring the share of women among heads of institutions in the Higher Education Sector felt slightly short of that threshold. Nevertheless, given the importance of this indicator, it was considered further for inclusion and assessed further in the overall correlation structure (see Section 5).

Table 19 Principal Components – Eigenvalues and cumulative proportion of variance for representation in decision-making positions

	Eigenvalue	Cumulative proportion of variance
Component 1	1.91	0.64
Component 2	0.83	0.91
Component 3	0.26	1.00

Notes: Indicators: Share of women among heads of institutions in the Higher Education Sector (HES); Share of women on boards, as leaders; Share of women on boards, as members and leaders

Table 20 Principal Components – Rotated factor loadings for representation in decision-making positions

	Component 1	KMO
Share of women among heads of institutions in the Higher Education Sector (HES)	0.45	0.50
Share of women on boards, as leaders	0.67	0.50
Share of women on boards, as members and leaders	0.59	0.50
Overall KMO		0.50

Notes: Values below |0.3| are blanked

Research participation

The PCA analysis for the dimension of research participation confirmed that the three indicators examined represented a single underlying construct, with the first component explaining 71% of the (Table 21).

All rotated factor loadings are above the threshold value of 0.5, with a KMO value of 0.66 suggesting a good fit (Table 22).

Table 21 Principal Components – Eigenvalues and cumulative proportion of variance for research participation

	Eigenvalue	Cumulative proportion of variance
Component 1	2.14	0.71
Component 2	0.57	0.90
Component 3	0.29	1.00

Notes: Indicators: Share of women among authors on publications in all fields of R&D; Distribution of patent applications by sex composition of the inventors' team; Share of women beneficiaries of research funding.

Table 22 Principal Components – Rotated factor loadings for research participation

	Component 1	KMO
Share of women among authors on publications in all fields of R&D	0.61	0.61
Distribution of patent applications by sex composition of the inventors' team	0.53	0.77
Share of women beneficiaries of research funding	0.58	0.65
Overall KMO		0.66

Notes: Values below |0.3| are blanked

The gender dimension in R&I content

Three indicators for the dimension of the gender dimension in research and innovation content (GDRIC) were examined in depth. The first indicator examines the share of publications at country level with a gender dimension in their content. Two further indicators examined the extent to which Horizon projects include a gender dimension, and intersectional perspective, respectively.

The first PCA (Table 23) suggests that, as the second eigenvalue is marginally lower that the cut-off of 1, both a one- and a two-component solution should be examined. The one-component solution explains only 49% of the variance, though all three rotated factor loadings are above or near the threshold of 0.5 (Table 24). The two-component solution explains a greater proportion of the variance (82%), however the rotated factor loadings show a cross-loading, implying that this set of indicators is not able to accurately distinguish between two different sub-constructs (Table 25).

Looking at the indicators, it appears that the share of Horizon projects with a gender perspective is both related to the percentage of a country's publications with a gender dimension and with the proportion of Horizon projects with an intersectional perspective. This makes this solution unsuitable for inclusion in the She Figures Index. In addition, these solutions have a relatively low fit as indicated by a KMO value of 0.47.

Table 23 Principal Components – Eigenvalues and cumulative proportion of variance for the gender dimension in R&I content

	Eigenvalue	Cumulative proportion of variance
Component 1	1.47	0.49
Component 2	0.98	0.82
Component 3	0.55	1.00

Notes: Indicators: Percentage of a country's publications with a sex or gender dimension in their research content, by field of R&D; Percentage of country's Horizon projects integrating a sex or gender dimension; Percentage of country's Horizon projects integrating intersectionality

Table 24 Principal Components – Rotated factor loadings for the gender dimension in R&I content

	Component 1	KMO
Percentage of a country's publications with a sex or gender dimension in their research content, by field of R&D	0.48	0.46
Percentage of country's Horizon projects integrating a gender dimension	0.70	0.48
Percentage of country's Horizon projects integrating intersectionality	0.53	0.47
Overall KMO		0.47

Notes: Values below |0.3| are blanked

Table 25 Principal Components – Rotated factor loadings for the gender dimension in R&I content

	Component 1	Component 2	KMO
Percentage of a country's publications with a sex or gender dimension in their research content, by field of R&D		0.88	0.46
Percentage of country's Horizon projects integrating a sex or gender dimension	0.53	0.45	0.48
Percentage of country's Horizon projects integrating intersectionality	0.84		0.47
Overall KMO			0.47

Notes: Values below |0.3| are blanked

Alternatives are to consider either the two indicators focusing on Horizon projects together, or the indicator focusing on publications on its own. Both solutions are feasible, as demonstrated by the PCA results below. Using the indicators focusing on Horizon projects measure a single construct that accounts for 67% of the variance (Table 26). The rotated factor loadings meet the threshold of 0.5, with adequate fit as demonstrated by a KMO value of 0.5 (Table 27). Different alternatives were further considered in the assessment of the correlation structure to determine the most appropriate indicators to select (as discussed further in 5.12 and the following section).

Table 26 Principal Components – Eigenvalues and cumulative proportion of variance for the gender dimension in R&I content

	Eigenvalue	Cumulative proportion of variance
Component 1	1.34	0.67
Component 2	0.66	1.00

Notes: Indicators: Percentage of country's Horizon projects integrating a sex or gender dimension; Percentage of country's Horizon projects integrating intersectionality

Table 27 Principal Components – Rotated factor loadings for the gender dimension in R&I content

	Component 1	KMO
Percentage of country's Horizon projects integrating a sex or gender dimension	0.71	0.50
Percentage of country's Horizon projects integrating intersectionality	0.71	0.50
Overall KMO		0.50

Notes: Values below |0.3| are blanked

7.1.2. Assessment of the correlation structure and refinement of the measurement framework

To refine the measurement framework, initial tests were performed using a limited choice of methods for an initial assessment of the correlation structure. The use of different weights and aggregation method does not significantly modify the correlation structure, as the relative positions of scores at different levels are normally preserved. The same applies to metrics that rely on a linear or curvilinear approach. However, the use of an uncapped metric can produce slightly different results, as by construction, different values are obtained when women exceed the parity point. However, it was clear from the Steering Group for She Figures that there was a strong preference for an uncapped metric approach, and it was therefore used for the initial assessment (and later test in the multi-modelling assessment, see Section 5). Different options for metric to measure the level of gender dimension in R&I were tested, as one of the aims of the initial assessment was to understand which metric might be more suitable methodologically.

Three initial tests were conducted using all 19 short-listed, using the M100, M10 and MM metric in turn for indicators measuring the gender dimension in R&I content (Table 28 Test 1 – All Shares indicators (16) and gender dimension levels indicators (3) using M100 method). This decision was guided by a broader interpretation of indicators based on publications, rather than Horizon projects alone, as these projects represent only a small proportion of R&I content in Europe (Table 29 and Table 30, respectively). These revealed the five issues stated in 5.12.6 ([303]).

The next iteration tested the correlation structure after removing indicators measuring the share of women and men as doctoral students in Arts and Humanities, as well as in Business and Law (as shown in **Table 31**). This is followed by another iteration where the indicator measuring the share of women and men in Grade C positions is removed (as shown in **Table 32**). The final iteration consists of removing indicators related to the gender and intersectional dimensions in Horizon projects (as shown in **Table 33**). This decision was guided by a broader interpretation of indicators based on publications, rather than Horizon projects alone, as these projects represent only a small proportion of R&I content in Europe.

([303]) 1) moderate negative correlation between the share of women and men as doctoral graduates in Arts and Humanities and the share of women and men in Grades A and B positions; 2) low contribution of the indicator measuring the share of women and men in Grade C positions; 3) marginal contribution of the indicator measuring the shares of women and men as doctoral graduates in Business and Law and as Member and Leaders of boards; 4) no contribution from the two indicators measuring GDRIC and the indicator measuring the share of women and men in Arts and Humanities to the Index score; and 5) no contribution of the gender dimension in research and innovation content to the Index score in the case where M100 and M10 is used.

Table 28 Test 1 - All shares indicators (16) and gender dimension levels indicators (3) using M100 method

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	Doctoral Graduates - Arts and Humanities	1.00																									
2	Doctoral Graduates - Business and law	0.38	1.00																								
3	Doctoral Graduates - ICTs	0.15	0.37	1.00																							
4	Doctoral Graduates - Engineering manufacturing and construction	0.40	0.54	0.39	1.00																						
5	Researchers - Higher Education	0.10	0.25	0.54	0.54	1.00																					
6	Researchers - Government	0.44	0.50	0.36	0.63	0.52	1.00																				
7	Researchers - Business Enterprise	-0.20	-0.17	0.44	0.26	0.49	0.21	1.00																			
8	Grade A	-0.43	-0.19	0.24	0.11	0.49	0.03	0.56	1.00																		
9	Grade B	-0.48	-0.09	0.14	0.18	0.49	0.06	0.43	0.74	1.00																	
10	Grade C	0.47	0.52	0.33	0.72	0.48	0.67	-0.10	-0.12	-0.03	1.00																
11	Heads of institutions	0.05	-0.09	0.20	0.15	0.25	0.13	0.31	0.36	0.29	-0.03	1.00															
12	Leaders	-0.13	0.22	0.23	0.08	0.30	0.37	0.21	0.13	0.25	0.03	0.40	1.00														
13	Members and Leaders	-0.35	-0.02	0.18	-0.20	0.05	0.02	0.04	0.06	0.24	-0.01	0.25	0.66	1.00													
14	Authors	-0.13	0.09	0.44	0.51	0.78	0.45	0.68	0.71	0.64	0.38	0.31	0.18	0.03	1.00												
15	Patents	-0.05	0.06	0.27	0.51	0.59	0.20	0.48	0.41	0.55	0.16	0.14	0.02	-0.23	0.56	1.00											
16	Beneficiaries of funding	0.08	0.29	0.54	0.54	0.70	0.45	0.51	0.38	0.36	0.32	0.45	0.47	0.24	0.64	0.49	1.00										
17	Publications with a sex or gender dimension M100	0.00	-0.30	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	-0.25	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00									
18	Horizon projects with a sex or gender dimension M100	-0.08	-0.36	-0.23	-0.03	-0.09	-0.28	-0.22	0.08	0.21	0.05	0.07	-0.12	0.13	0.03	-0.20	-0.27	0.31	1.00								
19	Horizon projects with an intersectional dimension M100	0.17	0.29	0.46	0.25	0.22	0.19	-0.08	-0.19	-0.04	0.41	0.27	0.41	0.54	0.11	-0.17	0.23	0.02	0.34	1.00							
20	Segregation in the pipeline	0.55	0.82	0.70	0.80	0.52	0.66	0.15	-0.03	-0.01	0.69	0.10	0.19	-0.08	0.36	0.30	0.54	-0.29	-0.27	0.42	1.00						
21	Research careers and sectors	0.11	0.20	0.57	0.58	0.81	0.71	0.80	0.47	0.41	0.38	0.30	0.37	0.04	0.82	0.54	0.69	0.01	-0.27	0.11	0.53	1.00					
22	Career progression	-0.28	0.05	0.33	0.43	0.69	0.31	0.47	0.85	0.86	0.32	0.32	0.20	0.14	0.84	0.54	0.51	0.11	0.16	0.04	0.25	0.61	1.00				
23	Representation in decision-making positions	-0.18	0.09	0.26	0.03	0.26	0.26	0.23	0.21	0.31	0.01	0.64	0.92	0.80	0.21	-0.02	0.49	0.24	0.00	0.51	0.11	0.32	0.26	1.00			
24	Research participation	0.00	0.22	0.53	0.61	0.81	0.47	0.64	0.55	0.55	0.35	0.41	0.35	0.11	0.84	0.70	0.93	-0.02	-0.20	0.14	0.51	0.81	0.71	0.37	1.00		
25	Gender dimension in research and innovation content	-0.03	-0.37	-0.27	-0.07	0.03	-0.19	-0.13	0.08	0.33	-0.07	0.22	0.13	0.16	-0.01	-0.09	-0.15	0.79	0.82	0.32	-0.29	-0.15	0.16	0.20	-0.12	1.00	
26	Index	0.04	0.39	0.63	0.60	0.77	0.62	0.55	0.49	0.54	0.43	0.54	0.66	0.39	0.74	0.46	0.84	0.05	-0.14	0.42	0.62	0.82	0.70	0.68	0.86	-0.01	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Table 29 Test 2 - All shares indicators (16) and gender dimension levels indicators (3) using M10 method

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	Doctoral Graduates - Arts and Humanities	1.00																									
2	Doctoral Graduates - Business and law	0.38	1.00																								
3	Doctoral Graduates - ICTs	0.15	0.37	1.00																							
4	Doctoral Graduates - Engineering manufacturing and construction	0.40	0.54	0.39	1.00																						
5	Researchers - Higher Education	0.10	0.25	0.54	0.54	1.00																					
6	Researchers - Government	0.44	0.50	0.36	0.63	0.52	1.00																				
7	Researchers - Business Enterprise	-0.20	-0.17	0.44	0.26	0.49	0.21	1.00																			
8	Grade A	-0.43	-0.19	0.24	0.11	0.49	0.03	0.56	1.00																		
9	Grade B	-0.48	-0.09	0.14	0.18	0.49	0.06	0.43	0.74	1.00																	
10	Grade C	0.47	0.52	0.33	0.72	0.48	0.67	-0.10	-0.12	-0.03	1.00																
11	Heads of institutions	0.05	-0.09	0.20	0.15	0.25	0.13	0.31	0.36	0.29	-0.03	1.00															
12	Leaders	-0.13	0.22	0.23	0.08	0.30	0.37	0.21	0.13	0.25	0.03	0.40	1.00														
13	Members and Leaders	-0.35	-0.02	0.18	-0.20	0.05	0.02	0.04	0.06	0.24	-0.01	0.25	0.66	1.00													
14	Authors	-0.13	0.09	0.44	0.51	0.78	0.45	0.68	0.71	0.64	0.38	0.31	0.18	0.03	1.00												
15	Patents	-0.05	0.06	0.27	0.51	0.59	0.20	0.48	0.41	0.55	0.16	0.14	0.02	-0.23	0.56	1.00											
16	Beneficiaries of funding	0.08	0.29	0.54	0.54	0.70	0.45	0.51	0.38	0.36	0.32	0.45	0.47	0.24	0.64	0.49	1.00										
17	Publications with a sex or gender dimension MM	0.00	-0.30	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	-0.25	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00									
18	Horizon projects with a sex or gender dimension MM	-0.08	-0.36	-0.23	-0.03	-0.09	-0.28	-0.22	0.08	0.21	0.05	0.07	-0.12	0.13	0.03	-0.20	-0.27	0.31	1.00								
19	Horizon projects with an intersectional dimension MM	0.17	0.29	0.46	0.25	0.22	0.19	-0.08	-0.19	-0.04	0.41	0.27	0.41	0.54	0.11	-0.17	0.23	0.02	0.34	1.00							
20	Segregation in the pipeline	0.55	0.82	0.70	0.80	0.52	0.66	0.15	-0.03	-0.01	0.69	0.10	0.19	-0.08	0.36	0.30	0.54	-0.29	-0.27	0.42	1.00						
21	Research careers and sectors	0.11	0.20	0.57	0.58	0.81	0.71	0.80	0.47	0.41	0.38	0.30	0.37	0.04	0.82	0.54	0.69	0.01	-0.27	0.11	0.53	1.00					
22	Career progression	-0.28	0.05	0.33	0.43	0.69	0.31	0.47	0.85	0.86	0.32	0.32	0.20	0.14	0.84	0.54	0.51	0.11	0.16	0.04	0.25	0.61	1.00				
23	Representation in decision-making positions	-0.18	0.09	0.26	0.03	0.26	0.26	0.23	0.21	0.31	0.01	0.64	0.92	0.80	0.21	-0.02	0.49	0.24	0.00	0.51	0.11	0.32	0.26	1.00			
24	Research participation	0.00	0.22	0.53	0.61	0.81	0.47	0.64	0.55	0.55	0.35	0.41	0.35	0.11	0.84	0.70	0.93	-0.02	-0.20	0.14	0.51	0.81	0.71	0.37	1.00		
25	Gender dimension in research and innovation content	-0.03	-0.37	-0.28	-0.08	0.03	-0.19	-0.12	0.08	0.34	-0.09	0.22	0.13	0.16	-0.01	-0.08	-0.15	0.81	0.81	0.31	-0.30	-0.14	0.16	0.20	-0.11	1.00	
26	Index	0.04	0.36	0.61	0.60	0.77	0.61	0.54	0.50	0.56	0.43	0.55	0.67	0.40	0.74	0.46	0.83	0.10	-0.08	0.44	0.60	0.81	0.71	0.69	0.85	0.05	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Table 30 Test 3 - All shares indicators (16) and gender dimension levels indicators (3) using MM method

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	Doctoral Graduates - Arts and Humanities	1.00																									
2	Doctoral Graduates - Business and law	0.38	1.00																								
3	Doctoral Graduates - ICTs	0.15	0.37	1.00																							
4	Doctoral Graduates - Engineering manufacturing and construction	0.40	0.54	0.39	1.00																						
5	Researchers - Higher Education	0.10	0.25	0.54	0.54	1.00																					
6	Researchers - Government	0.44	0.50	0.36	0.63	0.52	1.00																				
7	Researchers - Business Enterprise	-0.20	-0.17	0.44	0.26	0.49	0.21	1.00																			
8	Grade A	-0.43	-0.19	0.24	0.11	0.49	0.03	0.56	1.00																		
9	Grade B	-0.48	-0.09	0.14	0.18	0.49	0.06	0.43	0.74	1.00																	
10	Grade C	0.47	0.52	0.33	0.72	0.48	0.67	-0.10	-0.12	-0.03	1.00																
11	Heads of institutions	0.05	-0.09	0.20	0.15	0.25	0.13	0.31	0.36	0.29	-0.03	1.00															
12	Leaders	-0.13	0.22	0.23	0.08	0.30	0.37	0.21	0.13	0.25	0.03	0.40	1.00														
13	Members and Leaders	-0.35	-0.02	0.18	-0.20	0.05	0.02	0.04	0.06	0.24	-0.01	0.25	0.66	1.00													
14	Authors	-0.13	0.09	0.44	0.51	0.78	0.45	0.68	0.71	0.64	0.38	0.31	0.18	0.03	1.00												
15	Patents	-0.05	0.06	0.27	0.51	0.59	0.20	0.48	0.41	0.55	0.16	0.14	0.02	-0.23	0.56	1.00											
16	Beneficiaries of funding	0.08	0.29	0.54	0.54	0.70	0.45	0.51	0.38	0.36	0.32	0.45	0.47	0.24	0.64	0.49	1.00										
17	Publications with a sex or gender dimension MM	0.00	-0.30	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	-0.25	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00									
18	Horizon projects with a sex or gender dimension MM	-0.08	-0.36	-0.23	-0.03	-0.09	-0.28	-0.22	0.08	0.21	0.05	0.07	-0.12	0.13	0.03	-0.20	-0.27	0.31	1.00								
19	Horizon projects with an intersectional dimension MM	0.17	0.29	0.46	0.25	0.22	0.19	-0.08	-0.19	-0.04	0.41	0.27	0.41	0.54	0.11	-0.17	0.23	0.02	0.34	1.00							
20	Segregation in the pipeline	0.55	0.82	0.70	0.80	0.52	0.66	0.15	-0.03	-0.01	0.69	0.10	0.19	-0.08	0.36	0.30	0.54	-0.29	-0.27	0.42	1.00						
21	Research careers and sectors	0.11	0.20	0.57	0.58	0.81	0.71	0.80	0.47	0.41	0.38	0.30	0.37	0.04	0.82	0.54	0.69	0.01	-0.27	0.11	0.53	1.00					
22	Career progression	-0.28	0.05	0.33	0.43	0.69	0.31	0.47	0.85	0.86	0.32	0.32	0.20	0.14	0.84	0.54	0.51	0.11	0.16	0.04	0.25	0.61	1.00				
23	Representation in decision-making positions	-0.18	0.09	0.26	0.03	0.26	0.26	0.23	0.21	0.31	0.01	0.64	0.92	0.80	0.21	-0.02	0.49	0.24	0.00	0.51	0.11	0.32	0.26	1.00			
24	Research participation	0.00	0.22	0.53	0.61	0.81	0.47	0.64	0.55	0.55	0.35	0.41	0.35	0.11	0.84	0.70	0.93	-0.02	-0.20	0.14	0.51	0.81	0.71	0.37	1.00		
25	Gender dimension in research and innovation content	0.07	-0.11	0.04	0.07	0.14	-0.03	-0.12	-0.04	0.23	0.14	0.30	0.32	0.39	0.04	-0.14	0.02	0.61	0.73	0.73	0.01	-0.03	0.14	0.42	0.00	1.00	
26	Index	0.06	0.30	0.56	0.55	0.72	0.54	0.44	0.42	0.54	0.43	0.57	0.69	0.48	0.66	0.36	0.74	0.24	0.12	0.61	0.55	0.70	0.65	0.74	0.75	0.52	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Table 31 Test 4 – After removing the share of women and men as graduates in Arts & Humanities and Business & Law

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Doctoral Graduates - ICTs	1.00																							
2	Doctoral Graduates - Engineering manufacturing and construction	0.39	1.00																						
3	Researchers - Higher Education	0.54	0.54	1.00																					
4	Researchers - Government	0.36	0.63	0.52	1.00																				
5	Researchers - Business Enterprise	0.44	0.26	0.49	0.21	1.00																			
6	Grade A	0.24	0.11	0.49	0.03	0.56	1.00																		
7	Grade B	0.14	0.18	0.49	0.06	0.43	0.74	1.00																	
8	Grade C	0.33	0.72	0.48	0.67	-0.10	-0.12	-0.03	1.00																
9	Heads of institutions	0.20	0.15	0.25	0.13	0.31	0.36	0.29	-0.03	1.00															
10	Leaders	0.23	0.08	0.30	0.37	0.21	0.13	0.25	0.03	0.40	1.00														
11	Members and Leaders	0.18	-0.20	0.05	0.02	0.04	0.06	0.24	-0.01	0.25	0.66	1.00													
12	Authors	0.44	0.51	0.78	0.45	0.68	0.71	0.64	0.38	0.31	0.18	0.03	1.00												
13	Patents	0.27	0.51	0.59	0.20	0.48	0.41	0.55	0.16	0.14	0.02	-0.23	0.56	1.00											
14	Beneficiaries of funding	0.54	0.54	0.70	0.45	0.51	0.38	0.36	0.32	0.45	0.47	0.24	0.64	0.49	1.00										
15	Publications with a sex or gender dimension MM	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	-0.25	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00									
16	Horizon projects with a sex or gender dimension MM	-0.23	-0.03	-0.09	-0.28	-0.22	0.08	0.21	0.05	0.07	-0.12	0.13	0.03	-0.20	-0.27	0.31	1.00								
17	Horizon projects with an intersectional dimension MM	0.46	0.25	0.22	0.19	-0.08	-0.19	-0.04	0.41	0.27	0.41	0.54	0.11	-0.17	0.23	0.02	0.34	1.00							
18	Segregation in the pipeline	0.85	0.82	0.65	0.59	0.42	0.21	0.19	0.62	0.21	0.19	0.00	0.56	0.46	0.65	-0.27	-0.16	0.43	1.00						
19	Research careers and sectors	0.57	0.58	0.81	0.71	0.80	0.47	0.41	0.38	0.30	0.37	0.04	0.82	0.54	0.69	0.01	-0.27	0.11	0.69	1.00					
20	Career progression	0.33	0.43	0.69	0.31	0.47	0.85	0.86	0.32	0.32	0.20	0.14	0.84	0.54	0.51	0.11	0.16	0.04	0.46	0.61	1.00				
21	Representation in decision-making positions	0.26	0.03	0.26	0.26	0.23	0.21	0.31	0.01	0.64	0.92	0.80	0.21	-0.02	0.49	0.24	0.00	0.51	0.18	0.32	0.26	1.00			
22	Research participation	0.53	0.61	0.81	0.47	0.64	0.55	0.55	0.35	0.41	0.35	0.11	0.84	0.70	0.93	-0.02	-0.20	0.14	0.68	0.81	0.71	0.37	1.00		
23	Gender dimension in research and innovation content	0.04	0.07	0.14	-0.03	-0.12	-0.04	0.23	0.14	0.30	0.32	0.39	0.04	-0.14	0.02	0.61	0.73	0.73	0.06	-0.03	0.14	0.42	0.00	1.00	
24	Index	0.60	0.56	0.73	0.51	0.50	0.46	0.57	0.41	0.57	0.65	0.46	0.69	0.40	0.75	0.22	0.13	0.59	0.70	0.73	0.68	0.71	0.77	0.50	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Table 32 Test 5 – After removing the share of women and men as in Grade C positions

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	Doctoral Graduates - ICTs	1.00																						
2	Doctoral Graduates - Engineering manufacturing and construction	0.39	1.00																					
3	Researchers - Higher Education	0.54	0.54	1.00																				
4	Researchers - Government	0.36	0.63	0.52	1.00																			
5	Researchers - Business Enterprise	0.44	0.26	0.49	0.21	1.00																		
6	Grade A	0.24	0.11	0.49	0.03	0.56	1.00																	
7	Grade B	0.14	0.18	0.49	0.06	0.43	0.74	1.00																
8	Heads of institutions	0.20	0.15	0.25	0.13	0.31	0.36	0.29	1.00															
9	Leaders	0.23	0.08	0.30	0.37	0.21	0.13	0.25	0.40	1.00														
10	Members and Leaders	0.18	-0.20	0.05	0.02	0.04	0.06	0.24	0.25	0.66	1.00													
11	Authors	0.44	0.51	0.78	0.45	0.68	0.71	0.64	0.31	0.18	0.03	1.00												
12	Patents	0.27	0.51	0.59	0.20	0.48	0.41	0.55	0.14	0.02	-0.23	0.56	1.00											
13	Beneficiaries of funding	0.54	0.54	0.70	0.45	0.51	0.38	0.36	0.45	0.47	0.24	0.64	0.49	1.00										
14	Publications with a sex or gender dimension MM	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00									
15	Horizon projects with a sex or gender dimension MM	-0.23	-0.03	-0.09	-0.28	-0.22	0.08	0.21	0.07	-0.12	0.13	0.03	-0.20	-0.27	0.31	1.00								
16	Horizon projects with an intersectional dimension MM	0.46	0.25	0.22	0.19	-0.08	-0.19	-0.04	0.27	0.41	0.54	0.11	-0.17	0.23	0.02	0.34	1.00							
17	Segregation in the pipeline	0.85	0.82	0.65	0.59	0.42	0.21	0.19	0.21	0.19	0.00	0.56	0.46	0.65	-0.27	-0.16	0.43	1.00						
18	Research careers and sectors	0.57	0.58	0.81	0.71	0.80	0.47	0.41	0.30	0.37	0.04	0.82	0.54	0.69	0.01	-0.27	0.11	0.69	1.00					
19	Career progression	0.21	0.15	0.52	0.04	0.53	0.95	0.91	0.35	0.19	0.15	0.72	0.51	0.40	0.21	0.15	-0.13	0.22	0.48	1.00				
20	Representation in decision-making positions	0.26	0.03	0.26	0.26	0.23	0.21	0.31	0.64	0.92	0.80	0.21	-0.02	0.49	0.24	0.00	0.51	0.18	0.32	0.27	1.00			
21	Research participation	0.53	0.61	0.81	0.47	0.64	0.55	0.55	0.41	0.35	0.11	0.84	0.70	0.93	-0.02	-0.20	0.14	0.68	0.81	0.59	0.37	1.00		
22	Gender dimension in research and innovation content	0.04	0.07	0.14	-0.03	-0.12	-0.04	0.23	0.30	0.32	0.39	0.04	-0.14	0.02	0.61	0.73	0.73	0.06	-0.03	0.08	0.42	0.00	1.00	
23	Index	0.58	0.51	0.73	0.46	0.54	0.54	0.63	0.59	0.65	0.46	0.71	0.42	0.74	0.25	0.14	0.54	0.66	0.72	0.62	0.71	0.78	0.49	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Table 33 Test 6 – After removing the indicators on the gender and intersectional dimensions in Horizon projects

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	Doctoral Graduates - ICTs	1.00																				
2	Doctoral Graduates - Engineering manufacturing and construction	0.39	1.00																			
3	Researchers - Higher Education	0.54	0.54	1.00																		
4	Researchers - Government	0.36	0.63	0.52	1.00																	
5	Researchers - Business Enterprise	0.44	0.26	0.49	0.21	1.00																
6	Grade A	0.24	0.11	0.49	0.03	0.56	1.00															
7	Grade B	0.14	0.18	0.49	0.06	0.43	0.74	1.00														
8	Heads of institutions	0.20	0.15	0.25	0.13	0.31	0.36	0.29	1.00													
9	Leaders	0.23	0.08	0.30	0.37	0.21	0.13	0.25	0.40	1.00												
10	Members and Leaders	0.18	-0.20	0.05	0.02	0.04	0.06	0.24	0.25	0.66	1.00											
11	Authors	0.44	0.51	0.78	0.45	0.68	0.71	0.64	0.31	0.18	0.03	1.00										
12	Patents	0.27	0.51	0.59	0.20	0.48	0.41	0.55	0.14	0.02	-0.23	0.56	1.00									
13	Beneficiaries of funding	0.54	0.54	0.70	0.45	0.51	0.38	0.36	0.45	0.47	0.24	0.64	0.49	1.00								
14	Publications with a sex or gender dimension MM	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00							
15	Segregation in the pipeline	0.85	0.82	0.65	0.59	0.42	0.21	0.19	0.21	0.19	0.00	0.56	0.46	0.65	-0.27	1.00						
16	Research careers and sectors	0.57	0.58	0.81	0.71	0.80	0.47	0.41	0.30	0.37	0.04	0.82	0.54	0.69	0.01	0.69	1.00					
17	Career progression	0.21	0.15	0.52	0.04	0.53	0.95	0.91	0.35	0.19	0.15	0.72	0.51	0.40	0.21	0.22	0.48	1.00				
18	Representation in decision-making positions	0.26	0.03	0.26	0.26	0.23	0.21	0.31	0.64	0.92	0.80	0.21	-0.02	0.49	0.24	0.18	0.32	0.27	1.00			
19	Research participation	0.53	0.61	0.81	0.47	0.64	0.55	0.55	0.41	0.35	0.11	0.84	0.70	0.93	-0.02	0.68	0.81	0.59	0.37	1.00		
20	Gender dimension in research and innovation content	-0.31	-0.14	0.09	-0.08	0.02	0.08	0.35	0.25	0.26	0.05	-0.07	0.08	-0.03	1.00	-0.27	0.01	0.21	0.24	-0.02	1.00	
21	Index	0.43	0.41	0.71	0.42	0.58	0.57	0.70	0.59	0.65	0.36	0.65	0.49	0.71	0.49	0.51	0.72	0.68	0.68	0.76	0.49	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

7.1.3. Multi-modelling results

The different alternatives considered explored the effects of weighting and aggregation methods, as well as the type of metric (un/capped and curvi/linear). Alternatives are only computed for the case where the MM method is used for the gender dimension, as the assessment of the correlation structure suggested it was the most appropriate indicator (see above). The principle of increased penalisation is applied, which means that only aggregation methods at the level of indicators that provide as much or more penalisation can be applied (i.e. if the geometric mean is applied at the level of indicators, it is only possible to use the geometric or harmonic mean at the level of dimensions).

The weighting methods considered are equal weights (W1), and two suggested weightings that were derived from a budget allocation exercise with members of the She Figures Steering Group, as shown in Table 34. Five submissions were returned, three of which expressed a preference for equal weights (W1). A second suggestion (W2) allocated more weight to the dimensions of progression and decision-making, on the basis that career advancement and participation play a key role in gender equality and are more in the limelight of politicians and the general public. The third suggestion (W3) allocated more weight to the gender dimension in R&I content, arguing that it is an emerging issue in research and innovation that should be pushed more than the other dimensions.

Table 34 Summary of weight budget allocation

Dimension	W1 (Equal weights)	W2	W3
Segregation in the pipeline	1/6	1/8	1/8
Research careers and sectors	1/6	1/8	1/8
Career progression	1/6	1/4	1/8
Representation in decision-making positions	1/6	1/4	1/8
Research participation	1/6	1/8	1/8
The gender dimension in research and innovation content	1/6	1/8	3/8

The list of alternatives considered is provided in **Table** 35.

Table 35 Different alternatives considered in the multi-modelling approach

Alternative	Metric	Aggregation method at the level of indicators	Aggregation method at the level of dimensions	Weighting method
1	Linear and capped (LC)	Harmonic mean (HM)	Harmonic mean (HM)	W3
2				W2
3				W1 (Equal weights)
4		Geometric mean (GM)	Harmonic mean (HM)	W3
5				W2
6				W1 (Equal weights)
7			Geometric mean (GM)	W3
8				W2
9				W1 (Equal weights)
10		Arithmetic mean (AM)	Harmonic mean (HM)	W3
11				W2
12				W1 (Equal weights)
13			Geometric mean (GM)	W3
14				W2
15				W1 (Equal weights)
16			Arithmetic mean (AM)	W3
17				W2
18				W1 (Equal weights)
19		Harmonic mean (HM)	Harmonic mean (HM)	W3
20				W2

21	Linear and uncapped (LU)			W1 (Equal weights)
22		Geometric mean (GM)	Harmonic mean (HM)	W3
23				W2
24				W1 (Equal weights)
25			Geometric mean (GM)	W3
26				W2
27				W1 (Equal weights)
28		Arithmetic mean (AM)	Harmonic mean (HM)	W3
29				W2
30				W1 (Equal weights)
31			Geometric mean (GM)	W3
32				W2
33				W1 (Equal weights)
34			Arithmetic mean (AM)	W3
35				W2
36				W1 (Equal weights)
37	Curvilinear and capped (CC)	Harmonic mean (HM)	Harmonic mean (HM)	W3
38				W2
39				W1 (Equal weights)
40		Geometric mean (GM)	Harmonic mean (HM)	W3
41				W2
42				W1 (Equal weights)

43				W3
44			Geometric mean (GM)	W2
45				W1 (Equal weights)
46				W3
47			Harmonic mean (HM)	W2
48				W1 (Equal weights)
49				W3
50		Arithmetic mean (AM)	Geometric mean (GM)	W2
51				W1 (Equal weights)
52				W3
53			Arithmetic mean (AM)	W2
54				W1 (Equal weights)
55				W3
56		Harmonic mean (HM)	Harmonic mean (HM)	W2
57				W1 (Equal weights)
58				W3
59	Curvilinear and uncapped (CU)		Harmonic mean (HM)	W2
60		Geometric mean (GM)		W1 (Equal weights)
61				W3
62			Geometric mean (GM)	W2
63				W1 (Equal weights)
64			Harmonic mean (HM)	W3

65				W2
66				W1 (Equal weights)
67			Geometric mean (GM)	W3
68		Arithmetic mean (AM)		W2
69				W1 (Equal weights)
70			Arithmetic mean (AM)	W3
71				W2
72				W1 (Equal weights)

Distribution of scores – 72 alternatives

The distribution of scores across the 27 Member States from the multi-modelling process are provided in:

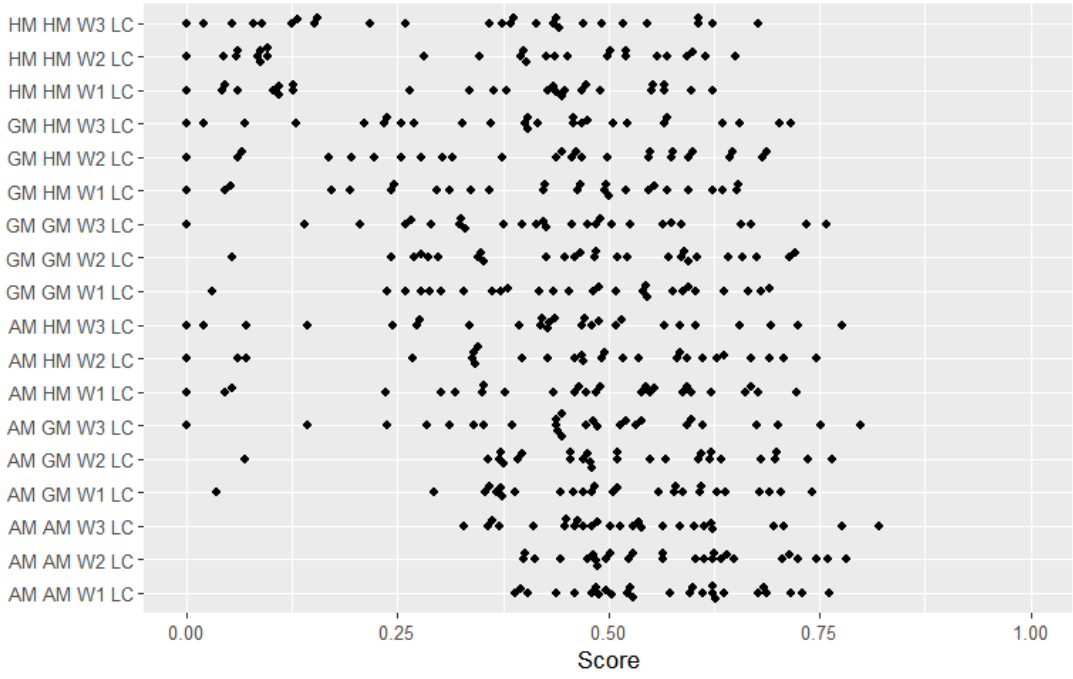
- Figure 4 for the linear capped metric
- Figure 5 for the linear uncapped metric
- Figure 6 for the curvilinear capped metric
- Figure 7 for the curvilinear uncapped metric.

As expected, by construction, the scores are lower overall for the linear metrics than the curvilinear metrics. However, capping or uncapping provides broadly similar distribution of scores. This reflects that fact that, across indicators, it is usually women that are at a disadvantage.

What is apparent is that the use of the geometric and harmonic means pulls down some of the scores quite significantly, as a result of the presence of low values close to 0. The results show that using aggregation methods that do allow for less compensation (geometric and harmonic means) generates 'penalties' for some Member States that are large. This is not seen as desirable, as it is cause by the presence of '0' (recoded as 0.01), with a disproportionate penalty for data points that tend to this value.

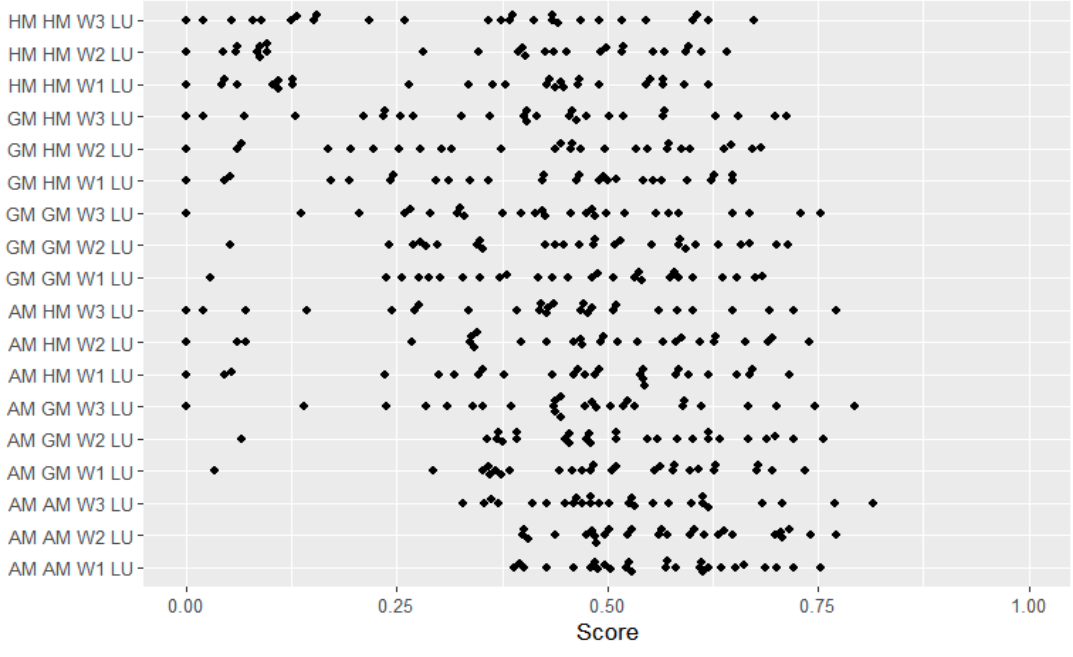
The distributions of scores and ranks by Member State are provided in Figure 8 and Figure 9 respectively. This confirms that the use of the geometric and harmonic means introduces a lot of variation in both scores and ranks, particularly in countries that perform less well in relation to gender equality in R&I. On the basis of the analysis, only the arithmetic mean as a model of aggregation was retained.

Figure 4 Distribution of scores by method – Linear capped



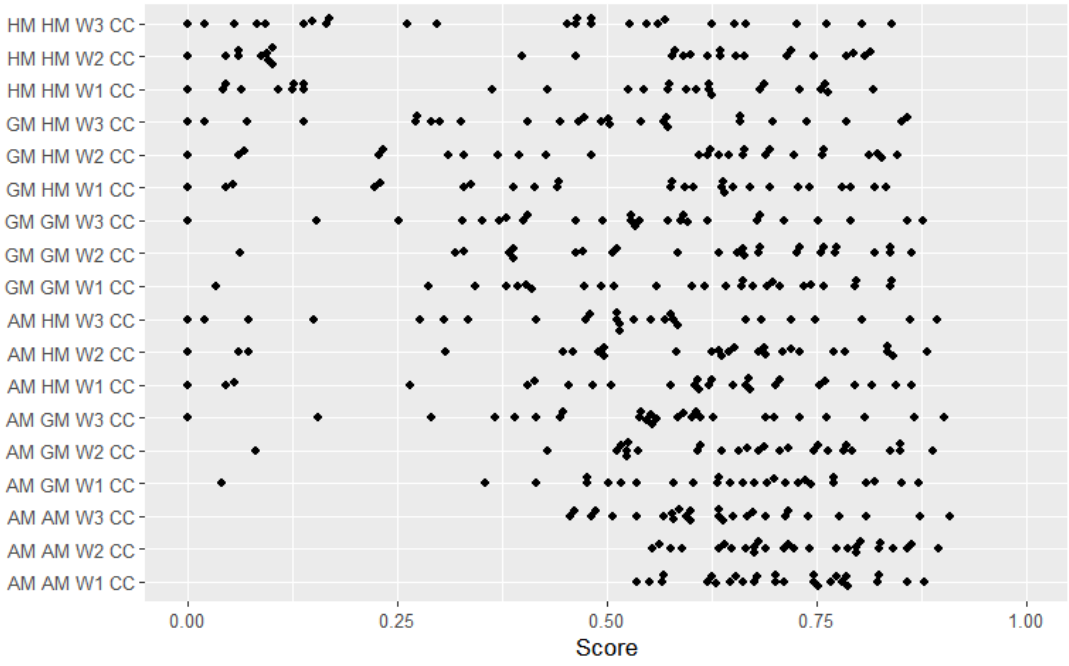
Notes: AM – Arithmetic mean; GM – Geometric mean; HM – Harmonic mean; W1-3 – Weight options 1 to 3; LC – Linear capped.

Figure **5** Distribution of scores by method – Linear uncapped



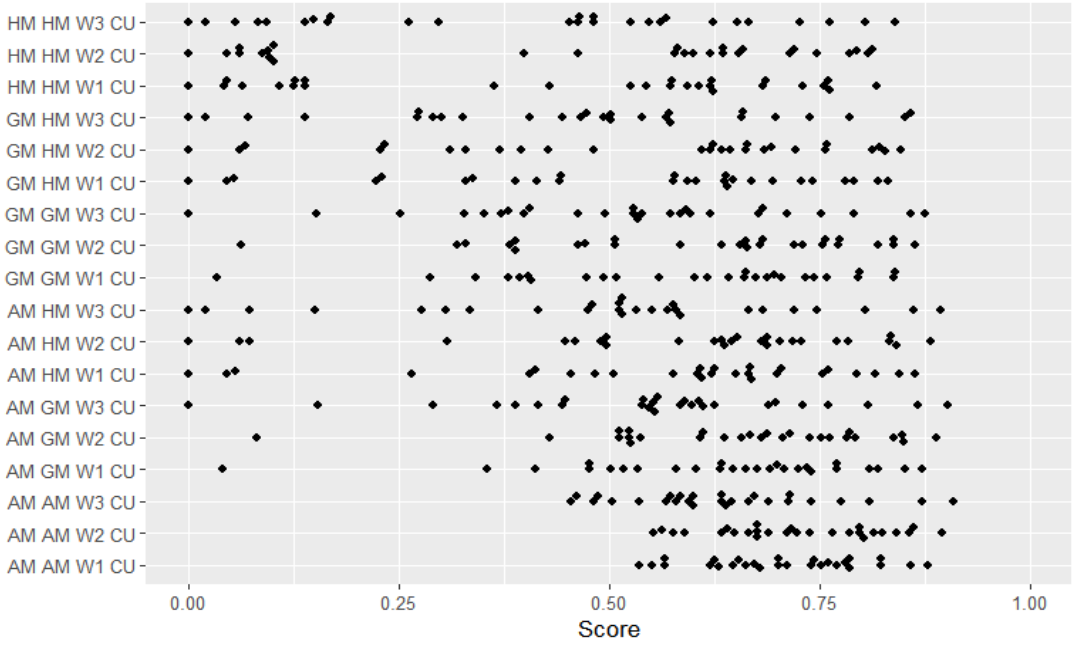
Note: AM – Arithmetic mean; GM – Geometric mean; HM – Harmonic mean; W1-3 – Weight options 1 to 3; LU – Linear uncapped.

Figure 6 Distribution of scores by method – Curvilinear capped



Note: AM – Arithmetic mean; GM – Geometric mean; HM – Harmonic mean; W1-3 – Weight options 1 to 3; CC – Curvilinear capped.

Figure 7 Distribution of scores by method – Curvilinear uncapped



Note: AM – Arithmetic mean; GM – Geometric mean; HM – Harmonic mean; W1-3 – Weight options 1 to 3; CU – Curvilinear uncapped.

Figure 8 Distribution of scores by Member States – 72 alternatives

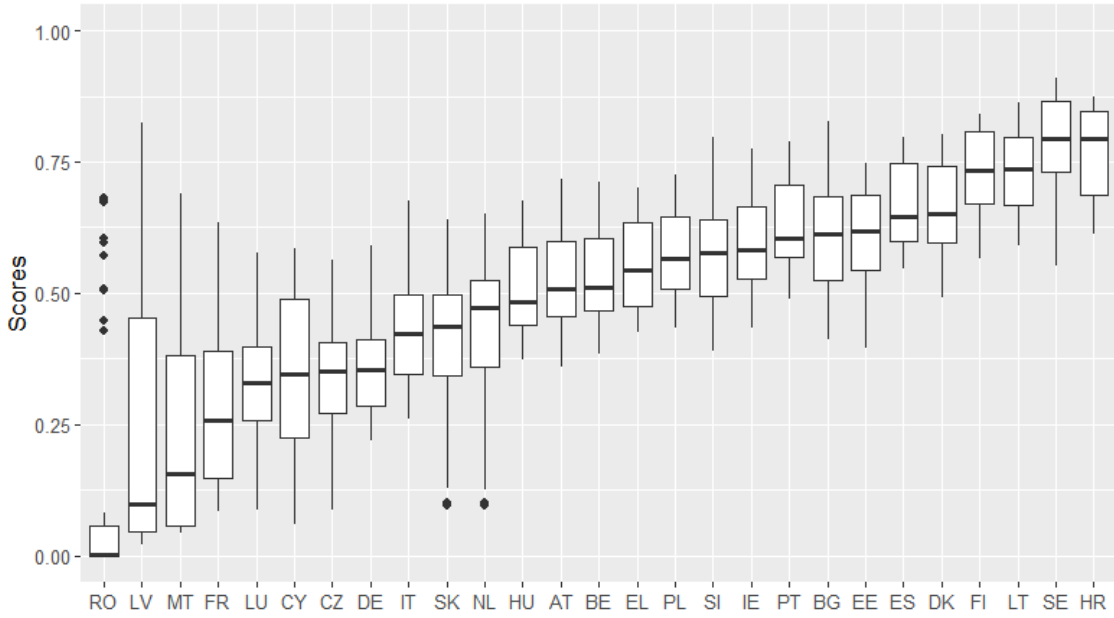
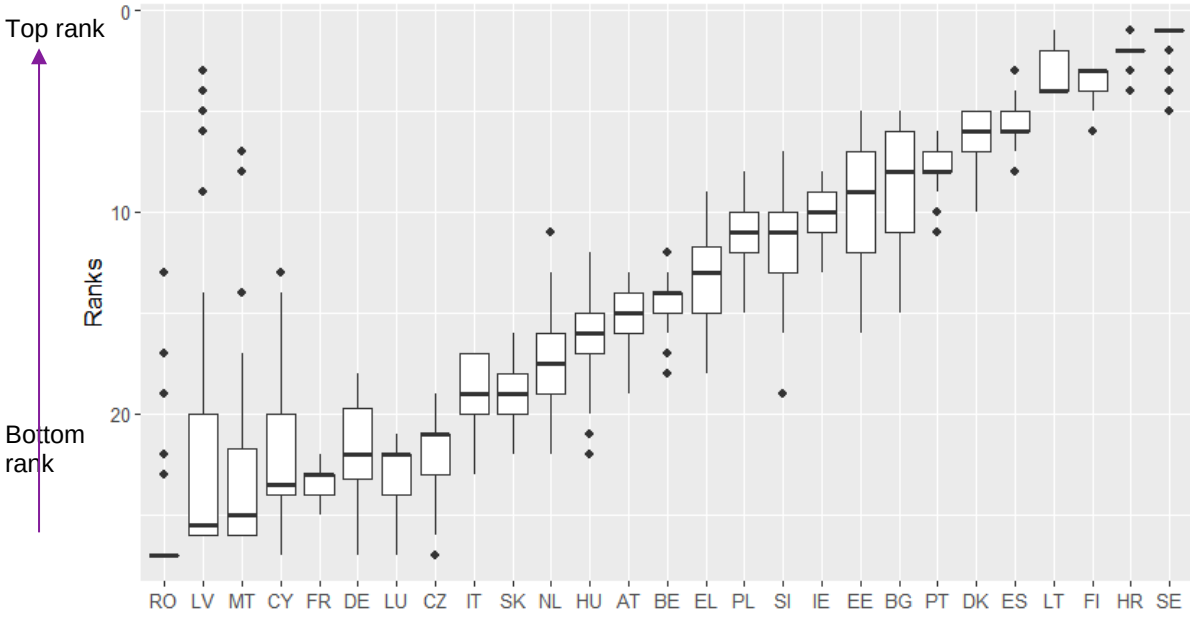


Figure 9 Distribution of ranks by Member States – 72 alternatives



Distribution of scores – 12 alternatives

The multi-model analysis was replicated with only the use of the arithmetic mean as a method of aggregation, yielding a narrower set of 12 alternatives. The associated distributions of scores and ranks by Member State are provided in Figure 9 and Figure 10, respectively. Using only the arithmetic mean as a method of aggregation decreases the variation in scores significantly, however, there is still substantial variation in ranks for some Member States.

Figure 10 Distribution of scores by Member States – 12 alternatives

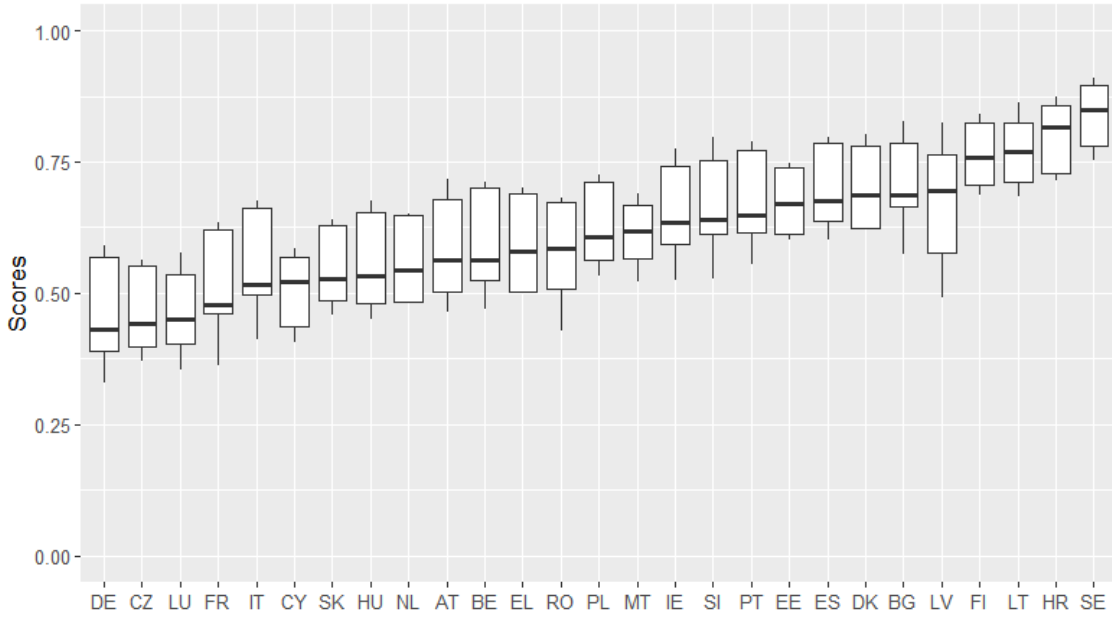
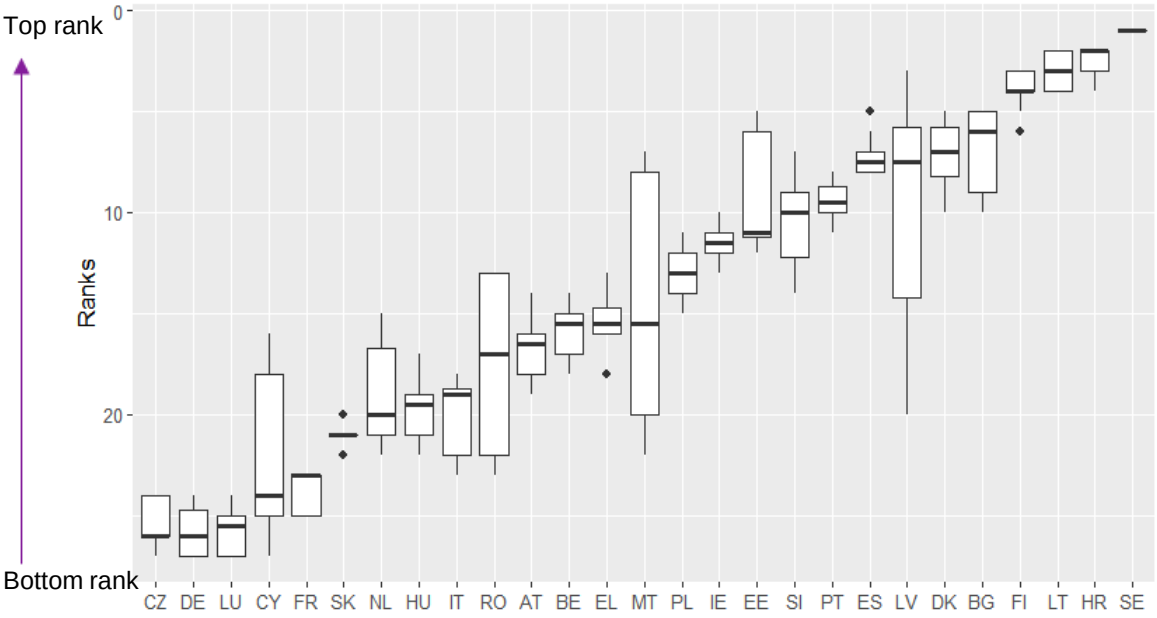


Figure 1 Distribution of ranks by Member States – 12 alternatives



A closer examination of the Member States for which these large variations in ranks are occurring show the origin of the issue. In the case of Latvia, the scores are lowest when the set of weights W3 is used (Figure 12), corresponding to ranks ranging from 14 to 20, well below the median rank of 7.5 (Figure 13). A similar pattern emerges for Romania, as illustrated in Figure 14 for scores and Figure 15 for ranks. In the case of Malta, however, the effects of using W3 weights is reversed, providing much higher scores and ranks (Figure 16 and Figure 17).

What these results demonstrate is that the calculation of the She Figures Index is very sensitive to the indicator measuring the gender dimension in R&I content. The use of the expert weight W3 places more emphasis on this aspect, amplifying the variation in some Member States (see also section 5 for further results on this aspect). Noting that this indicator relies on a different metric, and the variation observed in the multi-modelling analysis leads to the conclusion that it is more sensible to use a methodology for weighting that does not provide disproportionate weighting to this dimension, and hence instead using the option of equal weights (W1).

Figure 2 Example: Scores for Latvia across the 12 alternatives

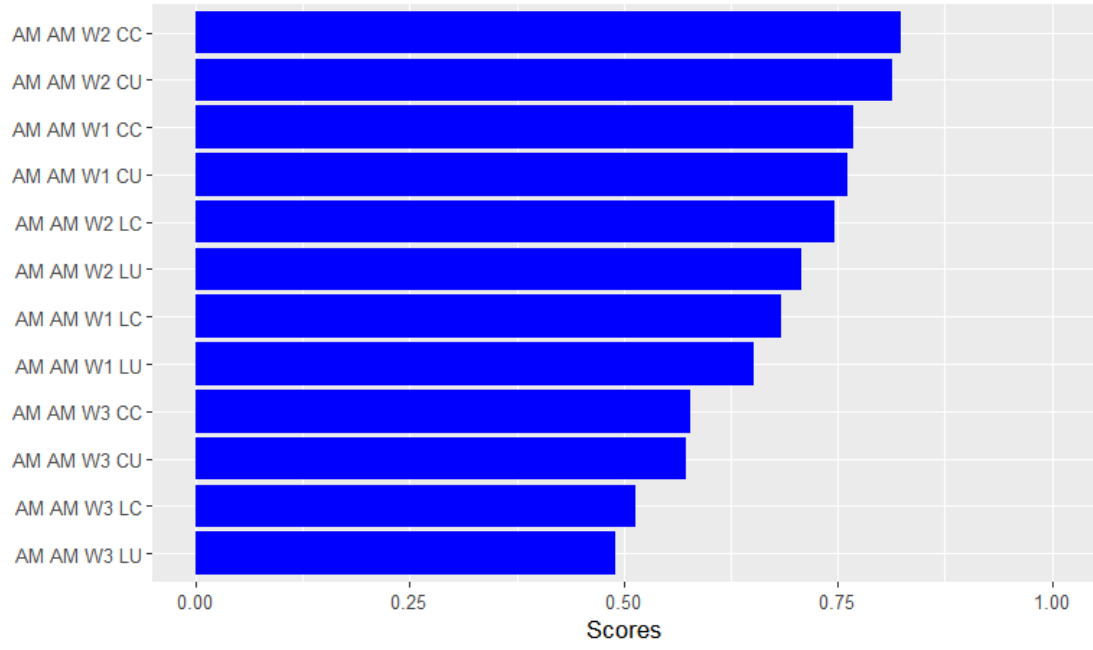


Figure 3 Example: Ranks for Latvia across the 12 alternatives

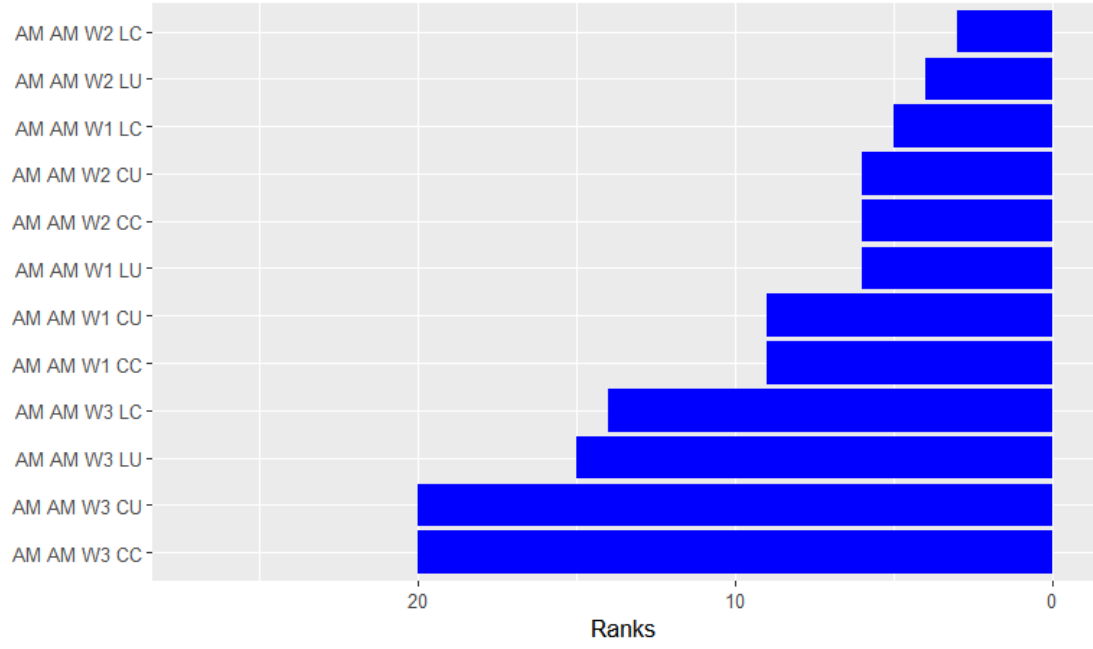


Figure 4 Example: Scores for Romania across the 12 alternatives

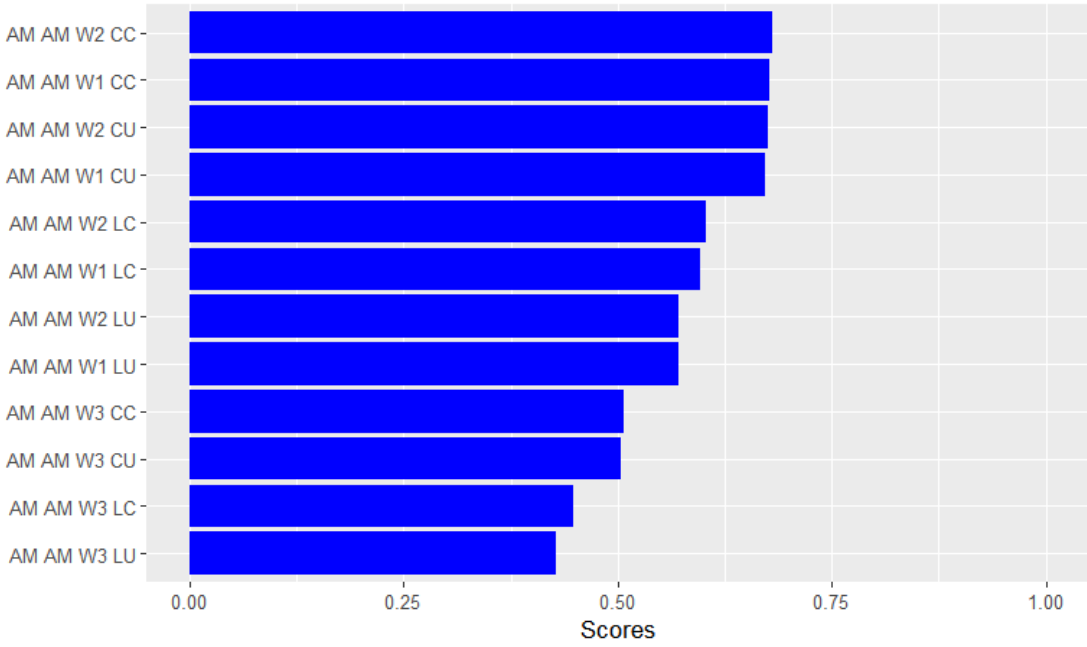


Figure 5 Example: Ranks for Romania across the 12 alternatives

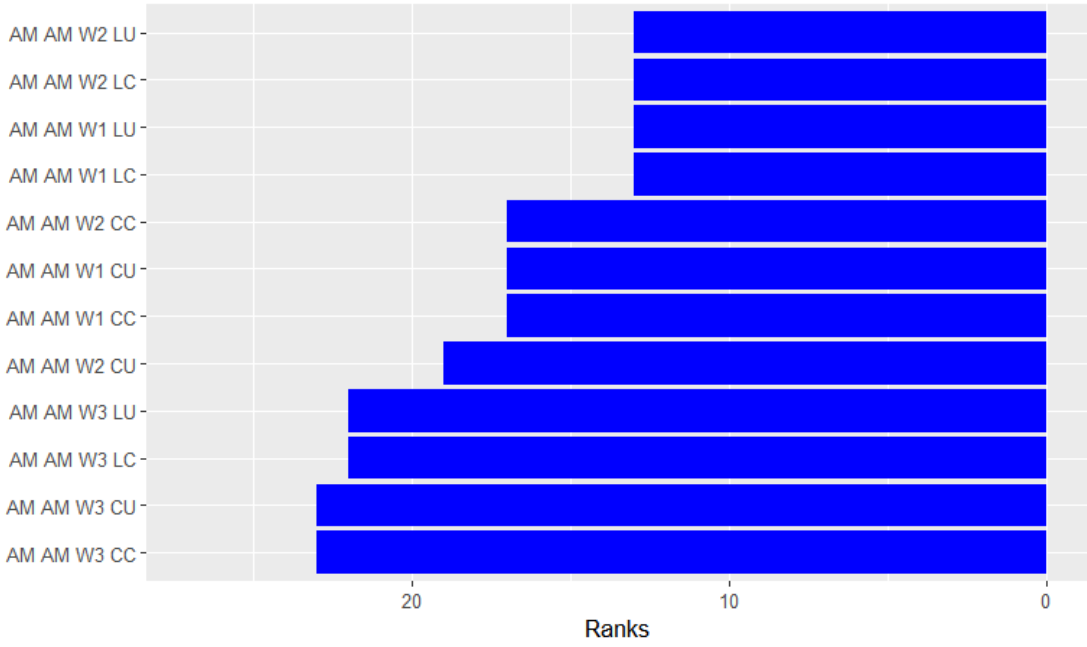


Figure 6 Example: Scores for Malta across the 12 alternatives

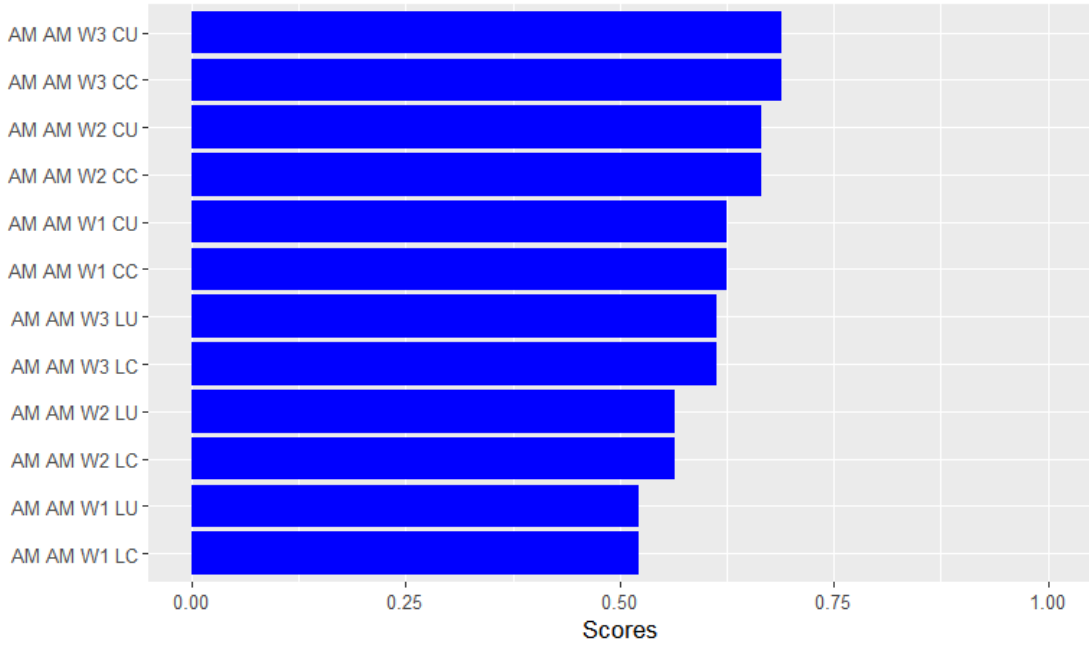
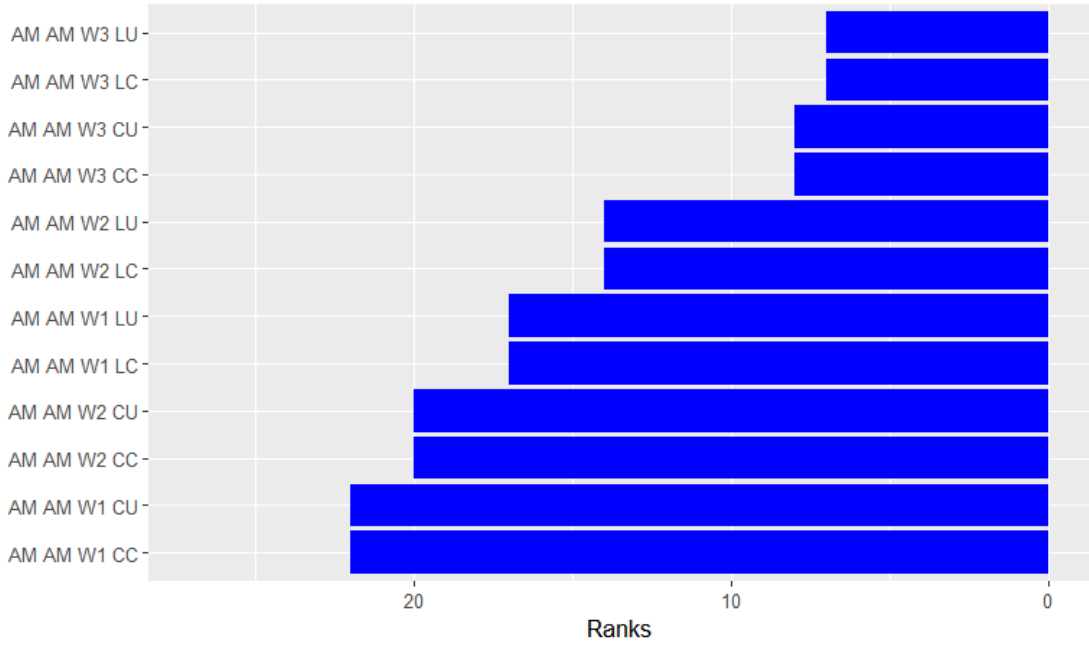


Figure 7 Example: Ranks for Malta across the 12 alternatives



7.1.4. Robustness checks

Assessments of the correlation structure for the 2021 and 2024 editions of the She Figures Index are shown in Figure 18 and Figure 19.

The results of a robustness check that show the contribution of each dimension can be found Figure 20.

Figure 18 Assessment of correlation structure – 2021 edition

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	Doctoral Graduates - ICT	1.00																				
2	Doctoral Graduates - Engineering manufacturing and construction	0.57	1.00																			
3	Researchers - Higher Education	0.69	0.65	1.00																		
4	Researchers - Government	0.62	0.64	0.60	1.00																	
5	Researchers - Business Enterprise	0.54	0.52	0.48	0.10	1.00																
6	Grade A	0.28	0.23	0.41	-0.11	0.42	1.00															
7	Grade B	0.28	0.33	0.40	0.00	0.44	0.73	1.00														
8	Heads of institutions	0.05	0.27	0.35	0.00	0.26	0.44	0.20	1.00													
9	Leaders	0.38	0.11	0.38	0.37	0.23	0.15	0.21	0.30	1.00												
10	Members and Leaders	0.24	-0.16	0.09	-0.08	0.04	0.04	0.10	0.14	0.50	1.00											
11	Authors	0.50	0.64	0.71	0.31	0.63	0.70	0.73	0.39	0.20	-0.01	1.00										
12	Patents	0.36	0.54	0.53	0.19	0.49	0.45	0.58	0.16	0.09	-0.27	0.59	1.00									
13	Beneficiaries of funding	0.60	0.52	0.72	0.56	0.37	0.37	0.33	0.32	0.50	0.14	0.54	0.48	1.00								
14	Publications with a sex or gender dimension MM	-0.24	-0.10	0.05	-0.20	0.09	0.12	0.26	0.28	0.21	-0.04	0.01	0.07	-0.08	1.00							
15	Segregation in the pipeline	0.94	0.81	0.76	0.70	0.59	0.29	0.34	0.15	0.31	0.10	0.61	0.48	0.64	-0.21	1.00						
16	Research sectors	0.74	0.72	0.73	0.50	0.91	0.36	0.41	0.26	0.37	0.02	0.71	0.53	0.59	0.01	0.82	1.00					
17	Career progression	0.30	0.28	0.43	-0.08	0.45	0.97	0.87	0.39	0.18	0.06	0.76	0.52	0.38	0.18	0.33	0.40	1.00				
18	Decision-making	0.34	0.13	0.41	0.23	0.26	0.27	0.24	0.60	0.91	0.64	0.27	0.05	0.49	0.23	0.29	0.34	0.28	1.00			
19	Research participation	0.60	0.67	0.78	0.46	0.56	0.56	0.60	0.34	0.35	-0.04	0.79	0.81	0.87	-0.01	0.69	0.72	0.61	0.35	1.00		
20	Gender dimension in research and innovation content	-0.24	-0.10	0.05	-0.20	0.09	0.12	0.26	0.28	0.21	-0.04	0.01	0.07	-0.08	1.00	-0.21	0.01	0.18	0.23	-0.01	1.00	
21	She Figures Index	0.51	0.47	0.69	0.30	0.59	0.55	0.62	0.56	0.67	0.26	0.63	0.48	0.61	0.55	0.55	0.67	0.61	0.72	0.68	0.55	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Figure 8 Assessment of correlation structure – 2024 edition

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	Doctoral Graduates - ICT	1.00																				
2	Doctoral Graduates - Engineering manufacturing and construction	0.05	1.00																			
3	Researchers - Higher Education	0.35	0.55	1.00																		
4	Researchers - Government	0.02	0.28	0.19	1.00																	
5	Researchers - Business Enterprise	0.27	0.64	0.50	0.15	1.00																
6	Grade A	0.12	0.38	0.37	0.23	0.16	1.00															
7	Grade B	0.07	0.55	0.40	0.20	0.26	0.70	1.00														
8	Heads of institutions	-0.14	0.28	0.54	0.28	0.23	0.38	0.28	1.00													
9	Leaders	0.32	0.07	0.62	0.08	0.09	0.12	0.06	0.32	1.00												
10	Members and Leaders	-0.08	-0.10	0.25	0.22	-0.04	-0.11	-0.02	0.28	0.42	1.00											
11	Authors	0.21	0.72	0.71	0.39	0.65	0.64	0.74	0.41	0.34	0.02	1.00										
12	Patents	0.32	0.44	0.46	0.01	0.58	0.30	0.45	0.14	0.17	-0.34	0.61	1.00									
13	Beneficiaries of funding	0.27	0.50	0.66	0.09	0.40	0.47	0.51	0.37	0.52	0.10	0.62	0.60	1.00								
14	Publications with a sex or gender dimension MM	0.01	0.12	0.24	0.26	0.14	-0.05	0.16	0.24	0.01	-0.29	0.19	0.25	0.12	1.00							
15	Segregation in the pipeline	0.93	0.41	0.52	0.12	0.48	0.25	0.26	-0.03	0.31	-0.10	0.45	0.45	0.43	0.05	1.00						
16	Research sectors	0.29	0.69	0.63	0.35	0.97	0.24	0.33	0.34	0.19	0.05	0.75	0.57	0.46	0.21	0.52	1.00					
17	Career progression	0.11	0.46	0.40	0.23	0.20	0.98	0.84	0.37	0.11	-0.09	0.71	0.36	0.51	0.01	0.27	0.28	1.00				
18	Decision-making	0.11	0.13	0.67	0.23	0.14	0.21	0.15	0.70	0.86	0.65	0.39	0.07	0.50	0.03	0.15	0.27	0.20	1.00			
19	Research participation	0.32	0.60	0.68	0.14	0.60	0.50	0.61	0.33	0.40	-0.11	0.79	0.88	0.88	0.22	0.51	0.65	0.57	0.35	1.00		
20	Gender dimension in research and innovation content	0.01	0.12	0.24	0.26	0.14	-0.05	0.16	0.24	0.01	-0.29	0.19	0.25	0.12	1.00	0.05	0.21	0.01	0.03	0.22	1.00	
21	She Figures Index	0.40	0.50	0.79	0.36	0.51	0.42	0.52	0.56	0.53	0.04	0.73	0.58	0.68	0.63	0.55	0.64	0.48	0.58	0.75	0.63	1.00

Note: The shading shows values of =<-0.4 and =>0.4, for red and green shading, respectively, to visualise the correlation structure

Figure 9 Assessment of robustness of the scores and ranks by dimension

	Scores							Ranks						
	She Figures Index	Without segregation in the pipeline	Without research sectors	Without career progression	Without decision-making	Without research participation	Without the gender dimension	She Figures Index	Without segregation in the pipeline	Without research sectors	Without career progression	Without decision-making	Without research participation	Without the gender dimension
SE	87.5	88.8	86.6	86.4	86.6	91.3	84.9	1	1	1	1	2	1	7
HR	85.0	85.6	83.2	82.2	90.1	84.9	84.3	2	3	2	2	1	3	8
LT	84.8	85.8	83.0	82.1	83.9	86.2	87.9	3	2	3	3	3	2	3
FI	81.0	82.0	80.0	78.5	78.9	83.6	83.0	4	4	4	5	5	4	10
ES	79.8	80.9	76.9	78.7	79.3	80.5	82.6	5	5	5	4	4	6	12
IE	78.5	77.4	75.9	75.9	75.0	81.1	85.8	6	7	6	7	9	5	5
DK	78.1	78.3	75.4	77.1	75.1	80.1	82.7	7	6	7	6	8	7	11
BG	78.0	75.5	74.2	73.7	75.9	78.9	89.9	8	11	8	10	7	8	2
PT	77.2	76.4	74.1	74.8	76.2	76.2	85.7	9	9	9	9	6	11	6
NL	76.8	76.5	74.0	75.1	72.5	78.6	83.9	10	8	10	8	14	9	9
EE	75.4	74.5	73.0	71.0	74.9	76.6	82.5	11	12	11	11	10	10	13
LV	75.2	73.7	71.2	70.3	70.4	75.3	90.2	12	14	12	12	15	12	1
PL	73.2	74.5	69.7	69.9	73.1	73.4	78.6	13	13	14	13	13	14	19
SK	72.8	72.9	70.4	68.1	70.2	74.7	80.8	14	15	13	15	16	13	15
RO	72.6	67.8	67.9	67.6	73.3	72.4	86.3	15	20	15	18	12	15	4
EL	71.8	68.0	67.6	69.8	74.5	71.7	78.8	16	19	16	14	11	19	18
SI	71.5	75.7	67.5	66.7	67.4	71.8	79.6	17	10	17	19	19	18	16
IT	71.1	69.2	67.2	67.6	67.8	72.4	82.1	18	17	18	17	17	16	14
BE	69.9	72.0	65.9	67.9	64.6	70.2	79.2	19	16	19	16	24	20	17
AT	69.1	69.0	65.8	66.2	65.4	72.3	75.9	20	18	20	20	22	17	20
HU	67.3	66.5	63.9	65.0	67.4	68.5	72.3	21	22	21	21	18	21	23
FR	63.9	60.9	59.2	58.3	66.2	63.3	75.4	22	25	22	24	21	25	21
DE	62.8	63.3	59.0	59.6	58.2	64.0	72.4	23	23	23	23	26	23	22
MT	62.6	66.6	57.5	56.0	63.9	63.6	68.2	24	21	24	25	25	24	24
CY	62.1	61.0	55.9	61.2	66.9	65.2	62.3	25	24	26	22	20	22	27
CZ	60.3	59.3	56.6	53.0	64.9	61.2	66.9	26	26	25	27	23	26	26
LU	58.9	57.8	54.8	55.8	57.6	59.3	67.8	27	27	27	26	27	27	25

Note: shading relates to the score/rank within each dimension, with green indicating the highest score/rank and red showing the lowest score/rank.

References

Aad, G., et al., Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7\ and\ 8$ TeV with the ATLAS and CMS Experiments, Phys. Rev. Lett. 114, article 191803, 2015, https://arxiv.org/pdf/1503.07589.pdf

All Science Journal Classification Codes (ASJC), https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/

Bagihole, B., *NORFACE gender equality workshop address*, 2005, Rejkjavik, Iceland.

Bar-Ilan, J., *Informetrics at the beginning of the 21st century—A review*, Journal of Informetrics, 2 (1), 2008, Pages 1-52.

Berry, A.K. and Zucker, I., *Sex bias in neuroscience and biomedical research*, Neuroscience & Biobehavioural Reviews, Volume 35, Issue 3, January 2011, Pages 565-572.

Burri, S. and Aune, H., European Network of Legal Experts in the Field of Gender Equality, Sex Discrimination in Relation to Part-Time and Fixed-Term Work, 2013, https://publications.europa.eu/en/publication-detail/-/publication/83fb80e9-9444-42bb-b35d-f9fa979b2900/language-en

Cameron, I., Synnott, J., Beisiegel, U., O'Carroll, C., Esposito, F., Harrap, K. A., Israel, N., Modjeska, N., Predescu, R., Prijic-Samarzija, S. and Vandevelde, K., *Shaping the future of the Human Resources Strategy for Researchers – HRS4R*, Brussels, 2015, https://cdn1.euraxess.org/sites/default/files/policy_library/experts-report-strengthened-hrs4r-9-2015_0.pdf

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A New European Innovation Agenda [2022] COM/2022/332. , https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0332

Council of the European Union, Council conclusions on women and the economy: Economic independence from the perspective of part-time work and self-employment [pdf], 2014, http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/lsa/143269.pdf

Council of the European Union, Council conclusions on enhanced measures to reduce horizontal gender segregation in education and employment [pdf], 2017, http://data.consilium.europa.eu/doc/document/ST-15468-2017-INIT/en/pdf

DG for Internal Policies, Precarious Employment in Europe Part 1: Patterns, Trends and Policy Strategies, Brussels, 2016, http://www.europarl.europa.eu/RegData/etudes/STUD/2016/587285/IPOL_STU(2016)587285_EN.pdf

DG Research, *Mapping the maze: Getting more women to the top in research* [pdf], Office for Official Publications of the European Communities, Luxembourg, 2008,

http://ec.europa.eu/research/science-society/document_library/pdf_06/mapping-the-maze-getting-more-women-to-the-top-in-research_en.pdf

DG Research, *Women in science and technology: Creating sustainable careers* [pdf], Office for Official Publications of the European Communities, Luxembourg, 2009, http://ec.europa.eu/research/science-society/document_library/pdf_06/wist2_sustainable-careers-report_en.pdf

DG Research and Innovation, *Towards a European Framework for research careers*, 2011, https://cdn5.euraxess.org/sites/default/files/policy_library/towards_a_european_framework_for_research_careers_final.pdf

DG Research and Innovation, *Researchers' Report 2012*, 2012, http://ec.europa.eu/euraxess/pdf/research_policies/121003_The_Researchers_Report_2012_FINAL_REPORT.pdf

DG Research and Innovation, *HORIZON 2020 in Brief* [pdf], Publications Office of the European Union, Luxembourg, 2014, http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_FinalBAT.pdf

DG Research and Innovation, Horizon Europe; The EU Research and Innovation investment programme (2021-2027), 2019, https://research-and-innovation.ec.europa.eu/system/files/2022-06/ec_rtd_he-investing-to-shape-our-future_0.pdf

DG Research and Innovation, *She Figures 2018*, Publications Office of the European Union, Luxembourg, 2019, https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en

DG Research and Innovation, *Horizon Europe Strategic Plan (2021-2024)* , Luxembourg, 2020, https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

DG Research and Innovation, *Horizon Europe strategic plan 2025-2027*, Publications Office of the European Union, 2024, https://data.europa.eu/doi/10.2777/092911

DG Research and Innovation, *She Figures 2021*, Publications Office of the European Union, Luxembourg, 2021,https://op.europa.eu/en/publication-detail/-/publication/d9fbd9da-4da0-11ec-91ac-01aa75ed71a1/language-en/format-PDF/source-293651619

DG Research and Innovation, *European Research Area Policy Agenda – Overview of actions for the period 2022-2024*, Publications Office of the European Union, Luxembourg, 2021, https://op.europa.eu/en/publication-detail/-/publication/490ee6ca-aa58-11ec-83e1-01aa75ed71a1

Elsevier, Gender in the Global Research Landscape, 2017, https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

European Commission's Expert Group on Gender and Employment (EGGE), *Gender segregation in the labour market: Root causes, implications and policy responses in the EU* [pdf], Publications Office of the European Union, Luxembourg, 2009, http://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=364&furtherPubs=yes

European Commission, Supporting growth and jobs – an agenda for the modernisation of Europe's higher education systems, 2011, COM(2011)567 final, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0567:FIN:EN:PDF

European Commission, Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of The Regions, A Reinforced European Research Area Partnership for Excellence and Growth, 2012, COM(2012)0392, https://eige.europa.eu/sites/default/files/era-communication_en_2012.pdf

European Commission, Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of The Regions Entrepreneurship 2020 Action Plan, 2013a, COM(2012)0795, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0795&from=EN

European Commission, Gendered Innovations – How gender analysis contributes to research: report of the expert group 'Innovation through gender', 2013, https://op.europa.eu/en/publication-detail/-/publication/d15a85d6-cd2d-4fbc-b998-42e53a73a449

European Commission, *Report on progress on equality between women and men in 2013* [pdf], SWD(2014) 142 final, 2014, https://op.europa.eu/en/publication-detail/-/publication/cae38103-6104-4a90-9771-17996c3d9762/language-en#:~:text=The%20Staff%20Working%20Document%20on,of%20equal%20value%3B%20equality%20in

European Commission, *Human resources in science and technology*, 2014, at: http://ec.europa.eu/eurostat/cache/metadata/en/hrst_esms.htm

European Commission, Innovation Union Scoreboard 2014, 2014, https://op.europa.eu/en/publication-detail/-/publication/69a64699-18d7-40b9-8f92-1db3226cd2ec

European Commission, *Gender Equality in Horizon 2020*, 2014, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/gender/h2020-hi-guide-gender_en.pdf

European Commission, *High-tech statistics*, 2015, http://ec.europa.eu/eurostat/statistics-explained/index.php/High-tech_statistics#Database

European Commission, Human resources in science and technology: Statistics explained, 2015, http://ec.europa.eu/eurostat/statistics-explained/index.php/Human_resources_in_science_and_technology#Women_in_science_and_technology

European Commission, Science, technology and innovation: Overview, 2015, http://ec.europa.eu/eurostat/web/science-technology-innovation/overview

European Commission, Social Dimension of Europe (Reflection paper), 2017, https://commission.europa.eu/system/files/2020-07/reflection-paper-social-dimension-europe_en.pdf

European Commission, Towards a sustainable Europe by 2030 (Reflection paper), 2019, https://commission.europa.eu/system/files/2019-02/rp_sustainable_europe_30-01_en_web.pdf

European Commission, *Women in Digital - Shaping Europe's Digital Future*, 2020, https://digital-strategy.ec.europa.eu/en/library/women-digital#:~:text=The%20future%20is%20about%20digital,equitable%20digital%20economy%20and%20society.\

European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the regions on achieving the European Education Area by 2025, COM(2020) 625 final, 2020, https://ec.europa.eu/education/sites/default/files/document-library-docs/communication-european-education-area.pdf

European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions – A new ERA for Research and Innovation, COM(2020) 628 final, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0628&from=EN

European Commission, Directorate-General for Research and Innovation. European Research Area Policy Agenda – Overview of actions for the period 2022-2024 [2021], https://research-and-innovation.ec.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

European Commission, Gendered Innovations 2: How Inclusive Analysis contributes to Research and Innovation, 2020, https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en

European Commission, Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL to strengthen the application of the principle of equal pay for equal work or work of equal value between men and women through pay transparency and enforcement mechanisms, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0093

European Commission, Gender Equality Strategy 2020-2025, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0152

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0152European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities, 2022, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0016&qid=1713516266208

European Institute for Gender Equality (EIGE), *Gender Statistics Database*, https://eige.europa.eu/gender-statistics/dgs

European Institute for Gender Equality (EIGE), *Gender Equality Index*, http://eige.europa.eu/content/gender-equality-index#/domain/power

European Institute for Gender Equality (EIGE), *Gender Equality in Academia and Research* [pdf], Publications Office of the European Union, Luxembourg, 2016, https://eige.europa.eu/sites/default/files/mh0716096enn.pdf

European Parliament, Commission Implementing Regulation on statistics science and technology No 995/2012, 2012, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0018:0030:EN:PDF

European Parliament, Gender pay gap: Parliament adopts new rules on binding pay-transparency measures (Press release). 2023, 20230327IPR78545_en.pdf (europa.eu)

Eurostat, *Key figures on Europe*, 2017 Edition, 2017, https://op.europa.eu/en/publication-detail/-/publication/ec0bf3c8-bac1-11e7-a7f8-01aa75ed71a1/language-en/format-PDF/source-199900535

Eurostat, Statistics Explained: Europe 2020 indicators – R&D and innovation, 2020, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_indicators_-_R%26D_and_innovation&oldid=335438

Eurostat, Gender pay gaps in the European Union –a statistical analysis, 2021, https://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/-/ks-tc-22-002

Eurostat, Statistics Explained: Education statistics, 2023, http://ec.europa.eu/eurostat/statistics-explained/index.php/Education_statistics#Data_sources_and_availability

Eurostat, Statistics Explained: *Tertiary education statistics*, 2023, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics

Gasser C.E. and Shaffer K.S., *Career Development of Women in Academia: Traversing the Leaky Pipeline*, The Professional Counselor, Volume 4, Issue 4, Pages 332-352, 2014, https://mdsoar.org/bitstream/handle/11603/5410/GASSER%20&%20SHAFFER%20-%20WOMEN%20IN%20ACADEMIA.pdf?sequence=3

Gvozdanović, J. and Maes, K., *Implicit bias in academia: A challenge to the meritocratic principle and to women's careers - And what to do about it*, League of European Research Universities (LERU) Advice Paper No 23, Leuven, 2018, https://www.leru.org/files/implicit-bias-in-academia-full-paper.pdf

International Labour Organization, *International classification of occupations (ISCO-08)*, 2012, http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_172572.pdf

Jappelli, T., Nappi, C.A., Torrini, R., *Gender effects in research evaluation*, Research Policy, 46 (5), 2017, 911-924

OECD/European Union/EC-JRC, Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Publishing, Paris, 2008, https://doi.org/10.1787/9789264043466-en

OECD, Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, 2015, http://dx.doi.org/10.1787/9789264239012-en

Snieska V et all, Education and unemployment in European Union economic cycles, 2015, https://www.sciencedirect.com/science/article/pii/S1877042815057833

Stanford University, '*Disparities between women and men*', https://genderedinnovations.stanford.edu/institutions/disparities.html [Last accessed on 4/10/2023]

Summary of the Strasbourg Agreement concerning the International Patent Classification, 1971, https://www.wipo.int/treaties/en/classification/strasbourg/summary_strasbourg.html

UN Women, *Concepts and definitions*, http://www.un.org/womenwatch/osagi/conceptsandefinitions.htm

UNESCO, International standard classification of education, 2012, http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

UNESCO, ISCED Fields of Education and Training 2013, 2014, http://uis.unesco.org/sites/default/files/documents/isced-fields-of-education-and-training-2013-en.pdf

Vinkenburg, C.J., van Engen, M., Peters, C. P., *Promoting new norms and true flexibility: Sustainability in combining career and care*. In: de Vos, A., & van der Heijden, B., Handbook of research on sustainable careers, 131-145, London: Elgar, 2015, https://ideas.repec.org/h/elg/eechap/15416_9.html

World Intellectual Property Organization (WIPO), *International Patent Classification. Version 2015.01*, 2015, http://www.wipo.int/classifications/ipc/en/

World Intellectual Property Organization (WIPO), *Identifying the gender of PCT inventors*, 2016, http://www.wipo.int/publications/zh/details.jsp?id=4125

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union.
You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website
(european-union.europa.eu).

EU Publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions,
go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes.
The portal also provides access to a wealth of datasets from European countries.

The She Figures Handbook (2024) provides methodological guidance on the calculation of indicators included in the She Figures 2024 publication. It is intended to strengthen the capacity of stakeholders to systematically produce meaningful, systematic data on gender in research and innovation.

Organised by data source, information provided on each indicator includes a brief definition, rationale, computation method and any comments or critical issues for the reader to take into account The handbook also includes a section on the verification and validation of data that outlines coherence checks and additional data considerations taken into consideration in the computation and interpretation of indicators. Finally, the annexes outline important information regarding classification standards to which data for several of the indicators are tied, as well as key terminology and definitions.

Studies and reports

Publications Office
of the European Union