

# The gender dimension in research

Mathias W. Nielsen  
Associate Professor  
Department of Sociology  
University of Copenhagen

UNIVERSITY OF  
COPENHAGEN



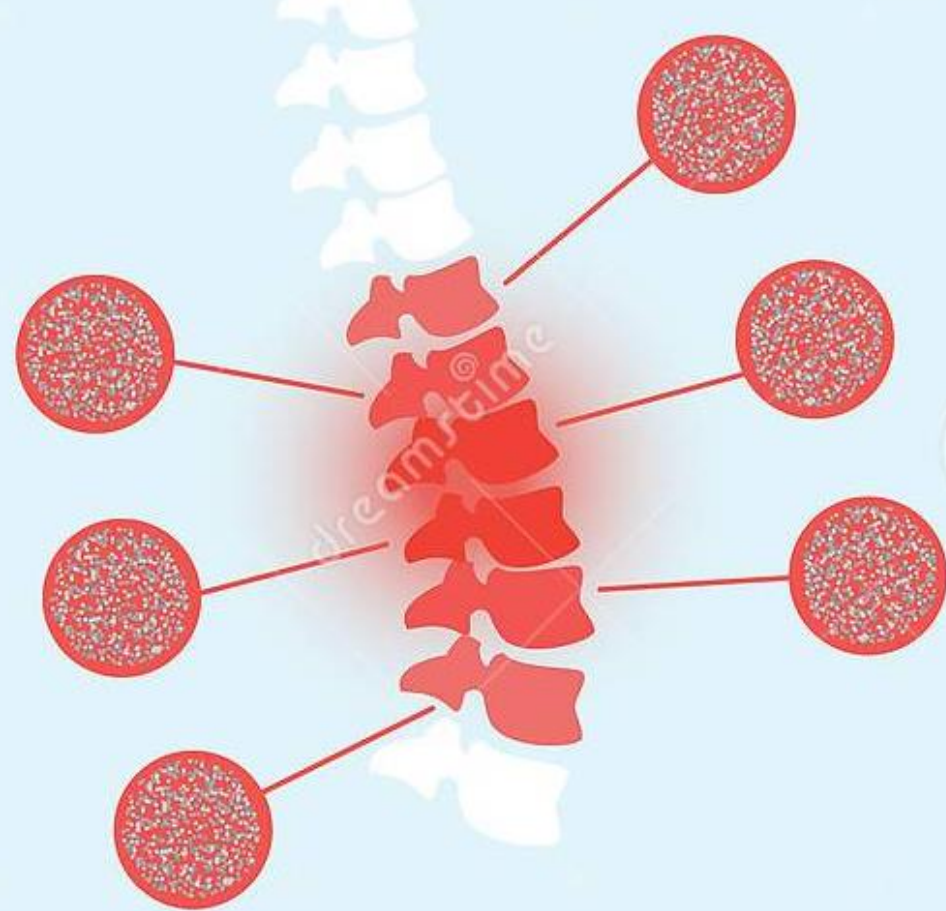
**Gendered  
Innovations** in Science,  
Health & Medicine,  
Engineering, and  
Environment



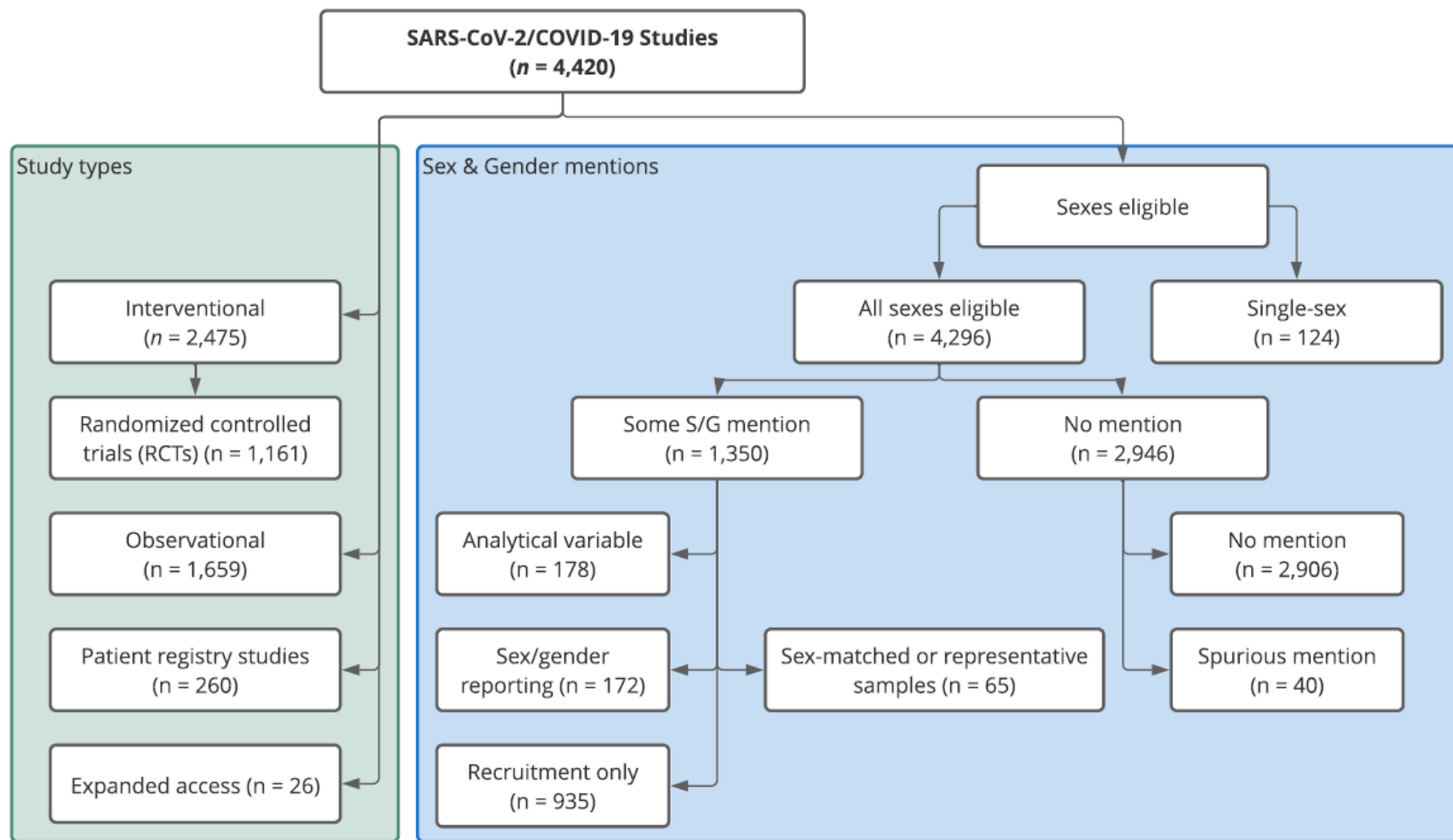
# Agenda

- 1. Concrete examples of the importance of the gender dimension
- 2. Methods on how to include the gender dimension in research
- 3. The mutual benefits of promoting gender equality and the gender dimension

# THE IMPORTANCE OF THE GENDER DIMENSION





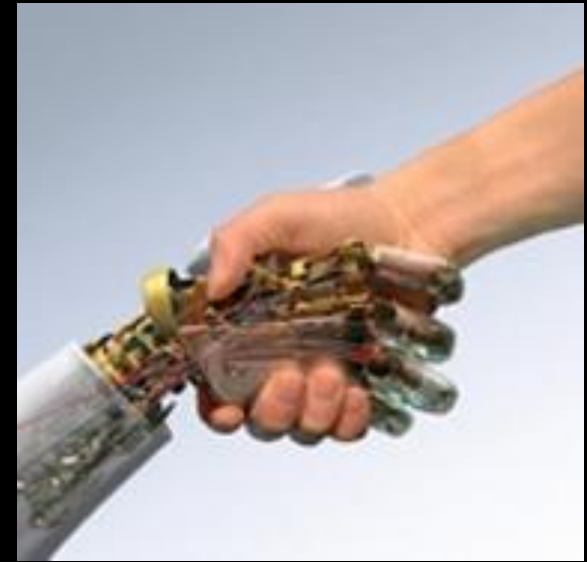


## Published trials:

8 of 45 pharmacological trials on COVID report sex disaggregated results.

Brady, E., Nielsen, M. W., Andersen, J. P., & Oertelt-Prigione, S. (2021). Lack of consideration of sex and gender in COVID-19 clinical studies. *Nature Communications*, 12(1), 1-6..





[genderedinnovations.stanford.edu](http://genderedinnovations.stanford.edu)

# METHODS ON HOW TO INCLUDE THE GENDER DIMENSION



What is **Gendered Innovations**?

## SEX & GENDER ANALYSIS

General Methods

**Specific Methods**

Terms

Checklists

## CASE STUDIES

Science

Health & Medicine

Engineering

Environment

## INTERSECTIONAL DESIGN

POLICY  
RECOMMENDATIONS

VIDEOS



der Methods for Design | Gendered Innovations [>](#)

ENVIRONMENT  
ENGINEERING  
HEALTH & MEDICINE  
SCIENCE



Why Gendered Innovations?

**Gendered Innovations** employs methods of sex, gender, and intersectionality.



## **GENDERED INNOVATIONS 2: How Inclusive Analysis Contributes to Research and Innovation**

Policy Review

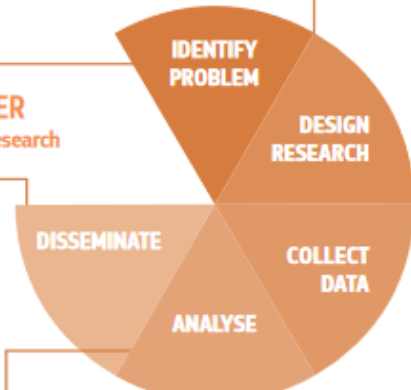


## ANALYSING GENDER

- Gender may play a role in all studies involving human (Tannenbaum et al., 2019).
- Perform literature searches with adequate terms for "gender" and "sex" (Oertelt-Prigione et al., 2010).
- Consider the project's relevance in a relation to different gender identities, norms, and relations.
- Consider relevant factors intersecting with gender (age, socio-economic status, ethnicity, etc.).
- Reflect upon your own gender assumptions in relation to the project.
- Consider what opportunities may be missed as a result of failing to analyse gender and intersecting factors.

- Consider how to involve diverse groups of research subjects/end-users at various steps in the project life-cycle to ensure inclusive solutions.
- Consider which method (qualitative and quantitative) are suited for examining the gender dimensions of relevance to your project.
- In quantitative research, calculate appropriate sample sizes for gender comparisons (Sell, 2017).
- When measuring gender in survey research, ensure that your instrument has been psychometrically validated in the target population (Steenkamp and Baumgartner, 1998).
- Inspect your analytical concepts, categories, and theoretical models for misguided or stereotypical assumptions.
- Consider the risk of stereotypical or excluding relevant groups.

## ANALYSING GENDER enhances all phases of research



- Report sample characteristics by gender, sex, and relevance intersecting variables.
- Report how information on gender identity was obtained.
- Disaggregate reported results by sex and gender.
- Report all results positive, negative, and inconclusive.
- Ensure that gender variations are properly in tables, figures, and conclusions.
- Avoid overemphasizing gender differences. Are the observed variations of practical significance? (Nelson, 2017).
- Consider following the SAGER publication guidelines (Heidari et al., 2016).

- Conduct analyses of relevant factors related to gender norms, gender identity and gender relations (Nielsen et al., 2020).
- When using existing data consider cultural or institutional contents in which the data were generated for potential gender biases.
- Examine similarities between groups (i.e. men, women, and gender-diverse individuals) and variations within groups (Hyde, 2005).
- Examine how observed differences between women, men and gender-diverse individuals relate to gender norms and relations.
- Examine how observed gender differences vary by factors such as age, ethnicity, socioeconomic status.
- In longitudinal studies, examine how observed gender variations evolve over time.
- In qualitative analysis, consider how gender norms, identities and relations intersect to shape people's experiences, opportunities and practices.

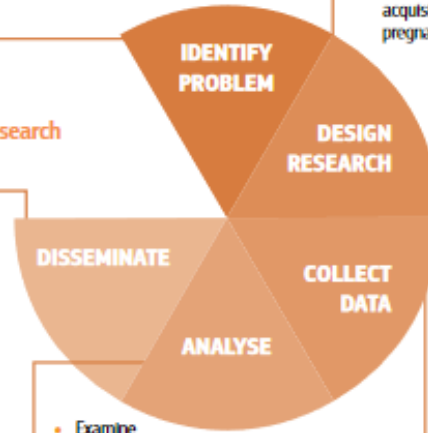
- Collect data across gender characteristics (e.g. gender norms, gender identities, and gender relations) in intersecting factors.
- In survey research, the two-step approach to collect use data on gender identity and birth sex (Deutsch et al., 2013). Ensure that all participants feel safe disclosing their gender identity.
- Ensure equal access for women, men and gender-diverse individuals. Is oversampling needed to ensure that a sufficient number of gender-diverse individuals participate in the study? (Vaughan, 2017).
- Consider how gender relations between researchers and participants may impact the data collection (Chapman et al., 2018).

## ANALYSING SEX

- Sex may play a role in all studies involving human or non-human animals.
- Perform a literature review to identify how sex may be of relevance to your study (Moerman et al., 2009).
- Consider whether sex is a covariate, confounder, or explanatory variable.
- Consider the relevance of sex-related characteristics (e.g. genetic, physiological, hormonal, anthropometric, biomechanical, injury thresholds, levels of pain tolerance, etc.) (Tannenbaum et al., 2019).
- Consider how sex-related factors interact with gender, ethnicity, age, socioeconomic status, lifestyle, etc.
- Consider what opportunities have been missed in the past as a result of failing to analyse sex.

- Sex may serve as a direct explanatory factor or act as a potential modulator for associations between other factors; drawing a causal diagram helps make underlying assumptions explicit (see e.g. Buckley et al., 2017).
- In experimental studies, consider factorial designs to reduce the sample size required for sex-based comparisons (Buch et al., 2017; Miller et al., 2019).
- Consider how sex should be conceptualised in data collection, including physiological, hormonal, anthropometric, or biomechanical aspects (Tannenbaum et al., 2019).
- In longitudinal research, consider how reproductive history may influence the cohort under investigation; e.g. will data acquisition be impacted if females get pregnant during the study?

## ANALYSING SEX enhances all phases of research



- Report the sex of your subjects, even in single-sex studies.
- Report the sex distribution of cells, animals and humans.
- Report how information on sex was obtained.
- Disaggregate reported results by sex.
- Ensure that sex variations are properly visualised in the tables, figures, and conclusions.
- Avoid overemphasising sex differences. Are observed sex differences of practical significance? (Maney et al., 2016; Ribbon et al., 2014).
- Report all results: positive, negative, and inconclusive.
- Consider following the SAGER publication guidelines (Heidari et al., 2016).

- Examine overlaps between and variations within groups of different sexes (see, e.g., Maney et al., 2016).
- Consider the source of any sex difference observed, including the role of environmental, genetic, hormonal, or anthropometric factors.
- When examining sex differences, adjust for possible intersecting and confounding factors (e.g. age). Overlooking confounding factors may result in overemphasising sex differences.
- In longitudinal studies, examine how observed sex variations evolve over time.
- Analyse how observed sex differences may vary by factors such as age, ethnicity, socioeconomic status.

- Consider how to collect information on intersex subjects and hermaphrodite animals.
- Include adequate samples of females and males and, where relevant, intersex or hermaphrodites of different configurations.
- Record information on factors that intersect with sex (e.g. age, life-style, socioeconomic status).
- In experiments, consider how the sex of the researcher may impact research outcomes (Chapman et al., 2018).
- In survey research, questions about gender should not be used as a proxy for birth sex.
- In product and systems design, data collection must consider anthropometric, biomechanical, and physiological factors that vary by sex (Tannenbaum et al., 2019; Jingwen et al., 2012).

## INTERSECTIONAL APPROACHES

An intersectional approach (see Section 2.3) is important to consider when setting research priorities, developing hypotheses and formulating study designs. Taking an intersectional approach can better predict variations in health outcomes and determine user needs, and ultimately lead to more inclusive research and engineering solutions (Faulkner, 2004; Weber and Fore, 2007). For example, sex, socioeconomics, gendered divisions of labour and language interact to determine how agricultural workers are exposed to endocrine disruptors (see Gendered Innovations 1 case study 'Environmental chemicals'). Recent research also demonstrates how an intersectional approach can improve the accuracy of AI-based facial recognition (see case study 'Facial recognition') and energy-efficiency measures (see case study 'Smart energy solutions').

### Identify problem

Intersectional approaches may be relevant in studies involving human subjects. While sex and gender are important concepts to consider (see 'Analysing sex' and 'Analysing gender' above), they are shaped by other social and biological factors. The way the research problem is formulated will determine which intersecting variables are required for analysis. The most important categories, factors and relationships cannot be determined a priori, but emerge in the process of investigation (Hankivsky, 2014).

Before beginning a study, researchers should conduct systematic literature searches to identify factors and categories of potential relevance. These categories and factors can be biological, sociocultural or psychological aspects of users, customers, participants, experimental subjects or cells (see box right). Even intersecting factors, such as diet or genetic variability, may be important to consider. Such factors may reveal subgroup differences among males and females that would have been obscured by using only sex as a variable (see Gendered Innovations 1 case study 'Nutrigenomics').

### Factors to consider in an intersectional analysis may include:

- ▶ gender
- ▶ disabilities
- ▶ ethnicity
- ▶ race
- ▶ age
- ▶ geographic location
- ▶ socioeconomic status
- ▶ nationality
- ▶ sexual orientation
- ▶ LGBTI+ identity
- ▶ religion
- ▶ educational background
- ▶ lifestyle
- ▶ language
- ▶ family configuration
- ▶ environment
- ▶ genetics
- ▶ sex hormones
- ▶ reproductive status
- ▶ body composition
- ▶ comorbidities
- ▶ body size.

In this phase, it is also important to consider the social contexts, including societal, institutional and community-level circumstances (e.g. laws, policies, healthcare providers, school systems, law enforcement, religious institutions,

# Field-specific methods

## Health and biomedicine

ANALYSING SEX IN BIOMEDICINE

ANALYSING GENDER IN HEALTH AND BIOMEDICINE

ANALYSING SEX IN TISSUES AND CELLS

ANALYSING SEX IN LAB ANIMAL RESEARCH

## Information and communication technologies

ANALYSING GENDER AND INTERSECTIONALITY IN MACHINE LEARNING

ANALYSING GENDER AND INTERSECTIONALITY IN SOCIAL ROBOTICS

## Climate change

ANALYSING SEX IN HERMAPHRODITIC SPECIES

## Urban planning/transportation

GENDER IMPACT ASSESSMENT

## Innovation

NORM-CRITICAL INNOVATION

# ASKING ABOUT GENDER AND SEX IN SURVEYS

## Box 1. The one-step method (example)

Are you male or female?

- ☐ Male
- ☐ Female

Source: American National Election Survey 2008–2009  
(Westbrook and Saperstein, 2015)

## Box 2. The two-step method

### Birth Sex

What sex were you assigned at birth?

- ☐ Female
- ☐ Male
- ☐ Intersex
- ☐ A sex not listed here  
(please specify)

☐ Prefer not to state

### Current Gender Identity

What is your current gender identity?  
(Please select all that apply)

- ☐ Woman
- ☐ Man
- ☐ Non-binary
- ☐ Genderqueer
- ☐ A gender identity not listed here  
(please specify)

☐ Prefer not to state



# Sex and gender analysis improves science and engineering

<https://doi.org/10.1038/s41586-019-1657-6>

Received: 8 January 2019

Accepted: 27 August 2019

Published online: 6 November 2019

Cara Tannenbaum<sup>1,2</sup>, Robert P. Ellis<sup>3,4</sup>, Friederike Eysset<sup>3,5</sup>, James Zou<sup>4,6,7</sup> & Londa Schiebinger<sup>4,8</sup>

The goal of sex and gender analysis is to promote rigorous, reproducible and responsible science. Incorporating sex and gender analysis into experimental design has enabled advancements across many disciplines, such as improved treatment of heart disease and insights into the societal impact of algorithmic bias. Here we discuss the potential for sex and gender analysis to foster scientific discovery, improve experimental efficiency and enable social equality. We provide a roadmap for sex and gender analysis across scientific disciplines and call on researchers, funding agencies, peer-reviewed journals and universities to coordinate efforts to implement robust methods of sex and gender analysis.



Anniversary  
collection:  
[go.nature.com/nature150](http://go.nature.com/nature150)

Integrating sex and gender analysis into the design of research, where relevant, can lead to discovery and improved research methodology. A deeper understanding of the genetic and hormone-mediated basis for sex differences in immunity, for example, promises insights into novel cancer immunotherapies<sup>1</sup>. Evidence that facial recognition systems misclassify gender more often for darker-skinned women than for lighter-skinned men has led to refinements in computer vision<sup>2</sup>. Understanding sex-based responses to climate change allows better modelling of demographic change among marine organisms and the downstream effects for humans<sup>3,4</sup>. Sex or gender analysis can be critical to the interpretation, validation, reproducibility and generalizability of research findings (Box 1).

The documented importance of sex and gender analysis in research has underwritten policy change at major funding agencies. New policies have been implemented at the Canadian Institutes of Health Research (2010), European Commission (2014), US National Institutes of Health (2016), German Research Foundation (2020), among others. Concurrently, peer-review journals have implemented editorial guidelines to evaluate the rigour of sex and gender analysis as one criterion among many when selecting manuscripts for publication. The goal is to increase transparency, promote inclusion and reset the research default to carefully consider sex and gender, where appropriate.

In this Perspective, we discuss how incorporating sex and/or gender analysis into research can improve reproducibility and experimental efficiency, help to reduce bias, enable social equality in scientific outcomes and foster opportunities for discovery and innovation. From highlighted examples, we extract decision-tree roadmaps for researchers across

disciplines. We consider the limits to sex and gender analysis and offer recommendations to researchers and funding agencies on how to move the field forward. Throughout this Perspective, we explore how integrating sex and gender analysis into research design has the potential to offer new perspectives, pose new questions and, importantly, enhance social equalities by ensuring that research findings are applicable across the whole of society.

## Reproducibility and efficiency

Reproducibility is important for scientific excellence. One important reason for a lack of reproducibility in experimentation is inconsistency in methodological reporting, which varies widely across disciplines from biology to chemistry, human–robot interaction, medicine, physics, psychology and beyond<sup>5,6</sup>. Sex- and gender-specific reporting is still limited in a range of scientific disciplines. In preclinical microbiology and immunology, a review of published studies using primary cells from diverse animal species (that is, humans and nonhuman vertebrates) revealed that the majority failed to report the sex of donors from which the cells were isolated<sup>7,8</sup>. In marine science, a review of experimental ocean acidification studies showed that only 3.9% of studies statistically assessed sex-based differences, while only 10.5% of studies accounted for possible sex effects by assessing females and males independently<sup>9</sup>. Similarly, in ecotoxicology, a review of omics studies showed that although most reported sex, only 23% (5 out of 22) examined the omics response of each sex to a toxicant<sup>10</sup>. In social robotics, the notion of robot gender, gender-stereotypical domains and their interaction with user gender has only recently become a target of scientific inquiry<sup>11</sup>. A lack of transparency in reporting sex and gender-related variables makes it difficult to reproduce experiments in which these variables affect experimental results.

## Disaggregating the data

Analysing experimental results by sex and/or gender is critical for improving accuracy and avoiding misinterpretation of data (Fig. 1). The common practice of pooling the response of females and males or

# Gender-related variables for health research

Mathias W. Nielsen<sup>1</sup>, Marcia L. Stefanick<sup>2</sup>, Diana Peragine<sup>3</sup>, Torsten B. Neilands<sup>4</sup>, John P. A. Ioannidis<sup>5</sup>, Louise Pilote<sup>6</sup>, Judith J. Prochaska<sup>7</sup>, Mark R. Cullen<sup>6</sup>, Gillian Einstein<sup>3</sup>, Ineke Klinge<sup>7</sup>, Hannah LeBlanc<sup>8</sup>, Hee Young Paik<sup>9</sup> and Londa Schiebinger<sup>10\*</sup>



## Abstract

**Background:** In this paper, we argue for Gender as a Sociocultural Variable (GASV) as a complement to Sex as a Biological Variable (SABV). Sex (biology) and gender (sociocultural behaviors and attitudes) interact to influence health and disease processes across the lifespan—which is currently playing out in the COVID-19 pandemic. This study develops a gender assessment tool—the Stanford Gender-Related Variables for Health Research—for use in clinical and population research, including large-scale health surveys involving diverse Western populations. While analyzing sex as a biological variable is widely mandated, gender as a sociocultural variable is not, largely because the field lacks quantitative tools for analyzing the influence of gender on health outcomes.

**Methods:** We conducted a comprehensive review of English-language measures of gender from 1975 to 2015 to identify variables across three domains: gender norms, gender-related traits, and gender relations. This yielded 11 variables tested with 44 items in three US cross-sectional survey populations: two internet-based ( $N = 2051$ ;  $N = 2135$ ) and a patient-research registry ( $N = 489$ ), conducted between May 2017 and January 2018.

**Results:** Exploratory and confirmatory factor analyses reduced 11 constructs to 7 gender-related variables: caregiver strain, work strain, independence, risk-taking, emotional intelligence, social support, and discrimination. Regression analyses, adjusted for age, ethnicity, income, education, sex assigned at birth, and self-reported gender identity, identified associations between these gender-related variables and self-rated general health, physical and mental health, and health-risk behaviors.

**Conclusion:** Our new instrument represents an important step toward developing more comprehensive and precise survey-based measures of gender in relation to health. Our questionnaire is designed to shed light on how specific gender-related behaviors and attitudes contribute to health and disease processes, irrespective of—or in addition to—biological sex and self-reported gender identity. Use of these gender-related variables in experimental studies, such as clinical trials, may also help us understand if gender factors play an important role as treatment-effect modifiers and would thus need to be further considered in treatment decision-making.

**Keywords:** Gender measures, Biomedical outcomes, Sex differences

\* Correspondence: [schiebinger@stanford.edu](mailto:schiebinger@stanford.edu)

<sup>10</sup>History of Science, Stanford University, Building 200, 450 Jane Stanford Way, Stanford, CA 94305, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the

<sup>1</sup>Institute of Gender and Health, Canadian Institutes of Health Research, Université de Montréal, Montréal, Québec, Canada. <sup>2</sup>College of Life and Environmental Sciences, University of Exeter, Exeter, UK. <sup>3</sup>Center of Excellence Cognitive Interaction Technology, Department of Psychology, Universität Bielefeld, Bielefeld, Germany. <sup>4</sup>Biomedical Data Sciences, Stanford University, Stanford, CA, USA. <sup>5</sup>Chan Zuckerberg Biohub, San Francisco, CA, USA. <sup>6</sup>History of Science, Gendered Innovations in Science, Health & Medicine, Engineering and Environment, Stanford University, Stanford, CA, USA. <sup>7</sup>These authors contributed equally: Cara Tannenbaum, Robert P. Ellis, Friederike Eysset, James Zou. \*e-mail: [schiebinger@stanford.edu](mailto:schiebinger@stanford.edu)



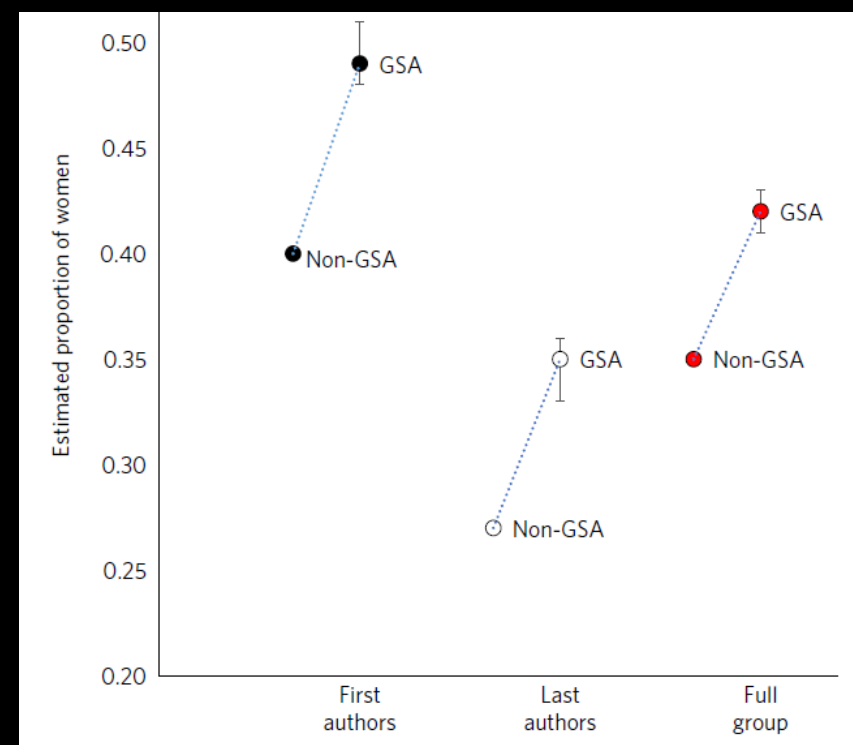
THE MUTUAL BENEFITS OF PROMOTING GENDER EQUALITY AND  
THE GENDER DIMENSION IN RESEARCH

# nature human behaviour

One and a half million medical papers  
reveal a link between author gender and  
attention to gender and sex analysis

Mathias Wullum Nielsen ✉, Jens Peter Andersen, Londa Schiebinger & Jesper W. Schneider

**Fig. 2 | Plot of estimated marginal means.** The estimated marginal means for  $f\_first$ ,  $f\_last$  and  $fw$  in models 1, 2 and 3 are shown. Error bars represent 95% CI (for estimate specifications, see Supplementary Table 5). The plots visualize the participation of women (relative to men) as first authors, last authors and overall representation in the byline for studies that do and do not involve GSA. The figure shows that women's estimated share of authorships is higher in GSA studies than in non-GSA studies for all three author variables.





## Factors affecting sex-related reporting in medical research: a cross-disciplinary bibliometric analysis

Cassidy R Sugimoto, Yong-Yeol Ahn, Elise Smith, Benoit Macaluso, Vincent Larivière

### Summary

*Lancet* 2019; 393: 550–59

See [Comment](#) page 497

School of Informatics,  
Computing, and Engineering,  
Indiana University  
Bloomington, USA  
(C R Sugimoto PhD,  
Y-Y Ahn PhD); École de  
Bibliothéconomie et des  
Sciences de l'Information,  
Université de Montréal, Canada  
(E Smith PhD, V Larivière PhD);  
and Observatoire des Sciences  
et des Technologies, Centre  
Interuniversitaire de Recherche

**Background** Clinical and preclinical studies have shown that there are sex-based differences at the genetic, cellular, biochemical, and physiological levels. Despite this, numerous studies have shown poor levels of inclusion of populations into medical research. These disparities in sex inclusion in research are further complicated by the absence of sufficient reporting and analysis by sex of study populations. Disparities in the inclusion of women in medical research substantially reduce the utility of the results of such research for the entire population. The absence of sex-related reporting are problematical for the translation of research from the preclinical to the applied health settings. Large-scale studies are needed to identify the extent of sex-related reporting and disparities are more prevalent. In addition, while several studies have shown the dearth of female researchers in science, few have evaluated whether a scarcity of women in science might be related to disparities in sex-related reporting. We aimed to do a cross-disciplinary analysis of the degree of sex-related reporting across sciences—from biomedical, to clinical, and public health research—and the role of author gender in sex-related reporting.

## Original Investigation

FREE

July 2018

# Association of Author Gender With Sex Bias in Surgical Research

Nicholas Xiao, BSc<sup>1</sup>; Neel A. Mansukhani, MD<sup>1</sup>; Diego F. Mendes de Oliveira, PhD<sup>2</sup>; [et al](#)

» [Author Affiliations](#) | [Article Information](#)

*JAMA Surg.* 2018;153(7):663–670. doi:10.1001/jamasurg.2018.0040

## RESEARCH

### HISTORY OF INNOVATION

# Who do we invent for? Patents by women focus more on women's health, but few women get to invent

Rembrand Koning<sup>1\*</sup>, Sampsa Samila<sup>2</sup>, John-Paul Ferguson<sup>3</sup>

Women engage in less commercial patenting and invention than do men, which may affect what is invented. Using text analysis of all U.S. biomedical patents filed from 1976 through 2010, we found that patents with all-female inventor teams are 35% more likely than all-male teams to focus on women's health. This effect holds over decades and across research areas. We also found that female researchers are more likely to discover female-focused ideas. These findings suggest that the inventor gender gap is partially responsible for thousands of missing female-focused inventions since 1976. More generally, our findings suggest that who benefits from innovation depends on who gets to invent.

RESEARCH ARTICLE

# Diversifying history: A large-scale analysis of changes in researcher demographics and scholarly agendas

Stephan Risi<sup>1</sup>, Mathias W. Nielsen<sup>2</sup>, Emma Kerr<sup>3</sup>, Emer Brady<sup>4</sup>, Lanu Kim<sup>5</sup>, Daniel A. McFarland<sup>6</sup>, Dan Jurafsky<sup>5</sup>, James Zou<sup>6</sup>, Londa Schiebinger<sup>7</sup>\*

**1** Program in Digital Humanities, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **2** Department of Sociology, University of Copenhagen, Copenhagen, Denmark, **3** Stanford Graduate School of Education, Stanford University, Stanford, California, United States of America, **4** Department of Political Science, Aarhus University, Aarhus, Denmark, **5** Department of Linguistics, Department of Computer Science, Stanford University, Stanford, California, United States of America, **6** Department of Biomedical Data Science, Stanford University, Stanford, California, United States of America, **7** Department of History, Stanford University, Stanford, California, United States of America

\* [schiebinger@stanford.edu](mailto:schiebinger@stanford.edu)



## OPEN ACCESS

**Citation:** Risi S, Nielsen MW, Kerr E, Brady E, Kim L, McFarland DA, et al. (2022) Diversifying history: A large-scale analysis of changes in researcher demographics and scholarly agendas. PLoS ONE 17(1): e0262027. <https://doi.org/10.1371/journal.pone.0262027>

**Editor:** Dorothy Porter, University of California San Francisco, UNITED STATES

**Received:** June 3, 2021

**Accepted:** December 15, 2021

**Published:** January 19, 2022

**Copyright:** © 2022 Risi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All files are available from the OSF database (<https://osf.io/v4ysh/>).

**Funding:** DJ, DAM, LS, JZ were funded for this work through a seed grant from the Stanford University Human-Centered Artificial Intelligence Institute (<https://hai.stanford.edu/>). LK was supported by NSF grant 1633036. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

### Background

In recent years, interest has grown in whether and to what extent demographic diversity sparks discovery and innovation in research. At the same time, topic modeling has been employed to discover differences in what women and men write about. This study engages these two strands of scholarship to explore associations between changing researcher demographics and research questions asked in the discipline of history. Specifically, we analyze developments in history as women entered the field.

### Methods

We focus on author gender in diachronic analysis of history dissertations from 1980 (when online data is first available) to 2015 and a select set of general history journals from 1950 to 2015. We use correlated topic modeling and network visualizations to map developments in research agendas over time and to examine how women and men have contributed to these developments.

### Results

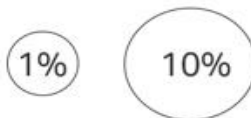
Our summary snapshot of aggregate interests of women and men for the period 1950 to 2015 identifies new topics associated with women authors: gender and women's history, body history, family and households, consumption and consumerism, and sexuality. Diachronic analysis demonstrates that while women pioneered topics such as gender and women's history or the history of sexuality, these topics broaden over time to become methodological frameworks that historians widely embraced and that changed in interesting ways as men engaged with them. Our analysis of history dissertations surface correlations between advisor/advisee gender pairings and choice of dissertation topic.

# Data:

- JSTOR : 10,367 full-text articles (1951 to 2014) from a set of core History journals published in the US: American Historical Review, Journal of American History, Journal of Modern History, Journal of Social History, etc.
- ProQuest: 21,548 history dissertation abstracts written in the US between 1980 (when dissertations are first digitized) and 2015.

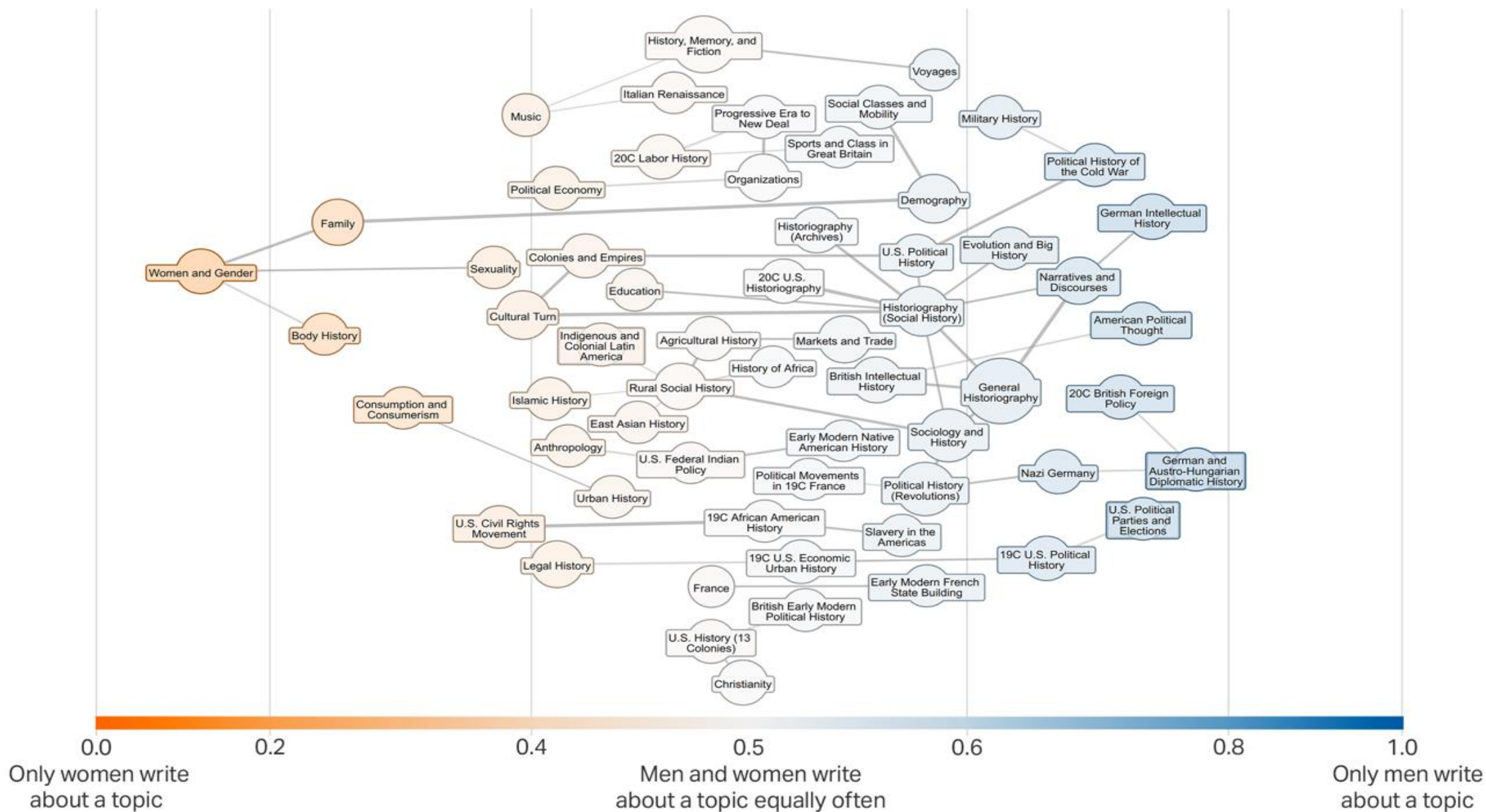
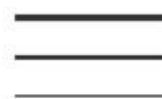


Average topic weight across all articles:



Correlation between topics:

greater than 0.2  
greater than 0.1  
greater than 0.05



THANKS !