

POLICY FORUM

DIVERSITY

A framework for sex, gender, and diversity analysis in research

Funding agencies have ample room to improve their policies

By Lilian Hunt¹, Mathias Wullum Nielsen²,
Londa Schiebinger^{3,4}

National research agencies are responsible for promoting excellent research that benefits all of society (1). Integrating sex, gender, and diversity analysis (SG&DA) into the design of research, where relevant, can improve research methodology, enhance excellence in science, and make research more responsive to social needs (2). National funding agencies—encouraged by scientists and social movements—have thus begun to implement policies to integrate sex, gender, and, more recently, diversity analysis into the grant proposal process, where these factors have been shown to play a role. We develop a five-part analytical framework for implementing and evaluating SG&DA policies, and use it to evaluate the quality of SG&DA policies for 22 major national funding agencies across six continents. By collecting emerging global practices for policy implementation, we seek to improve understanding of these policies and practices in efforts to enhance international collaborations and research excellence.

SG&DA is highly developed in health and biomedicine, machine learning, and artificial intelligence, and is emerging in other fields. Incorporating SG&DA into research design has enabled advancements across numerous disciplines. For example, research on pain documents both biological aspects (i.e., sex differences in electrical, ischemic, thermal, pressure, and muscle pain sensitivity) and cultural aspects (differences in how people report pain and how physicians understand and treat pain). Consensus analyses now in-

clude sex-based analysis in nonhuman animal and human preclinical and clinical research, gender-based analysis in patient-physician relationships and, more recently, analysis of different racialized groups and ethnicities in clinical treatment (3, 4).

SG&DA informs each phase of the research process—from establishing project objectives (e.g., considering the characteristics of target populations and the social implications of the project), to developing methodologies (e.g., ensuring appropriate and unbiased measures and instruments), gathering data (e.g., sampling sufficient participant numbers across categories), analyzing data (e.g., considering within- and between-group differences and intersecting factors) to reporting results (e.g., considering language use and specifying how categorical data were collected and annotated) (5).

There are three pillars of the science infrastructure that need to coordinate policies to achieve excellence in science (2). The first, funding agencies, encourage integrating SG&DA at the beginning of the research process. In 2003, the European Commission (EC) endorsed “questioning systematically whether, and in what sense, sex and gender are relevant in the objectives and methodology of projects.” Other public-funding agencies followed suit with policies implemented at the Canadian Institutes of Health Research (CIHR; 2010), German Research Foundation (DFG; 2020), and National Research Foundation of Korea (NRF; 2021), among others (6).

Pillar two, peer-reviewed journals, increasingly consider SG&DA when selecting manuscripts for publication (7). Pillar three, universities and research institutions, are responsible for developing methods for this type of analysis and for providing this expertise to future generations. Many faculties of science, medicine, and engineering have not yet integrated knowledge of SG&DA into their core curricula; hence, many researchers and research evaluators lack training in these types of analyses, and new guidance is needed to improve this.

This study focuses on funding agencies and develops an analytical framework to evaluate the uptake of policies for integrating sex, gender, and diversity—which covers intersectional characteristics such as age or life course, indigeneity, race and ethnicity, sexuality, socioeconomic status, and other axes of inequality—into research design. Previous studies have analyzed single funding agencies in depth (1, 8–9) or focused narrowly on particular regions such as sub-Saharan Africa, or more often Europe and North America (10, 11). One international study included questions on SG&DA policies in their larger survey focused on gender equity in research teams (12). Ours is the first to develop a cross-disciplinary analytical framework for policy development and to apply it globally.

DEVELOPING THE POLICY FRAMEWORK

We conducted an investigation in three stages: (i) document analysis, (ii) global survey, and (iii) policy analysis. Based on a documented analysis of SG&DA-related policies and guidelines and prior research (6), we developed our analytical framework, a five-part guide for implementing and evaluating SG&DA policy (see the box) (for a full outline of the development process, see supplementary materials (SM) S1). We convened an international advisory group that included representatives from public funders, expert researchers, and policy specialists from five global areas—Africa and the Middle East, Europe and Central Asia, Latin America and the Caribbean, North America, and South and East Asia and the Pacific—to discuss and improve the framework’s clarity, specificity, applicability, and global representativeness (SM S2).

We tested the applicability of our framework for evaluating SG&DA policies through a pilot survey targeting six large public funding agencies in Europe, East Asia, and North America—the CIHR, DFG, EC, Irish Research Council, US National Institutes of Health (NIH), and NRF Korea—four of which had representation on our advisory board. The survey was designed to collect all relevant information on SG&DA-related activities done by the agencies. Evidence consisted of publicly available policies and guidance documents, and, where agencies were in the process of developing policy, internal documents. We checked all self-reported responses against agency documents. The pilot study confirmed the applicability of our evaluation tool and also led to minor revisions in item phrasing.

For the global evaluation, we consulted with our advisory board to select 39 ad-

¹Equality, Diversity and Inclusion in Science and Health (EDIS), Culture, Equity, Diversity, and Inclusion, Wellcome Trust, London, UK. ²Department of Sociology, University of Copenhagen, Copenhagen, Denmark. ³History of Science, Stanford University, Stanford, CA, USA. ⁴Gendered Innovations in Science, Health and Medicine, Engineering, and Environment, Stanford, CA, USA. Email: schiebinger@stanford.edu

ditional funding agencies based on geographic spread with the goal of covering agencies from all continents that host such agencies (SM S4). We focused on public agencies as these are typically the largest, most prestigious, and agenda-setting funders at the national level. Of the 39 invited, 16 agreed, yielding a final sample of 22 (SM S5).

Response rates differed across regions. In North America, all 5 invited agencies participated; in Europe and Central Asia, 9 of 14 agencies responded; in Latin America and the Caribbean, 2 of 3 agencies responded; in South and East Asia and the Pacific, 4 of 7 agencies responded; and in Africa and the Middle East, only 3 of the 11 invited responded (table S1).

As part of our participatory design, funders were invited to complete an online questionnaire (SM S8 to S11) and asked to provide evidence (either publicly available or internally agreed upon) for each answer. Each agency was scored on their performance across the five parts of the policy framework by at least two evaluators using a scoring matrix (SM S6 and S12) vetted by the advisory board. The scores of individual funding agencies are kept confidential per agreement.

Overall, one agency scored in the first tier (81+ points out of 100) in SG&DA policies and measures; six agencies scored in the second tier (61 to 80 points), five in the third tier (41 to 60 points), two in the fourth tier (21 to 40 points), and nine are just beginning (0 to 20 points) (see the figure). Globally, almost half of agencies (9 of 22) provided definitions of terms of reasonable quality (first and second tiers) (see the figure) (table S2). Similarly, almost half of agencies (9 of 22) provided quality proposal guidelines for applicants. Only about a third of agencies (7 of 22) provided quality instructions to evaluations, and even fewer (5 of 22) provided quality trainings for applicants, evaluators, and staff. Evaluation of policy implementation was the weakest, with only two agencies scoring in the second tier and none in the first. Many agencies have begun implementing policy, but most have not considered evaluation. It is important that agencies plan to evaluate policies from the very beginning.

One complicating factor in comparing agencies is that some agencies span all fields of the human and natural sciences, technology, and health and biomedicine, whereas others are more specialized. When we divided our data by agency type, however, we found no pronounced differences, suggesting that funders with wide remit can successfully implement these policies (fig. S1).

Each component of our five-part frame-

A framework to implement and evaluate policies

The framework covers five aspects of public funding agencies' efforts to promote sex, gender, and diversity analysis (SG&DA).

Definition of terms

- Clear and quality definitions
- Definitions readily available

Proposal guidelines for applicants

- Instructions to applicants to include SG&DA
- Encourage or require?
- Examples given
- Specify how SG&DA is included at each stage of the research cycle—detail for yes and justify for no

Instructions for evaluators

- Instructions for reviewers to include SG&DA in their evaluations
- Assessment at each stage of the research process
- Monitoring

Trainings for applicants, evaluators, and staff

- Training, resources, and support available for applicants
- Training, resources, and support available for proposal evaluators
- Training, resources, and support available for relevant agency staff
- Training mandatory through certification
- Development of open access resources: courses and high-quality materials

Evaluation of policy implementation

- Number and proportion of proposals that include SG&DA
- Number and proportion of proposals that include quality SG&DA
- Quality of evaluators' scoring & comments
- Number of applicants, evaluators, & staff who engaged in training
- Number and proportion of publications from funded proposals that include SG&DA

work process for policy development is important to the success of policy implementation. A simple policy mandate to include SG&DA in research is itself not enough, and, when poorly executed, SG&DA can lead to harm (13). Further, funding agency policies are but one part of behavioral and cultural shifts in the research endeavor and need to be supported by coordinated change across the broader science infrastructure.

EMERGING GLOBAL PRACTICES

Each agency will develop country-specific policies that accommodate their cultural practices and regulatory landscapes. At the same time, agencies can share policies and

practices to enhance research collaboration across regions. Here we explore key considerations for each part of our policy framework and track emerging policies. These considerations are drawn from agencies' responses to our questionnaire (SM S7).

Overall, we found differences across countries as to where these policies sit within agencies, how they are implemented, and where accountability for them lies. Most funders in our study—for example, the DFG and EC—include SG&DA under their “excellence” criteria; i.e., SG&DA is judged in relation to its potential to create new knowledge. CIHR, for example, found that once their SG&DA policy was established, applicants who scored well on the SG&DA question scored well overall, i.e., SG&DA improved the overall quality of the proposal, and the proposal was more likely to be funded (9).

Definition of terms

The first step in policy development is definition of terms (see the box). It is important that the same definitions are shared with applicants, evaluators, and staff to support consistency across the agency. For example, the Canadian Tri-Agency provides guidance in both English and French describing Gender-Based Analysis Plus (GBA+) that links across agency materials (all references to policy documents are available in table S3). Which terms and how they are defined will be country specific. For “diversity,” for example, each agency will determine which factors to prioritize and for which funding areas. Some agencies, such as those in Australia, synchronize their guidelines with those set through national legislation. [See (14) for example definitions of key terms, and links between agency policies and this five-part framework.]

We chose to evaluate agencies on “sex” and “gender” analysis because sex and gender were historically the categories included in agency policy—for example, by the EC. Although sex is a biological characteristic of humans and numerous nonhuman organisms and is an important category to continue to call out, gender is at the same epistemic level as other aspects of sociocultural diversity and may, in the future, be folded into diversity analysis. We included both gender and diversity to allow us to evaluate developed and emerging policy practices. Second, we considered evaluating on “intersectional” analysis but, after policy reviews and consultation with our advisory board, judged “diversity” a more generally used term.

Proposal guidelines for applicants

Agencies take three basic approaches in their request to applicants to integrate SG&DA into their proposal, where relevant:

Most encourage applicants to integrate SG&DA; a few require this type of analysis; some only encourage applicants but instruct evaluators to score this element. In this study, more points were awarded to funders who required SG&DA; however, more research is needed to determine the differential impact of these various approaches. In all cases, the “where relevant” is crucial. No agency asks for SG&DA in pure mathematics, for example, where no body of literature has established its relevance.

The trajectory of the EC is of interest. Since 2003, the Commission has encouraged sex and gender analysis, referred to as the “gender dimension,” in research. To strengthen the policy, in 2014, Horizon 2020 flagged topics for which taking the gender dimension into account was mandatory. Since 2021, Horizon Europe has required all proposals to consider sex, gender, and/or intersectional analysis in research, unless otherwise specified.

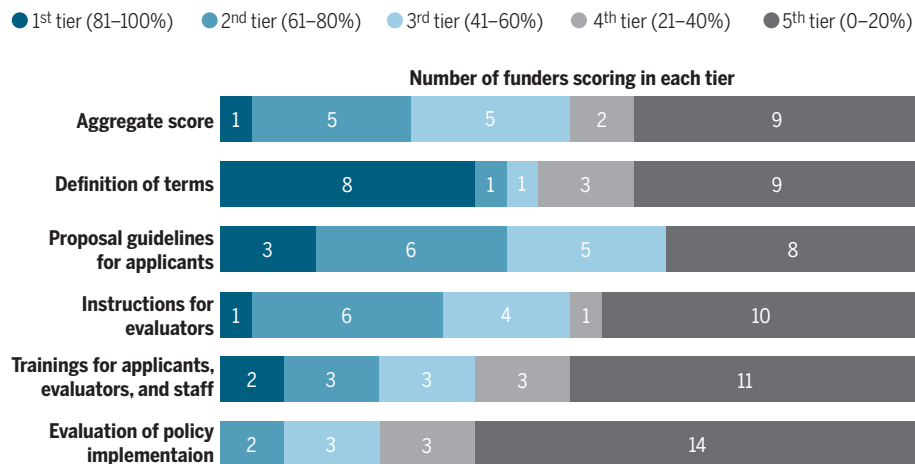
The US NIH has, since 2016, required all applicants to consider “sex as a biological variable” (SABV) and has detailed how this type of analysis supports excellence in science (1). This requirement supplements policies for inclusion of sex and gender, race, and ethnicity in clinical trials launched in the 1990s and inclusion of age (Inclusion Across the Lifespan) added in 2019.

The DFG implemented its SG&DA guidelines in 2020 after a 2-year study period. The DFG encourages but does not require applicants to consider SG&DA, emphasizing that it funds “proposals in curiosity-driven basic research” in fields selected by applicants where freedom of research is core. In so doing, the DFG taps into the issue of academic freedom, where principal investigators set their own research agendas. Notably, SG&DA is one methodological tool among many that researchers may employ. Like any other methodology, appropriate application is key to research outcomes. It is the job of researchers and research evaluators to determine when SG&DA may enhance research outcomes. Interestingly, the DFG requires research evaluators to take SG&DA into account, making SG&DA analysis de facto required if relevant to research outcomes.

In many countries, agency policy is governed by national legislation. In Japan, where the government renews the basic plans for Science, Technology, and Innovation (STI) every 5 years, the sixth STI Basic Plan in 2021 included integrating the gender perspective and gender analysis into research and technology development. Similarly, in 2021, the Republic of Korea

Agencies' performance on the policy framework

Using a scoring matrix (see supplementary materials), funding agencies were awarded points reflective of their performance on each of five parts of the sex, gender, and diversity analysis (SG&DA) policy framework (bottom five bars). Those five component-level scores were aggregated into an overall score (top bar). For each component and the aggregate, agencies were assigned to five tiers based on the percentage of the total possible points that they received for that component or aggregate. Bars reflect the number of agencies (out of 22 in total) whose scores fall within a given tier for that component or aggregate.



passed an amendment of the Framework Act on Science and Technology to include integrating sex and gender into research, which allowed the NRF Korea to develop new policies.

In our study, we evaluated whether applicants are instructed to detail how SG&DA analysis is incorporated into all phases of research. If SG&DA is not relevant to the proposed research, applicants should be asked to provide literature to demonstrate that no sex, gender, or relevant diversity differences have been found.

Instructions for evaluators

Evaluators are crucial to the success of SG&DA policies. CIHR found that “targeting applicants alone to adopt new science policies without concomitant pressure by evaluators...may not be effective” (9). Since 2018, CIHR has required evaluators to rate the quality of the SG&DA as a “strength,” “weakness,” or “not applicable” and to provide a rationale for their rating along with recommendations to applicants for improvement.

Funders should provide applicants and evaluators similar forms and instructions for consistency across the research process. Some agencies, such as the EC, are limited in the overall instructions they can provide on this particular requirement given the number of topics that need to be covered. Agencies may provide “good research guides” that reference assessing SG&DA alongside other elements of peer review, such as ethics and reproducibility.

Agencies must monitor the evaluation

process to confirm that SG&DA is addressed in reviewer comments and that those comments are high quality.

Trainings for applicants, evaluators, and staff

SG&DA is not yet consistently part of university curricula in the physical and life sciences, health and biomedicine, and engineering. Until universities step up to the task, funding agencies need to fill this gap. The most comprehensive online, interactive agency trainings to date are in health and biomedicine. Trainings are also available for certain areas of the natural sciences, engineering, computer science, and environmental sciences, but more are needed. Agencies can evaluate the effectiveness of the training by including a pre-test and a post-test. CIHR, for example, found that the majority of participants who completed training modules demonstrated improved knowledge of sex and gender analysis (15). Most trainings are voluntary; however, some funders require applicants to submit a certificate of completion for large, strategic competitions.

Use of the same training materials by applicants, evaluators, and agency staff helps to ensure consistency in policies, terminology, and expectations. Agencies can coordinate and share trainings internationally; there is no need to duplicate efforts, except where specific cultural practices require a particular approach.

Some agencies foster training in this area through research institutions. The US NIH, for example, has invested \$160 million in

Specialized Centers of Research Excellence across 25 research institutions to “train researchers in experimental design and analyses that consider sex and/or gender” (1).

Evaluation of policy implementation

Only three agencies in our study had performed policy implementation evaluations. A further nine were in the planning stages; 10 had no plans in place. We strongly recommend that agencies implement evaluation plans as they develop policies to facilitate appropriate quantitative and qualitative evaluation.

Our framework sets out five aspects for the evaluation of policy implementation:

1) The number and proportion of proposals that include SG&DA. CIHR found that from 2011 to 2019, the proportion of grants including sex analysis increased from 22 to 83% and grants including gender analysis from 12 to 33%. The level of integration differed across sectors, with the lowest in biomedical and the highest in clinical research (8).

2) The number and proportion of proposals that include quality SG&DA. The EC conducted a mid-term evaluation of sex and gender analysis for Horizon 2020 in 2017 to consider the proposal quality, methods, impacts, dissemination, and also whether the project advanced methodology for sex and gender analysis in the particular field in which the proposal was submitted. Overall, the EC concluded that the quality of the gender dimension in project proposals was not high and that more training was needed (9). The EC also experimented with computer-assisted textual analysis given the volume of applications per year. These evaluation methods are in their infancy and require further development.

3) The quality of evaluators’ scoring and comments. This aspect emerged as a central point in the analysis of agencies’ self-reported activities. We did not score funders on this point but will in future iterations. CIHR manually sampled 5% of evaluators’ comments to check the quality of responses (9). The EC reviewed the effectiveness of review panels and found that only 36% considered the gender dimension and, of those, 70% included a gender expert, suggesting that review panels require guidance from experts (8).

4) The number of applicants, evaluators, and staff who engaged in trainings and in what type of training. If possible, the correlation between applicant training and the success of applicant proposals should be assessed. Some funders, such as the Spanish Carlos III Health Institute, reported in our questionnaire that they

monitor the number of applicants who participate in SG&DA training and are setting targets for improvement.

5) The number and proportion of peer-reviewed publications (or other recognized modes of dissemination) that result from funded proposals that incorporated SG&DA. To monitor this, funders will need to track papers and research outputs using grant numbers. The Science Foundation Ireland reported in our study that they collect researcher-reported publication data to check that proposals that included sex and/or gender analysis reported those results in publications.

A trend that we continue to watch is the broadening of sex and gender analysis to include other social dimensions. Already, the EC has added “intersectional” analysis to their gender dimension; these policies, however, remain under the broader gender equality strategy. The DFG started with sex, gender, and “diversity” on equal terms. The US NIH has included “age,” which they term “Inclusion Across the Lifespan.” A number of funding agencies, such as the Natural Sciences and Engineering Research Council of Canada, incorporate research design policies under a broader equity, diversity, and inclusion (EDI) umbrella. In the past, EDI has typically focused on “who” is doing the research, not on “how” research is done. This means that special care will be needed to expand EDI to include research methodologies.

FUTURE FRAMEWORK DEVELOPMENT

Our evaluation of SG&DA policy implementation has some limitations that restrict its generalizability. First, our coverage of public funding agencies, though geographically comprehensive, is not exhaustive. Participation rates were especially low for agencies in Africa and the Middle East, and we were unable to establish contacts with funding agencies in Russia and China. Second, we excluded private funders in this first round. Future studies might consider large international private funders, such as the Bill and Melinda Gates Foundation, Novo Nordisk Foundation, Oswaldo Cruz Foundation, and Wellcome Trust. Third, we might add elements to each of the five parts of our framework. For example, are applications that include SG&DA funded at rates and amounts comparable to those of other similar projects? Fourth, policies that encourage scientists to change their practices can spur resistance. Future frameworks might include assessing how funding agencies deal with such resistance. Finally, our study details how well and how consistently SG&DA policies are implemented, but further quantitative and

qualitative analysis is needed to correlate the framework’s scoring with the impact of agency policies on research outputs. Current funder evaluation policies do not support this type of analysis.

CONCLUSION

National public funding agencies seek different ways to support the progress of science and technology and advance health, prosperity, and well-being by promoting discovery and innovation, excellence, and societal relevance in research. SG&DA is one set of methods among many that researchers will deploy to help achieve these goals. Funding agencies are rapidly developing policies to promote SG&DA. The goal of our work is to assist in collaboratively achieving efficient implementations of SG&DA-related funding policies globally. ■

REFERENCES AND NOTES

1. M. E. Arnegard, L. A. Whitten, C. Hunter, J. A. Clayton, *J. Womens Health* **29**, 858 (2020).
2. C. Tannenbaum, R. P. Ellis, F. Eysse, J. Zou, L. Schiebinger, *Nature* **575**, 137 (2019).
3. K. E. Boerner *et al.*, *Pain* **159**, 2137 (2018).
4. M. R. Janevic *et al.*, *J. Pain* **23**, 707 (2022).
5. L. Schiebinger *et al.*, “Methods of sex, gender, and intersectional analysis,” *Gendered Innovations*; <http://genderedinnovations.stanford.edu/methods-sex-and-gender-analysis.html>.
6. L. Schiebinger *et al.*, “Sex and gender analysis policies of major granting agencies,” *Gendered Innovations*; http://genderedinnovations.stanford.edu/sex-and-gender-analysis-policies-major-granting-agencies_8_6_21.html.
7. R. Merriman *et al.*, *BMJ Glob. Health* **6**, e005672 (2021).
8. Directorate-General for Research and Innovation (European Commission), “Interim evaluation: gender equality as a crosscutting issue in Horizon 2020” (Publications Office of the European Union, Luxembourg, 2017).
9. J. Haverfield, C. Tannenbaum, *Health Res. Policy Syst.* **19**, 941 (2021).
10. J. C. Jackson, J. G. Payumo, A. J. Jamison, M. L. Conteh, P. Chirawu, *Front. Res. Metr. Anal.* **7**, 814600 (2022).
11. S. Y. Håkansson, J. Sand, “The gender dimension in research and innovation: Results from a global survey on research funding organisations” (Swedish Secretariat for Gender Research University of Gothenburg, Gothenburg, 2021).
12. GRC Gender Working Group, “Gender-disaggregated data at the participating organisations of the global research council: Results of a global survey” (Global Research Council, 2021).
13. J. W. Rich-Edwards, U. B. Kaiser, G. L. Chen, J. E. Manson, J. M. Goldstein, *Endocr. Rev.* **39**, 424 (2018).
14. L. Schiebinger *et al.*, “Mapping emerging policy practices of public granting agencies,” *Gendered Innovations*; <http://genderedinnovations.stanford.edu/sex-and-gender-analysis-policies-major-granting-agencies.html>.
15. C. Tannenbaum, K. van Hoof, *Biol. Sex Differ.* **9**, 1 (2018).

ACKNOWLEDGMENTS

We thank the members of our advisory board for their guidance and the participating funding agencies for their time.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abp9775

10.1126/science.abp9775

A framework for sex, gender, and diversity analysis in research

Lilian HuntMathias Wullum NielsenLonda Schiebinger

Science, 377 (6614), • DOI: 10.1126/science.abp9775

View the article online

<https://www.science.org/doi/10.1126/science.abp9775>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)