GRANteD
GRANT ALLOCATION DISPARITIES FROM A GENDER PERSPECTIVE

Grant Allocation Disparities from a Gender Perspective: Literature Review. Synthesis Report

Laura Cruz-Castro and Luis Sanz-Menéndez

CSIC Institute of Public Goods and Policies, Madrid

DELIVERABLE 1.1 (MAIN REPORT), NOVEMBER 2019

How to cite: Cruz-Castro, Laura and Luis Sanz-Menéndez (2019). ***Grant Allocation disparities from a Gender Perspective: Literature Review. Synthesis Report***. GRANteD Project D.1.1. DOI: http://dx.doi.org/10.20350/digitalCSIC/10548 and Permanent URI : http://hdl.handle.net/10261/200024

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Project no.	824574
Project acronym:	GRANteD
Project title:	Grant Allocation Disparities from a Gender Perspective
Instrument:	H2020-SwafS-2018-2020/H2020-SwafS-2018-1
Start of the project:	01.01.2019
Duration:	50 months
Work Package:	WP1
Due date of deliverable:	Month 8
Organisation name of lead contractor for this deliverable:	CSIC
Other partners involved:	JR, ORU, DZHW, TMC
Document version:	v02

The **methodology** of this report has consisted in documentary research and analytical reading and comparative synthesis of the secondary sources and references, mostly original research publications, but also books and relevant reports. In addition to the main report, an annex was produced that contains an annotated bibliography. This bibliography was developed gathering the descriptions that partners produced following a common template which was designed by leaders of WP1.

How to cite: Cruz-Castro, Laura and Luis Sanz-Menéndez (2019). ***Grant Allocation disparities from a Gender Perspective: Literature Review. Synthesis Report***. GRANteD Project D.1.1. DOI: http://dx.doi.org/10.20350/digitalCSIC/10548 and Permanent URI : http://hdl.handle.net/10261/200024

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Abstract

The purpose of this review is to provide the GRANteD project with robust analytical approaches and methodological insights that take into account the state of the art, but it also acknowledges and aims to overcome the main shortcomings and point out the gaps of the existing literature; it is also our contention that knowledge could be used to improve and refine, and also to strengthen, gender equality policies and the management of diversity in teams and research institutions. The substantive focus of the literature review refers to the central event of the "allocation of grants by means of peer review evaluation, observed as a process and an outcome". We also consider "the impact of grants in career advancement in the context of hiring and promotion decisions". However we can learn significantly from other areas in which gender research have made significant contributions, like occupation segregation, discrimination, stereotypes, etc.

Authors and contributors:

Name	Organisation	Contribution
Laura Cruz-Castro	*CSIC*	*Main author 1*
Luis Sanz-Menéndez	*CSIC*	*Main author 2*
Florian Holzinger	*JR*	*Written feedback*
Liisa Husu	*ORU*	*Written feedback*
Monica Gaughan	*ASU and SAB*	*Oral feedback*
Team members	*GRANteD*	*Oral feedback*

Dissemination level	
PU	Public

Consortium: The GRANteD consortium consists of five partners:

JOANNEUM RESEARCH Forschungsgesellschaft mbH (JR) (project leader, Austria),

OREBRO UNIVERSITY (ORU) (Sweden),

AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (CSIC) (Spain),

DEUTSCHES ZENTRUM FUR HOCHSCHUL UND WISSENSCHAFTSFORSCHUNG GMBH (DZHW) (Germany),

TERESA MOM CONSULTANCY BV (TMC) (Netherlands)

Keywords: research grants allocation, peer review, gender, inequalities, bias, research careers, women and science.

To address questions and comments please contact: Laura.Cruz@csic.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

TABLE OF CONTENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

1 INTRODUCTION, SUMMARY AND APPROACH

1. The issue of women in science and inequality in various dimensions gained salience at the end of the sixties in the U.S.. Several factors contributed to it: firstly, the expansion of the number of granted PhDs to women in previous years; secondly the limited number of women in the highest reputation positions in science, and thirdly the personal accounts of women being treated unfairly in the science system. Even if there were some previous empirical work from sociologists focused on the issue (Reskin 1976) or literature reviews (Zuckerman and Cole 1975) the book "Fair Science. Women in the scientific community" (J. R. Cole 1979) probably represented the first monograph addressing the issue from a general perspective.

2. The approach, as much of the sociology of science at the time, aimed to identify the existence of some systematic departure from the predominant universalism that characterises the science system as codified by Merton (Merton 1973) and the analytical perspective was the "social stratification of science" (J. R. Cole and Cole 1973) to understand if the mechanisms identified (e.g. processes of accumulation advantages and disadvantages) applied to the situation of women. Because answering the question of whether the differences could be the result of social or/and self-selection was not easy, Cole (J. R. Cole 1979) criticised claims of discrimination lacking measures of productivity, and the grounds of some affirmative action initiatives but, at the same time, he accepted the existence of some discrimination. Despite the service provided to attracting attention and research to the topic, he was criticised on the basis of methodological ambiguity or contradictions (White 1982) and for not being able to clearly state the null hypothesis.

3. Forty years later research in this domain is facing similar realities and the same challenges and methodological tensions identified in the past. Today **there are some indisputable facts:** women are a small portion of the faculty members of research universities especially in the STEM fields. Women have lower representation at the top of academic profession and in leadership positions in academic institutions their share is low to the relative number of women that qualify (European Commission 2019). Gender segregation in science has been documented (Caprile et al. 2012). And, in most countries and Research Funding Organisations (RFO) women are getting fewer research grants, or less than expected, they received lower scores and amounts of funding. All that, despite the fact that (in countries with more mature science systems) they get better grades in school and in general are more likely to graduate from universities.

4. **The purpose of this review** is to provide the GRANteD project with robust analytical approaches and methodological insights that take into account the state of the art, but it also acknowledges and aims to overcome the main shortcomings of the existing literature; it is also our contention that knowledge could be used to improve and refine, and also to

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

strengthen gender equality policies and the management of diversity in teams and research institutions.

5. The substantive **focus of the literature review** refers to the central issue of the "allocation of grants by means of peer review evaluation, observed as a process and an outcome. We also consider the impact of grants in career advancement in the context of hiring and promotion decisions. However we can learn significantly from other areas in which gender research has made significant contributions, like occupation segregation, discrimination, stereotypes, etc.

6. In this report we want to go beyond general discourses; we would also like to go further than average/mean values of inequalities and enter in the composition of fields, areas, disciplines or positions and go beyond most common assumptions (sometimes not well grounded) about gender and its relationship with research competitive funding and careers.

7. **The issue of whether gender bias or discrimination**[1] **occurs in grant allocation is rather complex**; firstly because sometimes bias is implicit, but mainly because it raises relevant methodological and measuring challenges. Indeed, the production of convincing evidence of discrimination or bias has been elusive, even using sophisticated regression methods dealing with endogeneity problems, mainly due to the difficulties of observational data to deal with causality.

8. This is why it is fair to say that previous research about bias has produced **contradictory evidence**; most of the diversity of findings could be related with the difficulty to control, in the research design, of "the similar characteristics and similar circumstances" of individuals and groups, and the features of processes and practices of allocating grants by RFO; unobserved heterogeneity in the research designs can be highly confusing and lead to problems of endogenity and internal validity.

9. Reporting evidence and labelling it is controversial. For example, in the U.S. a pathbreaking report of the National Academy of Sciences (2007) stated that it was not lack of talent, but **unintentional biases and outmoded institutional structures** that were hindering the access and advancement of women (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine 2007). However that evidence presented on women discrimination associated to social processes was challenged by some others, for example in some of the papers collected by Sommers (Hoff Summers 2009), they suggest that the origin of disparities could be not in bias in science, but that gender disparities were related to "characteristics of the gender preferences grounded in biological differences", an approach that has attracted much criticism itself. It is neither rare to find in the media, public opinion surveys, policy documents (Gender Action 2019) statements

[1] We provisionally define discrimination as a situation in which *members of a minority group (women) are treated differentially (less favourably) than members of a majority group with otherwise identical characteristics in similar circumstances*.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

mentioning that bias, discrimination and nepotism are the key forces driving patterns of inequality in science related to gender.

10. However, quite often, in support of this type of statements, there are mostly normative principles regarding equality/equity; the empirical evidence on *de facto* gender equality or inequality in a certain domain is another issue which requires empirically robust facts constructed and based either on multivariate analysis (which control for relevant intervening variables such us the capacity, merit or performance) or on experimental designs controlling for competing explanations. When we simply look at the distributions and success rates in getting grants we find, most of the times, differences by gender, but little can be said about which phenomena and its interpretation or what are the causal mechanisms behind, understood as those which bring about change in some variable.

11. But problems relate not only to the measure of disparitites and its quality, but also to the use of terminology, because sometimes the same wording is used to refer to different concepts, terms and facts. In this context the review needs to introduce some clarification about concepts and categories too.

1.1 Specific objectives

12. To address the research challenges of GRANteD we need:

a) **To determine whether there are disparities/inequalities or bias/discrimination or both**. As already mentioned, to discriminate is to treat two subjects with equal attributes and in equal circumstances in a different manner breaking the principle of merit. Differentiation is a precondition of discrimination. Additionally, it could also be the case that inequalities generated as a result of an action constitute a form of discrimination, but disparities could also be the result of differences in preferences.

b) **To revise the robustness of the findings of prior research**, since most of it **argues in favor of the existence of bias** (sometimes confounding the identification of the differences with the existence of discrimination), and to determine **if the evidence provided is incomplete** (e.g observational and without control of other intervening variables and with problems of endogeneity) **or non conclusive** (mixed results).

c) **To contextualise the** research findings in time and space; what could be robust evidence in one country or funding agency may not apply to a different country and context. The same applies to the time reference of the findings, what seemed indisputable facts 20 years ago could have changed radically more recently.

d) Even in cases of confirmed bias or discrimination, determining the underlying factors or causal processes is not automatic; **we need robust theories and adequate research designs** (preferably experimental).

e) To **analyse gender equality strategies** and their corresponding policies at RFO. This implies entering the arena of solutions and of the evaluation of very diverse policies and interventions to reduce differences in outcomes and improving the fairness of the processes. Policies, intervention and organisational practices have been analysed less.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Some of the **main areas of intervention** are related to: a) affirmative action or positive discrimination; b) requiring gender parity/gender balance or increase representation of women in committees; c) changes in evaluation criteria, metrics and scales; d) anonymisation of candidates' gender and/or evaluation of projects only; e) establishing quotas; f) raising awareness; etc. It is important to acknowledge that policies have effects and change reality so they should be a feature in the understanding of processes.

1.2 Clarifications of definitions and concepts

13. One of the problems we have found in the literature is certain confusion in the terminology and the construction of concepts. This is why we should try to establish a much more clear understanding of the various definitions. We can distinguish three levels of analysis:

    1. Disparities, Differences/Differentiation, Inequalities, Diversity, Gaps, etc. (Individual/Social level).
    2. Segmentation, Stratification, Segregation, etc. (Structural/Organisational level).
    3. Bias, Prejudice, Partiality, Discrimination, Unfairness, etc. (Individual/Organisational level).

14. In all levels of analysis we should not confuse the **outcomes** with the **processes** (S. Cole and Fiorentine 1991).

1.2.1 Disparities, differences/differentiation, inequalities, diversity, gaps.

15. **Disparities** refer to the existence of distinct characteristics (of a group of individuals) in some of their attributes, treatments, or **results** according to classes or types, which do not imply value judgements about "fairness", "efficiency" o "social welfare". In our case, **we use the term disparities mainly to highlight the fact that men outnumber women in grant allocations**.

16. **Differentiation** refers to the social processes that mark certain personal characteristics as important. It is a precondition of segregation. Although **social differentiation does not inevitably lead to unequal treatment of members of different categories**, differentiation is a necessary precursor for social stratification -systematic inequality in the distribution of socially valued resources- on the basis of some of the people's personal characteristics. Of the characteristics upon which societies and organisations differentiate and stratify their members, one of the most basic have been sex[2].

1.2.2 Segmentation, Segregation, Stratification.

17. **Segmentation** refers to the existence of systemic patterns (social structures) that condition results of processes and reproduce those structures (e.g. cumulative advantage or disadvantage). These structures (opportunity structures) contribute to the reproduction of

[2] In this report we use the term sex and gender indistinctively.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

inequalities, either due to structural effects, or by other means such the reproduction of power positions of the majority group (men or women) in the social structure.

18. **Segregation** (separation) is a fundamental process in social inequality. The characteristics on which groups are sorted could symbolise dominant or subordinate status and become the basis for differential treatment. Indeed, segregation facilitates unequal treatment by subjecting groups to different reward systems (opening the door to discrimination).

19. There is one type of segregation that is very important to the project. **Occupational segregation** (by professions, disciplines, jobs, institutions, organisations, etc.) is a form of segmentation in which there is a specialisation of some groups in some professions or fields, and that is based on attributes other than merit, competence, capacity or performance.

1.2.3 Bias, prejudice, partiality, discrimination, unfairness

20. **Discrimimation** refers to the different treatment of individuals and groups with identical or similar characteristics of circumstances. **Discrimination** occurs when we find that "members of a **minority group** (women, immigrants, blacks, etc.) are treated differentially (less favourably) than members of a majority group with otherwise identical characteristics in similar circumstances" (Bertrand and Duflo 2016).

21. **Bias** refers essentially to cognitive or evaluative processes where the judgement of the evaluator is wrong (and this might be due to reasons related to preferences, prejudices or incomplete information). **Bias is** "any feature of an evaluator's cognitive or attitudinal mind-set that could interfere with an objective evaluation" (Shatz 2004). Bias can be conscious or unconscious.

22. In the context of quantitative research on **bias in peer review**, reviewer's bias is understood as the **violation of impartiality in the evaluation** of a submission or application. Lee et al. (C. J. Lee et al. 2013) define **impartiality** in peer evaluations as the ability for any reviewer to interpret and apply evaluative criteria in the same way in the assessment of a submission. Ideally, impartiality ensures that evaluations are independent of the author's and reviewer's social identities and independent of the reviewer's theoretical biases and tolerance for risk. There are many reasons to challenge this ideal notion of impartiality in peer review [(Lamont 2009), (C. J. Lee et al. 2013)].

1.3 Measuring disparities, segregation, bias and discrimination

23. To properly measure discrimination we need to avoid the **confusion between outcome and process** (S. Cole and Fiorentine 1991). More precisely, these authors warn about **the error of using inequality itself, an outcome, as evidence of a process: discrimination or lack of equal opportunity**. The premise should be that discrimination is only one possible cause of any observed inequality between two or more different groups.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

24. Cole (J. R. Cole 1979) called **"naive residualism"** to the practice of attributing any difference in rewards received by men and women scientists to discrimination. Most scholars today are aware that simple zero-order correlations between sex and some outcome can not be called discrimination, and the control variable approach may be considered an example of a more "sophisticated resisualism". The logic behind is that if we have taken into account all of the other variables that may influence a particular reward, then any reward difference would have to be the result of the unmeasured variable. The most obvious problem with this approach is the impossibility to control for all or even the most important variables that can cause inequality, most importantly, self-selectivity, which leaves us with the question of whether the control variable approach is completely valid.

25. For Cole and Fiorentine (S. Cole and Fiorentine 1991) the answer **depended on the nature of the dependent variable**. Some dependent variables are clearly more influenced by selectivity than others. If the particular reward under study is not or little influenced by self-selection, it can be studied by "**sophisticated residualism**", if adequate controls of performance are included among the controls. But it is important to note that rulling out self-selection is easier in cases where individuals do not have to apply, which is not the case of research grants.

26. This was more or less the state of the sociological reasoning about the topic of discrimination some thirty years ago, but in the last years we observe that many research works have forgotten these foundations as we will examine in section 2.1.1. As we will see, the literature and the methods have evolved substantially, but the critique to studies that use data on differences to conclude about the processes of creating inequality is still valid.

27. The same caveat applies when we observe that members of certain groups are less represented or less successful in specific job types (e.g women in academia) or reward processes (e.g grants) and we implicitly conclude that this must be the result of valid differences in individual merit and achievements. Van der Lee and Ellemers (R. van der Lee and Ellemers 2018) refer to this as "individual merit ideology" according to which if women are less successful than men something must be lacking in the competences, efforts or priorities of these women.

28. As regards the measurement of segregation Gross (Gross 1968) was the first to introduce the concept of "sex segregation" to describe women's and men's concentration in different occupations. Since then, researchers have shown that more often than not, **men and women are concentrated in different organisations or hold different jobs within the same organisations;** this is also the case in some research domains where we can find a gender-based specialisation. The question is: what factors are associated with the presence and decline of sex segregation in the workplace?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

29. In Economics there is a l**arge body of research on occupational segregation** that focuses on two main approaches to the problem: supply and demand. From these perspectives segregation stems in part a) from the **difference in preferences, skills, experience** and past performance that the sexes bring to the labour market; but b) it also results from organisations or **employers' preferences and practices.** Because the previous research has generally ignore it, we should acknowledge it and try to take the issues of segregation, discrimination and bias out of the exclusive level of individuals (researchers, evaluators) and to incorporate the organisational level into the analysis.

30. We need to recall that the three most commonly investigated indicators to empirically account for **gender inequality in organisations** have been:

    a) the **gender gap in wages** [(Goldin 1990) , (Blau and Kahn 2000) , (Blau and Kahn 2017)]
    b) the **gender gap in management and leadership positions** [(A. H. Eagly and Karau 1991), (A. Eagly and Carli 2007), (Kalev, Kelly, and Dobbin 2006)], and the closely related
    c) **Sex segregation of occupations** [(Bielby and Baron 1986) , (Reskin et al. 1990) , (Roos and Reskin 1992) , (Martell, Emrich, and Robison-Cox 2012)].

31. Research on **gender disparities in grant allocations is less common but it is relevant both for theory and for policy for several reasons; firsly, it** speaks directly to the unexplained **gender gap in career advancemen**t by illuminating potential effects of gender on productivity, reputation and compensation; secondly, research on gender in grant allocations also offers potential explanations for the distribution of **other types of organisational resources and career opportunities** (e.g., budgets, training, developmental work experiences), which contributes to the gender gap in management and leadership positions. Finally, research on this topic offers valuable insights for researchers interested in the **mechanisms of gender inequality in organisations** (see Reskin's (2003) call for research on the mechanisms of ascriptive inequality).

32. As regards the measures, when analysing the allocation of grants (Bornmann, Mutz, and Daniel 2008) not only the potential but also the scientific performance (track record) of the applicants should be measured in the analyses. In this way a distinction can be made between the influence of the applicants' achievements up to the date of application and the potential sources of bias in the selection decisions.

1.4 The importance of mechanisms: Causality and explanatory theories

33. The three analytical levels in which the concepts presented are located involve different degrees of presence of normative values of equity, justice and efficiency. In all three levels, the main problems for the analysis are the identification of the causal mechanisms that produce the phenomena of interest and the integration of different explanatory theories. Establishing that inequalities are related to or are the result of segregated

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

structures, discrimination practices or biased evaluations requires addressing the underlying mechanisms and moving into the sphere of available theories.

34. In general, we can provisionally say that existing disparities could be the result of discrimination if they do not originate from individual choices. In the same way, outcome differences could be related to bias if the assessment or evaluation is wrong, e.g. it is not reliable, fair, or lacks predictive validity. In other words, inequality or disparities are outcomes. When those outcomes do not result from self-selection we could suspect the existence of "unfair" mechanisms or processes: segmentation (structural), discrimination (behavioral/practices), or bias (cognitive evaluation). Therefore, the analysis needs first to **confirm the existence of the phenomenon** and then explain it; for that we need theory and a proper research design because if the explanation is incorrect the policies and interventions will be poorly founded, and even have unintended effects detrimental to the policy goal.

35. Gender is one among a possible set of variables that could contribute to the explanation of the outcomes in terms of allocations. In modelling the effect of gender, the variable should be compared with other potential ascritive factors.

36. Clarifying the **sources of gender disparities** (or any other of the relevant events such as segregation or discrimination) remains a high priority since the effectiveness of organisational policies seeking to reduce inequality depends on an accurate understanding of the organisational mechanisms that produce inequality [(Bielby 2000), (Reskin 2000a), (Reskin 2000b)]. This is why, in this review, we seek to advance our knowledge of **how social and organisational processes affect the allocation of research grants and contribute to gender inequalities later in the career.**

37. There is some empirical information about the **differences in grant application** and **success rates** between male and female researchers; additionally, it is possible to find scholarly work and official reports where some data on variables that seem relevant and associated to gender differences has been gathered; however, the number of explanatory theories or causal analysis addressing the existence of discrimination or bias is much more limited; in this sense, as already mentioned, much of the previous work is either data-driven or normatively-guided.

38. Instead, we think it can be useful to recall from economics and sociology some insights about the most abstract approaches to gender disparities, segmentation and discrimination. In Economics (Bertrand and Duflo 2016) two main approaches could contribute to link the existence of disparities with **discrimination**. The first is "**taste-based discrimination**" theory (Becker 1957) based on the discriminatory preferences of employers, fellow colleagues or custumers, meaning that employers have preferences over the "expected productivity" or performance of different groups. The second one is the "**statistical discrimination**" [(Phelps 1972), (Arrow 1973)] that refers to behaviour under uncertainty and limited information; in this case employers or evaluators use of information that

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

characterises the group of reference, for example, on the average performance of reference groups (e.g. women), that could be correct or biased; in this case, when the individual information available is incomplete, group specific membership could provide the employer (or evaluator) with information (but not necessarily correct) on the expected productivity of the individual. While taste-based discrimination is inefficient and unfair, statistical discrimination could be considered as socially efficient, but unfair.

39. Implicitly, employers, colleagues or evaluators often use "***group membership"*** (in our case of interest is gender) in their decisions, but at the same time individuals may have different group memberships (or identities) and some of them could compete among themselves (e.g. university of PhD award, advisor, gender, etc.) and they could have contradictory effects.

40. Despite the value of taking preferences into account, sociologists like Barbara Reskin (Reskin 2000a) warn us about the difficulties of explaining gender disparitities (and other kind of ascriptive inequalities) based on allocators' motives. In her opinion, this approach had been mostly inconclusive because **motive-based theories can not be empirically tested**. Without entering into debate on whether or not this claim is correct and whether motives or preferences can be observed and measured, her proposal to analyse mechanisms instead of motives is worth noting. Focusing on mechanisms implies a **shift from asking "why" questions to asking "how" or "under what conditions" questions**.

41. Mechanisms can be interpersonal, social and organisational. **Interpersonal mechanisms** can affect the amount of gender inequality in a given organisational context converting allocators' preferences or motives into differential behaviour towards men and women. **Social mechanisms** include normative expectations, social stereotypes, laws and regulations; finally, **organisational mechanisms** are several and include recruitment practices (Reskin and McBrier 2000), formalisation of procedures, level of managerial discretion, visibility or blindness of ascriptive characteristics (e.g. gender) to allocators, transparency of evaluation, accountability of allocators, and publicity of outcomes, among others.

42. In principle, bias can take place at any of the three mentioned levels of mechanisms. According to Lee et al. (C. J. Lee et al. 2013) quantitative research can be categorised by differences in their conception of the **primary source of bias**: (a) error in assessing a submission's "true quality," (b) social characteristics of the author, (c) social characteristics of the reviewer, and/or (d) content of the submission. Social bias is the differential evaluation of an author's submission as a result of her/his perceived membership in a particular social category. Social bias challenges the thesis of impartiality by suggesting that reviewers do not evaluate submissions—their content and relationship to the literature—independently of the author's (perceived) identity. A problem with a substantial part of the literature in this area is that it assumes that the quality of work by individuals across different social groups (e.g., prestigious vs. not, men vs. women) is, in the aggregate, roughly comparable. This might be the case, but it should not be an

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

assumption, but rather an empirical issue. When the assumption is made, the usual claim is that success rates of less powerful social groups should be proportionate to their representation in submission rates. Researchers infer the existence of bias when a difference is found and infer the lack of bias when no difference is found. In our view, this approach is incorrect (S. Cole and Fiorentine 1991) .

43. Fewer studies are able to demonstrate that their submission pools are similar to or representative of the larger population of researchers; the ones which do so, usually control for factors that may be correlated with quality. This is a very important issue that deserves attention in our project [(Bornmann 2011),(C. J. Lee et al. 2013)]. However, **after controlling for quality, the persistence of unexplained gender gaps suggests, but it does not automatically prove** the existence of discrimination; similarly, from general trends towards closing gaps we should not infer a decrease in discrimination.

44. In sum, to explain variation in levels of gender disparities across settings, and over time, we should analyse data at the individual level but also at the organisational level, and this data should include allocation mechanisms.

1.5 Methods and research designs (and consequences for the robustness of “explanations”)

1.5.1 Observational versus Experimental approaches

45. From the methodological point of view there are two main approaches in the literature: a) observational and b) experimental. Bearing this distinction in mind, we should address the need to critically question the assumptions and simplifications that sometimes affect gender equality issues.

46. Observational approaches usually address the extent of disparities and the existence of discrimination using different forms of multivariate or regression analysis, sometimes even with not small data sets. The exercise is to compare otherwise identical men and women and estimate the average outcome differences. Regressions are run with a particular outcome, such as the allocation of grants or promotions, as the dependent variables and a set of independent variables that most of the time includes only the personal characteristics of the applicants, although sometimes, contextual characteristics are also accounted for.

47. Estimating discrimination or bias in this way is problematic. Firstly, it is quite often impossible to account for all the factors affecting the outcome. In the case of allocation of grants for example, we want to compare two equally productive individuals, but productivity may have mutifaceted dimensions, not observable or difficult to measure; additionally the criteria defined for the assessment could include include merit and worth dimensions. Secondly, if some of the inequity in allocations (or promotions) is itself a result of previous inequalities in, for example, education or reputational status, and those inequalities are themselves a result of discrimination or segmentation, residual estimates of

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

discrimination after differential education or reputation is accounted for probably underestimate the extent of discrimination.

48. While descriptive accounts of the grant applications and allocations by gender show differences, the main issue is how to develop a research strategy that could contribute to the confirmation (or not) of the existence of discrimination processes and biased outcomes. According to Marsh et al. (Marsh, Jayasinghe, and Bond 2008) most peer review research is correlational and provides weak bases for causal inferences particularly regarding bias. As highlithed by Bornmann and Daniel (Bornmann and Daniel 2005) the lack of experimentally derived findindgs makes it impossible to determine without doubt whether work from a particular group of researchers receives better reviews (and thus has a higher approval rate) due to biases in the review and decision-making process, or if favourable review and greater success in the selection procedure is just a consequence of the scientific merit of the corresponding group of applicants.

49. The main advantage of the various types of "**experimental approaches**" (Gerber and Green 2011) in comparison with the descriptive or observational ones (even with a lot of controls in the context of the econometric evidence), is that they allow for higher levels of internal validity, or demonstration of causality. It is important not to equate experimental design only with laboratory studies, which are often criticised as being not applicable to real world organisations, since decision making contexts in a laboratory setting are abstracted form the institutional environment of allocation decisions (Bielby 2000).

50. Experiments also have limitations. To have external validity, experimental methods have to meet certain sampling and generalisability requirements. The important differences in the design of laboratory experiments and the selection of participants raised the issue of to what extent differences in findings and lack of generalisability are not the result of the variation in the substantive methodological approaches. We believe that research in the area of gender bias is likely to benefit from field experiments (D. C. Mutz 2011) which provide greater external variation in the organisational environment.

51. This is why one of the key dimensions of the literature review is the distinction between the methodological approach **(observational versus experimental**) of previous work. Although experimental studies are better suited to confirm the existence of discrimination or gender bias, determining their causes is not automatic, it is a further step. These issues will be addressed in more detail in section 2.3.

1.5.2 Static and dynamic approaches

52. There is also a second dimension of the analysis that refers to the issue of change over time, because most of the evidence presented in the literature is constructed in **static settings.** A rich literature has documented the existence of differences between men and women in science [(Ceci et al. 2014), (Kahn and Ginther 2018)] and has addressed the problem of the factors that could contribute to explain them. Additionally, some research has documented the existence of discrimination in a wide range of contexts; however,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

these empirical studies have mostly focused on static settings where individuals are evaluated based on the quality of their output in the context of a single interaction, with no information on prior evaluations in similar contexts. Moreover, from such static settings it is difficult to identify the underlying source of discrimination, as different sources may generate the same patterns of observable behaviour.

53. Therefore, it is important to consider timing and the need to introduce the idea of sequence or process (Abbott 2016) in order to account for the changes and adjustments resulting from transformation of context (and learning) including the effects of policies and interventions, and the effects of interactions among groups or the learning mechanisms of actors.

54. In particular settings, **evaluators are faced with a lack of information** about the candidates or with no records of reputation (or complete information of past performance). These circumstances lead to difficulties to evaluate the expected performance of an individual correctly; therefore, one needs to take into account the effects of successive interaction processes between evaluators and candidates, the learning processes, as well as the operation of identitary elements of a collective nature (gender, field, institutional affiliation).

55. The dynamic dimension is especially important in settings where the subjects of evaluation have diverse histories and records of previous evaluations. Those lacking such history could be more likely to be evaluated on the basis of uninformed stereotypes (Bohren, Imas, and Rosenberg 2019). If there was learning over time and as a product of interactions, "reverse discrimination" could take place. The occurrence of reverse discrimination can yield information about the sources of discrimination, because, theoretically, learning is possible in cases of belief-based discrimination or bias, but more difficult in cases where discrimination is based on preferences.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2 GRANT ALLOCATION GENDER DISPARITIES AS OUTCOMES AND PROCESSES: THE LOGIC MODEL

56. This is the core section of this synthesis report and it is organised from the point of view of the process (its phases) and the factors that influence (or could influence) the outcomes of our main event of interest: the allocation of grants by the Research Funding Organisations (RFO). The main focus of the report relates to the allocation of grants (in a context of evaluation or peer review); the allocation is just one stage of the process. The focus of several analyses is on the relationship **between the proportion of applicants and the success rates, or on the relationship between funding and ex-ante productivity to determine the fairness,** but there are factors and antecedent conditions that should be taken into account to understand the outcomes, factors that shape and influence the outcome and its gender distribution.

57. Understanding the outcomes of the allocation of grants requires considering two differentiated steps: the first one is the application and second is the evaluation/allocation. As we have seen in section 1, there is self selection. The probability of geting the grant is conditional on the probability of applying for funding, and the decision to apply is determined by antecedent factors that we also need to consider in the model.

58. In a second phase the project needs to analyse the impact of grants in career advancement; an analysis that requires considering not only granted and not granted researchers but also to determine what are the factors that influence (among of them grants) career advancement. A further complexity is introduced by the fact that, methodologically, to really talk about **career impact**, it would be important to **compare subgroups of successful applicants not only with rejected ones but also with subgroups of non-applicants** (potential ones in terms of eligibility) (Neufeld and Hornbostel 2012). A longitudinal approach is required to really fulfil this task and its complexity raises issues of feasibility.

59. At the same time we need to go deeper into the differences (by area, disciplines, categories, etc.), and to go **beyond the use of mean values** that tend to hide the change in the ways segmentation operates. We should advance into a fine grain approach with domain or field disaggregation.

60. In this section we will analyse some antecedents of the differences between men and women that are usually presented as related with the emergence of the further disparities and bias related to funding and employment. Accounting for the structure of prior differences is essential for understanding subsequent processes of inequality in research funding and career advancement. In the most simple understanding past differences in (t_{-1}) could be at the origin of inequalities in (t_0). We devote the first part of the section to the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

literature about the previous structure of inequalities that could influence and shape further processes related to application and allocation of grants. In the rest we revise some literature in the specific topic of grants' applications and evaluation, and we present some evidence more related to the evaluation of researchers and academics in the context of career advancement, always with a focus on gender. We end with a section on policies and interventions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.1 The effect of previous differences in the structure of opportunities and behaviour

61. Our starting point is that differences in grant application and success could be shapped and conditioned by previous inequalities. The literature on this topic is large. Differences in the prevalence of men and women by fields of science, particularly the underrepresentation of women in math-intensive fields of science and engineering have been extensively documented (Ceci et al. 2014). It is also a well known fact that there has been a secular decline in the overrepresentation of men both in participation and in the upper tail of achievement distributions in some countries. However, despite growing equality in high school and secondary education, **wide gaps in STEM participation remain in higher education,** especially among degree recipients in engineering, the physical sciences, math and computer science [(Xie, Fang, and Shauman 2015), (Ministerio Ciencia e Innovacion 2011)].

2.1.1 From school to research careers: the pipeline model (a critical view) and the glass celling

62. Using longitudinal data from 1960-1990, Xie and Shauman (Xie and Shauman 2003) developed a life course approach departing from the common "science pipeline" approach which they consider inadequate. They argue that "it would be naive to presume that S&E occupations are closed to women simply through discriminatory practice and structural barriers" (p.2). "The inadequate supply of interested and qualified women has been as much, if not more, of a hindrance to the feminisation of science as the influence of the demand factors". They stress the conceptual and methodological limitations of much of the previous literature that restrict the significance and usefulness of its findings.

63. According to these authors the "pipeline" model" is conceived as a developmental process where explanation is related to the rate of attrition and the policy action is focused on blocking the leakage; the problem, in their view, is that the model restricts the questions that are addressed by researchers. A summary of the criticisms posed to the model includes the following: firstly, it does not capture the complexity of the educational and career outcomes; secondly, the model is developmental, trajectory is assumed to follow progress, otherwise it is considered a failure; thirdly, the pipeline is independent of the timing and character of other life course events. Forthly, on the methodological side, the model is often substantiated on single data sources, non representative samples, with problems of "left censoring" and "selectivity bias".

64. In sum, dynamics can not be analysed by crosssectional data. The alternative of life course analysis brings up the ideas of sequence and interdependence of events. The very idea of inequality in the labour force is the manifestation of a process of stratification that occurs through-out the life course leading to well known processes of cumulative advantage. (Merton 1973). Cohort analysis and longitudinal studies are needed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.1.1.1 Gender differences in STEM

65. Historically, a large proportion of women who were interested in science and engineering careers were lost at every educational transition, and this is still the case in many systems. Yazilitas et al. (Yazilitas et al. 2013) categorise the state of the art on this topic (main factors) in three categories: micro level focus (self-efficacy beliefs, subjective value expectancy, role models) macro-level focus (socioeconomic conditions, socialisation) and institutional focus (education system features). Early literature often attributed gender gaps in tertiary STEM education to **gender differences in pre-college math and science achievement and participation**. However, more recent studies show that better achievement in math and science of male students are attributable to differences in the attitudes towards those disciplines rather than to **capacity**. As the gaps are closing capacity-based explanations have lost power. Contrary to essentialist views, social-psychological perspectives and socio-cultural perspectives provide more nuanced explanations of the interest gap and highlight the importance of **stereotypes**, as well as parents, teachers and significant others' expectations. In particular, expectation-states theory (Ridgeway 2014) argues that cultural **stereotypes structure inequality by generating implicit bias in evaluation and association preferences that segregate networks**. While the biasing impact of gender beliefs may be small in any one instance, the consequences acumulate over individuals' lives and result in substantially different outcomes for men and women (Ridgeway and Correll 2004).

66. In the same vein some economists (Kahn and Ginther 2018) concentrate on the **environmental factors that influence ability, preferences and rewards** for choosing a career in STEM, and favour **psychological and preferences explanations,** including the effect of stereotypes, role models, competition, risk aversion and interests. Because of the observed increase in the access to science by younger cohorts of women, it could be argued that it is simply a matter of time to achieve gender equity. However, testing the hypothesis that vertical gender segregation in scientific and academic careers is transient **requires longitudinal studies** of men and women who have entered academia and/or research organisations during a given point of time (Palomba 2004); if gender inequality persists, when seniority and other factors are held constant, then the "it is a matter of time" explanation cannot be considered valid.

67. Therefore it is important to acknowledge that women are not equally present across fields of science and to the extent that RFOs operate on the basis of field-specific programmes and calls, this has effects on gender disparities in grant allocations. But the problem is not simply the pipeline; in several fields the pipeline has reached gender parity and even in some areas women are the majority, as reflected in the proportion of female doctorates; yet at the top of research institutions percentages drop dramatically and do not mirror the closing gaps in PhDs granting.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.1.1.2 Occupational segmentation: Lower presence in high level positions and top rank institutions

68. Turning to a later phase in the career, there is also evidence of **occupational segmentation** and underrepresentation of women in most **high-level positions in organisations** [(Gino, Wilmuth, and Brooks 2015a), (Gupta et al. 2004)]. Men also outnumber women in positions of formal power, authority and high income [(Xie and Shauman 2003), (Timmers, Willemsen, and Tijdens 2010)]. Likewise, the underrepresentation of women at the top of the academy, the so-called the "glass celling" (Rosser 2004), is a persistent issue, but most of the measurements concerning the scarcity of women at the top of the academic and scientific hierarchy are cross-sectional and not longitudinal.

69. As explained, segregation is the difference in two groups' percentage distributions across some set of categories, for instance, organisations, occupations, ranks, fields, etc. Sociologists have stressed **sex segregation** because it is a central mechanism **linking sex to unequal individual career rewards** (productivity, promotion and pay). And the economic literature has also highlighted that gender differences in location in the labour market continue to be important for explaining pay gaps (more than human capital factors) and that some developments in the labour market for highly skilled workers in recent decades, such as wage penalties for temporal flexibility, have favoured men (Blau and Kahn 2017).

2.1.1.3 Some empirical facts

70. It is not new to acknowledge that women are likely to find obstacles along their research careers, and these barriers have differential impact by field and career stage. At the international level, a report (National Research Council 2010) for the U.S. analysed the transitions from assistant to associate professors and from associate to full professors in science and engineering departments; the survey revealed some interesting facts: in all fields, **women were underrepresented among tenure candidates** with respect to the number of women in the department, and especially in those subfields with a higher number of female professors: the probability that a woman got tenure increased when the proportion of tenured female faculty in the department was low; field or size of department did not have an effect; taken together all departments and fields 90% of men and 88% of women who a**pplied for a full professorship got it eventually, with no statistically significant differences.** In Europe, Ooms et al. (Ooms, Werker, and Hopp 2019) in a small scale study at two European universities of technology found an advantage of male academics in obtaining full professorships that stemed from advantages in earlier transitions, but they also reported that men had a much higher publication record and innovation output. The less favourable conditions of women in the academic labour market have also been reported in national case studies (Waaijer et al. 2016).

71. It may be the case that women who stay in academic research get tenure eventually, but many data sources indicate that **women are more likely than men to exit science before getting tenure**. As regards causes, Goulden et al. (Goulden, Mason, and Frasch 2011)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

found that, in the U.S., family formation accounted for the largest drop outs from graduate school to the acquisition of tenure for women in the sciences; similar results have been reported by Cech and Blair-Loy (Cech and Blair-Loy 2019) who, using a nationally representative 8-year longitudinal sample of U.S. STEM professionals, found **substantial attrition of new mothers**: 43% of women left full-time STEM employment after their first child, while the figure was 23% for men. Many studies in the U.S. find the majority of the gender retention gap is due to women leaving the labour force entirely, and that this exit is highly correlated with child bearing (Kahn and Ginther 2015).

72. The **distribution of women across academic organisations** with more or less reputation and resources is segregated. Women tend to work at **universities with lower reputation** and focus more on teaching (elsevier 2017). Sheltzer and Smith (Sheltzer and Smith 2014) observed that one cause of the leaky pipeline in biomedical research in the U.S. may be the exclusion of women, or their self-selected absence, from certain prominent or elite laboratories; elite male faculty members tended to employ fewer female graduates and postdoctoral researchers than female faculty members did.

73. Although the literature is dominated by case studies of the U.S., there is some research in European universities with results along the same lines (Caprile et al. 2012); for instance, Conti and Visentin (Annamaria Conti and Visentin 2015) analysing PhDs in Science and Engineering from two major universities in Sweden and Switzerland, found that **women are less likely than men to be employed in highly ranked universities**, even after controlling for their research outputs, but they also reported that gender differences in Ph.D.s' appointment to professorship were explained by the Ph.D.s' publication output and the quality of their postdoctoral training.

74. A caveat about comparative arguments is pertinent here; women's career trajectories cannot be compared across countries without taking the structural national features of the labour market into account. As pointed out by Le Feuvre (Le Feuvre 2018), in most national contexts, the academic profession is becoming increasingly fragmented and internally segmented. Although there is clearly a gender dimension to this re-segregation, it is not homogenous across countries.

2.1.1.4 The mechanisms

75. In this review we take the approach of a logic model for addressing gender disparities in grant funding that includes the prior or current gender composition of organisations and fields. The **determinants of the gender composition of organisations** include: the composition of the labour supply (in this case the scientists), the preferences of employers (including the qualifications they require), the response of majority groups within organisations, the organisation's attractiveness, organisational size, slack resources (Tolbert and Oberfield 1991) and recruitment methods (Reskin and McBrier 2000). In turn, gender **composition affects organisations themselves**, including their performance, hiring

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

and promotion practices and gaps, earning gaps, stereotyping and evaluation [(Reskin, McBrier, and Kmec 1999), (Reskin and Bielby 2005)].

Collective behaviour and queuing

76. Regarding the **mechanisms of segregation in occupations**, Reskin et al. (Reskin et al. 1990) proposed and developed a queuing theory of occupations' sex composition, arguing that the labour market comprises a "gender queue" with employers preferring male to female workers in most jobs; workers also rank jobs into a job queue. As a result, the highest ranked workers monopolise the most desirable jobs. The authors use this perspective to explain why several male occupations opened their doors to female participation in the seventies and present case studies of feminisation of specific occupations.

77. Some authors (Martell, Emrich, and Robison-Cox 2012) have tried to **link gender segregation in organisations (a macro-level phenomenon) to gender bias in performance assessment (a micro-level phenomenon) and have proposed** that: a) gender segregation in organisations arises from the collective behaviour of individuals who express a small bias in favour of men, in concert with the signals governing promotion decisions; b) the emergence of a gender-segregated organisation is often unintentional and bottom-up and top-down processes that produce segregation are difficult to observe; and c) agent-based modelling is well-suited for addressing the dynamics of bias that produce gender segregated organisations.

Organisational structures

78. Smith-Doer (Smith-Doerr 2004) argued that networks and hierarchies, as organisational forms, provide different employment experiences for female scientists; analysing the careers of over 2000 U.S. life scientists in the biotechnology **industry,** the study concluded that firms governed by networks allow for greater gender equity. The importance of the **institutional context** was also reported by Corley and Gaughan (Corley and Gaughan 2005) comparing U.S. university research centres with departments: although, as expected, women were younger, less likely to be tenured, and at a lower rank than their male colleagues, they appear to have greater research equality in research centers (compared to the departmental setting). In particular, men and women in research centers spent the same amount of time writing grant proposals, conducting both grant-supported and unfunded research, and administering grants.

Self assessment and perception

79. From a more social psychology point of view [(Correll 2001), (Correll 2004)] highlights the role of biased self-assessment and gender status beliefs as mechanisms behind segmentation. Cultural beliefs about gender are argued to **bias individuals' perceptions of their competence** at various career-relevant tasks, controlling for actual ability. To the extent that individuals then act on gender-differentiated perceptions when making career

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

decisions, cultural beliefs about gender channel men and women in substantially different career directions. She evaluates these hypotheses considering how gendered beliefs about mathematics impact individuals' assessments of their own mathematical competence, which in turn lead to gender differences in decisions to persist on a path towards a career in science, mathematics or engeeniering.

Stereotyping

80. Meyer et al. (Meyer, Cimpian, and Leslie 2015) proposed the idea of **the FAB** (field specific ability belief) whereby women are likely to be underrepresented in fields thought to require raw intellectual talent or brilliance-a sort of talent that **women are stereotyped to posses** less than men. Leslie and colleagues [(Leslie et al. 2015) (Cimpian and Leslie 2015)] tested this hypothesis in a U.S. nationwide survey of academics over three competing ones: The first concerned possible gender differences in willingness to work long hours: the more demanding a discipline in terms of work hours, the fewer the women. The second competing hypothesis was related to possible gender differences at the high end of the aptitude distribution which might cause greater gender gaps in highly selective fields: the more selective a discipline, the fewer the women. The third competing hypothesis was that the more a discipline prioritises systemising and abstraction over empathising, the fewer the women. According to their findings, the FAB hypothesis, unlike these three competitors, was able to predict women's representation across all of academia. Although inspiring, one has to acknowledge, however, that their study was survey-based, and cross-sectional, and thus limited as regards causal explanations.

2.1.2 Performance and productivity

81. The structure of prior differences or inequalities also includes performance, as a form of ability or competence measurement. Prior performance is essential when analysising gender disparities in grant allocations, especially considering that the idea of merit represents the universalistic normative principle that legitimates the allocation of resources.

2.1.2.1 Gender differences in productivity

82. The literature on **publications and productivity** and impact differences by gender is substantial although most of the evidence refers to the U.S.; classical sociology of science and blibliometric research have extensively addressed the issue. The traditional empirical claim has been that men on average publish more papers and receive more citations than female scientists [(J. Cole and Zuckerman 1984), (Long 1992), (Xie and Shauman 1998), (Nakhaie 2002), (Taylor, Fender, and Burke 2006), (Abramo, D'Angelo, and Caprasecca 2009), (Symonds et al. 2006), (van den Besselaar and Sandström 2016) among others]. A trend towards closing the gap has also been reported in the literature (Xie and Shauman 1998) as well as the finding that productivity of both men and women increases with scientific rank (Mauleón, Bordons, and Oppenheim 2008). More recent research reports

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

that, in highly selective research institutions, the direct relationship between gender and publications is relatively small for PhD students compared to faculty (Pezzoni et al. 2016).

83. In a recent meta-analysis of the findings of over a hundred papers Astegiano et al. (Astegiano, Sebastián-González, and Castanho 2019) still conclude that men published more articles per capita and had more scientific outputs (articles, grants, research positions) as a group, and the difference is not smaller in the 21th century than in the 20th century.This finding is mainly based on group comparisons, that do not account for overrepresentation of men in science. No differences in success rates in accepted papers were found- suggesting that women simply submit fewer papers; interestingly, men did better when peer committees played a role (grants, jobs, committee member). Finally, men were more cited, but without self-citations this effect disappeared.

84. The gender productivity gap is more acute at the top of the academic hierarchy. Up to date evidence shows that there is a considerable gender productivity gap among "stars" in favour of men across fields. Specifically, the underrepresentation of women is more extreme as we consider more elite ranges of performance (i.e., top 10%, 5%, and 1% of performers (Aguinis, Ji, and Joo 2018) . Although in all fields authorship moves towards gender parity, in quite a few STEM fields there is still a long way to go, and the change is very slow (Holman, Stuart-Fox, and Hauser 2018). The main problem with most of the studies of the relationship between gender and productivity is related to the establishment of causality, since most of the findings and conclusions are conditional about what may cause the observed effects.

85. Recent research confirms that women authors have been persistently underrepresented in high-profile journals. Analysing research article publication records from 15 high-profile multidisciplinary and neuroscience journals for 2005-2017, Shen et al. (Y. A. Shen et al. 2018) studied the representation of women over time, as well as its relationship with journal impact factor. They found that the percent of women first and last authors was negatively associated with journal impact factor. Since publishing in high-profile journals is a gateway to academic success this underrepresentation of women may contribute to the lack of women at the top of the academic ladder.

2.1.2.2 Why women appear to be less productive and advance more slowly?

86. Let us start by acknowledging that productivity is not only an independent characteristic of individuals but also a reflection of their positions in the academic social structure and the access to resources that those positions provide. Analytically, publication productivity is both cause and effect of status in science. Having said this, the explanations for women lower productivity can be classified broadly in two categories associated with what Sonnert and Holton (Sonnert and Holton 1995) called the "difference and the deficit" models**;** the **difference model** states that women act differently because they are different in motivations and commitment to scientific career; these differences could be innate or social. Most sociology of science has disposed of the idea of innate differences and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

stresses the importance of cultural and social influences of women to make certain educational and career choices. In contrast, **the deficit model** states that women have the same goals and aspirations than men but that they are treated differently. Therefore, their lower performance is due to lower opportunities, obstacles in their career, more problems in raising funds and in collaborating with other scientists. However, both type of factors are compatible and may reinforce each other (Mairesse and Pezzoni 2015).

87. Within those two general approaches, several **explanations of the productivity gap have been proposed**: scientific ability, self-selection, social selection, and accumulated disadvantage (Zuckerman 2001). The first explanation has been losing ground. The second explanation is often related to the evidence that **family affects women more than men,** leaving them with less time for research and therefore lower scientific output especially in the early career [(Long 1992) (Symonds et al. 2006) (Stack 2004)]. Some authors argue that only later in their career women more or less catch up with male researchers (Long 1992, Symonds et al. 2006) but the previous early productivity has negative effect on career (Hunter and Leahey 2010). **Attrition** is also very important since it has been recently found that drop out rates explain a large propotion of the career wise differences in productivity and impact (Huang et al. 2019); catching up is only a possibility for those who stay.

88. The relevance of studying whole cohorts and not only successful scientists was already stressed by Preston (Preston 2004) in her book "leaving science"; there she focused on the interrelations between perception of discrimination, behaviour and outcomes (p.145); when the whole cohort was considered, the previous finding of Cole and Zuckerman (J. R. Cole and Zuckerman 1991) of lack of perception of discriminatory treatment did not hold. Among the reasons behind attrition from scientific careers, both for men and women are: inadequate salary and opportunities, difficulties to balance family and career opportunities, lack of mentoring and mismatch between individual's interests and the requirements of scientific profession. Preston confirms that women are more at risk of being affected by the structural factors. On the contrary, other research finds that in the early career, women and men perform about equal but after the age of 38 the differences in productivity and impact increase quickly, especially in health sciences; in the natural sciences and engineering (NSE) and social sciences and humanities (SSH), the differences are smaller (Larivière et al. 2011).

89. **Family formation** has also been claimed to underlie differences not only in productivity but also in the structure of academic ranks by gender. For the U.S. universities Wolfinger et al. (Wolfinger, Mason, and Goulden 2008) found that family and children accounted for the lower rate at which women obtain tenure-track jobs in the humanities and social sciences. Single women without young children did better than their male counterparts on the market for assistant professorships. However, family formation could not account for women's difficulties at later career stages—namely, tenure and promotion to full professor. Here again, there is not a consensus in the literature and other studies argue that

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

the academic career does not suffer from parenthood [(J. R. Cole and Zuckerman 1991), (Fox and Faver 1985), (Fox 2005)] or that children are not a strong predictor of productivity, although the influence that they do have follows a gendered pattern (Stack 2004). Along similar lines, Kelly and Grant (Kelly and Grant 2012) focusing on salaries, find more evidence for fatherhood premiums than for motherhood penalties, especially for non-STEM faculty. The criticism that can be made to many of these studies is their consideration of a very limited set of variables. Qualitative research consistently shows that academic women who are married and have children face considerable stress and time-management pressure in their everyday lives (Caprile et al. 2012) .

90. Academic gender imbalances may also be result of **gender differences in life choices and social pressures** (Ceci and Williams 2011). Disentangling self-selection (choice) and social-selection processes is not easy, and one should not underestimate the constraining power of cultural expectations. Van Arensbergen et al. (2012) (van Arensbergen, van der Weijden, and van den Besselaar 2012) note that there are several differences between male and female researchers' performance which are partly based on one or on the other. For instance, some [(Erin Leahey 2006), (E. Leahey, Crockett, and Hunter 2008), (A. Conti, Denas, and Visentin 2014)] have argued that female researchers lose out an important means of increasing their productivity due to **their lower specialisation**, although Abramo et al. (Abramo, D'Angelo, and Costa 2018) urge caution in identifying research diversification as a co-determinant of the gender productivity gap between man and women.

91. The roots of the difference in publication productivity also include **access to resources**. The much cited work of Xie and Shauman (Xie and Shauman 1998) where they analysed gender differences in productivity in the U.S. during the 70s and 80s, already highlighted that the primary factor affecting women scientists' research productivity was their overall structural position, such as institutional affiliation and rank: when type of institution, teaching load, funding level and research assistance were controlled for, the productivity gender gap disappeared. And more recent evidence (Rørstad and Aksnes 2015) also stresses the importance of the academic position and availability of research funds. An important methodological consideration is that interaction effects with gender are worth testing.

2.1.3 Collaboration and mentorship

92. In bibliometrics the relationship between productivity and collaboration is well known. Acording to the Elservier report "Gender in the global research landscape" (Elsevier 2018) women tend to have less developed international collaboration and co-authoring networks. Using data from UK business schools, Brooks et al. (Brooks, Fenton, and Walker 2014) show important differences in the rated quality of journals than men and women publish in across sub-disciplines, with the former publishing in more rated ones. They also report that women who are able to use networks to co-author with individuals outside their institution

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

are able to publish in higher rated journals, although the same is not true for men. **Disentangling the effect of gender from the effect of status in the motivations to collaborate and collaboration dynamics** that generate scientific outputs is not easy. Gaughan and Bozeman (Gaughan and Bozeman 2016) came across the fact that both status and gender were used as interpretative framework of collaborative behaviour in STEM research, but that the status hierarchy exerted more clear effects. Another interesting finding related to gendered patterns of collaboration has been recently published by Salerno et al. (Salerno et al. 2019) showing, for Latin American ecology and zoology fields, that research groups led by women published with over 60% female co-authors whereas those led by men published with less than 20% female co-authors.

93. Whether female and male researchers in STEM disciplines **differ in their collaboration propensity** is an empirical question. Zeng et al. (Zeng et al. 2016) report on an analysis of the publication records of 3,980 faculty members in six STEM disciplines at select U.S. research universities. They find that female faculty has significantly fewer distinct co-authors over their careers than men, but that this difference can be fully accounted for by women's lower publication rate and shorter career lengths. They also showed that female scientists had a lower probability of repeating previous co-authors than men. Similarly, Beaudry and Lariviere (Beaudry and Larivière 2016), report, for Quebec, that researchers who collaborate with a higher proportion of female co-authors are consistently less cited in both the health and in STEM fields than if they were publishing with a male dominated group of co-authors.

94. **Mentorship (and its quality)** has also been a factor studied in relation to scientific performance and career advancement. However scholars have also acknowlged the problems to construct a proper theory (Bozeman and Feeney 2007) . More recently Ooms et al. (Ooms, Werker, and Hopp 2019) have analysed to what extent different factors (like research orientation, gender, or disciplinary and cultural differences) that affect the relationship with PhD supervisors, help or hamper academics' careers depending on heterogeniety. Their results confirm that heterogeneity stemming from research orientation is helpful for career advancement, but heterogeneity stemming from gender hinders women careers; mentor–mentee heterogeneity only helps in early career transitions, but hampers advancement later on.

95. Some insights can also be drawn from policy intervention. For example, Blau et al. (Blau et al. 2010) evaluated the success of a program (CeMENT) aimed at assisting female junior faculty in preparing themselves for the tenure hurdle, supported by the NSF ADVANCE initiative. They use a randomised trial. Applicants were randomly assigned to be in the "treatment" (mentees who attended the workshop) or "controls" who did not participate. The study compared the academic performance (i.e., papers, grants) of these two groups in the following years. They found significantly increased publication rates and successful grant applications for the treated group, while it was too early to assess the eventual effect on tenure.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.1.4 Recognition and impact

96. **Research impact** is one important dimension of gender differences in performance because it conditions career attrition and opportunities for grant success and promotion. One reason why male-authored research may be more cited is self-citation; another reason is that researchers that author more papers over their careers tend to produce higher impact work, advantaging men for higher productivity, dominance of senior roles and fewer career breaks. According to Lariviere and Sugimoto (Larivière and Sugimoto 2017) using Elsevier data, the average impact factor of journals for men and women are much closer to parity than citations and, systematically, **the gap in citations is much greater** than the gap in impact factors, and always in favour of men. In other words, even when women publish their works in high impact factor journals, this work is cited less, both in absolute and relative terms.

97. The quantitative evidence for the EU regarding gender differences in publications is not as abundant as in the U.S.. In the last decade there have interesting country-cases scholarly contributions. For instance, trying to understand the common argument that female researchers publish on average less than male researchers do, but male and female authored papers have an equal impact, van den Basselaar and Sandström (van den Besselaar and Sandström 2017) studied the publications and citations of 47,000 Swedish researchers; their analysis reveals that, in order to have impact, quantity does make a difference for male and female researchers alike, but women are vastly underrepresented in the group of most productive researchers; they find that gender differences in age, authorship position, and academic rank do explain quite a part of the productivity differences. Moreover, Sandström (Sandström 2009a) stressed the paradox, for medical research in Sweden, that female scientists may be less productive in both normal and fractional paper counts, but their impact in field-normalised citations was better than that of their male colleagues. A further complexitiy is given by the way in which credit is allocated to multiple co-authors (perceived or actual contribution) (H.-W. Shen and Barabasi 2014)

98. However, the results of more recent data analysis (2011-2018) by Thelwall (Thelwall 2018) conflict with prior studies that found a male citation advantage. Interestingly, the most likely reason this author provides for the difference in the findings is the use of arithmetic mean-based normalisation of all prior studies, which gives a substantial male advantage in comparison with geometric mean-based normalisation. He also mentioned the relevance of data analysis with higher levels of domain disaggregation. Differences by nations and fields and subfields are also hidden depending on the level of aggregation: international findings do not necessarily apply to individual countries and *vice versa.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.2 Application for funding: preferences versus expectations

99. Most of the policy initiatives that a few decades ago raised the awareness about the gender differences in science brought to our attention the existence of differences in the success rates (and soon after on applications) in funding between men and women. Later on the issue was formulated into a research question: **What factors condition the probabilities of success?**

100. Before we are able to address the issue of success we first need to understand that the success probabilities are conditional on the probability to apply. Evidence shows that **women apply less** even in sectors where gender distribution is nearly equal like the medical sciences in the U.S. (Ley and Hamilton 2008). In short, application rates do not mirror the potential pool (European Commission 2009). This also seems to be the case more generally in countries like Germany where Auspurg and Hinz (Auspurg and Hinz 2010) indicated that female scientists were more reluctant to apply for individual grants than males scientists.

101. Is there a theory about what are the factors and conditions that explain researchers' application behaviour? Why do people apply for funding? Needs, reputation/status, norm alligment, support, mentoring, integration in networks? Are there differences between men and women? Who applies for funding? How does supply and demand adjust?

102. **Many factors influence grant application behaviour**; some of them are structural and related with the prior segregation by occupation or rank. Survey results reported by Blake and La Valle (Blake and La Valle 2000) for applications to UK Research Councils and the Wellcome Trust show that women were as successful as men in getting the grants they apply for, but were less likely to apply because of their status in the institution and the support they received. The main influences on grant application behaviour were: **seniority, employment status, tenure, type of institution, professional profile, institutional support, career breaks and family circumstances**. Additionally, women were less likely to be working in universities which were the main recipients of research funding.

2.2.1 The mechanisms

103. As noted by Stephan and Ganainy (Stephan and El-Ganainy 2007), it might be analytically useful to **distinguish the contextual (structural) explanation** that women are underrepresented in the type of positions from which typically they would be more likely to engage in a particular activity (e.g application behaviour of women is much more affected negatively by elegibility criteria), from other type of explanations based **on factors affecting supply** (attitudes towards competition, risk, preferences about work-life

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

balance) **and demand** (role of networks, preferences of evaluators and employers, gender discounting[3], etc.).

104. **Regarding supply factors**, some research has explored the hypothesis that men and women view professional advancement differently and their views affect their decisions to climb (or not) the corporate ladder, and seeing professional advancement as equally attainable, but less desirable (Gino, Wilmuth, and Brooks 2015b) with women associating high level positions with conflict and tradeoffs more than men.

105. In this regard, De Paola et al. (De Paola, Ponzo, and Scoppa 2017) found that, controlling for productivity and a number of individual and field characteristics, **women had a lower probability of applying** for promotion of about 4 percentage points. They argue that determinants of this gap seem to be gender differences in **risk-aversion and self-confidence** as well as women's f**ear of discrimination**: the lower tendency to apply is especially relevant for women in the lower tail of the distribution of scientific productivity and in fields in which productivity is not easily measurable; furthermore, women are less likely to apply for promotion in fields in which promotions of women in the past were rare.

106. The so-called "**aspiration gap"** is a supply side explanation of application differences that has been countered by the argument that it is not that women have intrinsically lower levels of career aspirations than men, but that their **experience with gender inequality in organisations leads them to have lower expectations** about their potential for success (Fernandez-Mateo and Kaplan 2018). In other words, t**he supply is affected by the demand**, where demand can be understood as derived from decisions of evaluators, employers, etc. A question arises of what organisations can do to change expectations and/or gendered outcomes that may influence application behaviour.

2.2.1.1 Elegibility of funding instruments, employment conditions and incentives to apply

107. While many factors affect both men and women, some disproportionately stop women from making applications, especially if they are formalised in the **elegibility criteria of the calls**. For instance, rank or employment criteria designed by research funders to help define who can apply for research funding can produce a gender disadvantage at the application stage, because more women than men are employed on fixed-term contracts, part-time posts and are at lower academic ranks. In the Blake and La Valle survey (Blake and La Valle 2000) women were less likely than men to be eligible for the grants due to their type of employment. As for the role in application, men were more likely to apply as principal investigators than women, a result that coud be explained by differences in seniority; interestingly, women were more likely than men to have applied for the salary to be paid by the grant, which suggests a higher representation of women in non permanent posts. In sum, many of the gender differences in application behaviour identified in the survey were rooted in higher education institutions' **employment practices**. That is why

[3] For Stephan and El-Ganainy (Stephan and El-Ganainy 2007), gender discounting occurs when, holding everything else equal, the acomplishements of women are seen differently than those of men, and receive differential treatment.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

we need to know the employment practices of institutions to elaborate a model of the application for grants (European Commission 2009) .

108. The eligibility requirements of funding schemes play a crucial role since review systems only take effect when it comes to deciding on submissions and actual applicants; however, each potential applicant makes an **active decision on whether to apply or not**. Here, **self-selection** plays a relevant role. This is the approach taken by Neufeld and Hornbostel (Neufeld and Hornbostel 2012) in their study of the Emmy Noether Programme of the German Research Foundation; they aim to address the question of whether the best apply, and report that, with respect to bibliometric indicators, the actual applicants' group shows higher performance than the potential applicants' group in the early phase of their career, and that this could be explained by the lack of low performers in the applicants' group. Potential applicants perform less well in terms of quantity and impact, but it remains open what will happen in the long term. They conclude that the answer to the question "do the best apply"? depends very much on the time window considered and the criteria used to define "the best". Furthermore, these authors find interesting differences by field that make them wonder who is actually making the application decision: is it the potential **applicants or the mentors** who advice their well prepared mentees to apply for funding in the relevant programme? The career stage also matters.

109. Ley and Hamilton (Ley and Hamilton 2008) data suggest that a large fraction of female medical scientists choose to leave U.S. NIH-funded career pipeline at the transition to independence (i.e., in the late postdoctoral and early faculty years). Since men and women have near-equal NIH funding success at all stages of their careers, it is very unlikely that, in this case, **women's attrition** is due to negative selection from NIH grant-funding decisions.

2.2.1.2 Attitudes and preferences towards competition and risk aversion

110. Can the argument be made about women's higher risk aversion in relation to decisions to apply to competitive research funding and fellowships? The hypothesis of women being less competitive /having less competitive attitudes has also been studied in some experimental work. However, in many instances, evidence about choices and preferences is not conclusive. For instance, in a small sample comparative study of German and U.S. university, Sieveriding et al. (Sieverding et al. 2018) analysed **lifestyle preferences** and the gender gap in higher positions in academia, and they found some effect of preferences about ideal work hours of early career female researchers with children in German universities but not in the U.S..

111. Indeed, the claim that **women shy away from competition and are more risk-averse** than men has become widespread in the economic literature [(Croson and Gneezy 2009), (Niederle and Vesterlund 2007)]; for example Niederle and Vesterlund (2007) found, **with an experimental approach**, that although there were no differences in performance, men selected the tournament twice as much as women when choosing their compensation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

scheme for the next performance; they concluded that women are more likely to shy away from competition. In a later study, Niederle and Vesterlund (Niederle and Vesterlund 2011) reviewed the robustness of the differences and stated that laboratory and field studies largely showed that gender **differences in competitiveness** tend to result from men is greater **overconfidence**. They also found that gender differences in risk aversion, however, seemed to play a smaller and less robust role. But this does not mean that women are not able to perform in competitive environments per se.

112. For instance, in a laboratory experiment, Gneezy et al. (Gneezy, Niederle, and Rustichini 2003) observed a significant **increase in performance for men, but not for women when there was an increase in the competitiveness of the environment**. This resulted in a significant gender gap in performance in tournaments, while there was no gap when participants were rewarded according to piece rate. Moreover, this effect was stronger when women had to compete against men than in single-sex competitive environments.

113. Notwithstanding previous evidence, Nelson [(Nelson 2014), (Nelson 2015)] makes a critical review of some of the economic literature on gender and risk and argues that the **claims about gender and risk aversion have gone far beyond** what can be justified by the actual quantitative magnitudes of detectable differences and similarities that appear in the data. She points out that cultural beliefs combined with generally accepted but non-rigorous methodological practices has drifted a whole **literature in a particular direction and identifies six kinds of biases:** 1) inaccurate citations of earlier literature which favours essentialist beliefs, 2) generalisations from samples of individuals to categories of the population 3) overemphasis on differences within a study's own results, unsupported by the actual magnitudes of such differences 4) publication and confirmation bias that tend to favour significant over non significant results and those consistent with a prior belief, 5) failure to consider confounding variables such as social pressures to conform to expectations, and 6) examination of a narrow range of risks (mostly lottery, gambling and investment scenarios).

2.2.1.3 Expectations and interaction feedback (learning)

114. In a similar vein, Kugler et al. (Kugler, Tinsley, and Ukhaneva 2017) put into question recent work that suggests that women are **more responsive to negative feedback** than men in certain environments; when they analyse the effect of low grades in major related courses in switching out from male dominated STEM majors among students, they do not find support for such claim.

115. Thus, for some authors a more realistic account of the empirical evidence do not support the claim of intrinsic differences between men and women in this issue; factors like minority status, and different reactions to men's and women's choices may produce behaviours that look like risk aversion but that in fact derive from the different incentives or payoffs to men and women to taking action (Fernandez-Mateo and Kaplan 2018).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

116. **In sum, we must acknowledge that the literature reports contradictory findings and is rather inconclusive about the causes of differential application**. One reason for this lack of robustness may be that studies seldom take feedback dynamics into account. Interestingly, Brands and Fernandez-Mateo (Brands and Fernandez-Mateo 2017) study **how rejection experiences shape women's decisions to compete for executive roles**. They theorise and show that women are less likely than men to consider another job with a prospective employer that has rejected them in the past. The mechanism would be that rejection triggers uncertainty about women's self-assessed general belonging in the executive domain, which in turn leads women to place greater weight than men on fair treatment and negatively affects their perception of the fairness of the treatment they receive; this dual process would make women less inclined than men to apply to a firm that has rejected them before.

117. What is interesting about this result is that, as mentioned Fernandez-Mateo and Kaplan (Fernandez-Mateo and Kaplan 2018) **demand effects are often confused with supply effects**. In their view, what looks like a supply problem -that women choose not to aspire to top positions or to jobs in top paying fields- might actually be a demand problem -organisations or jobs look unappealing to women because of past histories of not hiring or promoting women into leadership roles or of making work–life balance appear to be impossible-.

118. Although the dominant approach in this literature is focused on the individuals, we also need to understand the diversity of propensity to apply that the different organisational conditions of academic institutions creates. The majority of analyses do not incorporate the **organisational conditions that incentivise researchers** within particular institutions or career conditions to apply for competitive funding. It might be the case that those organisational conditions, for instance the research intensity of the institution, are gender neutral[4], but to the extent that there are differences by gender in the employment conditions, organisational incentives to apply may be mediated by gendered employment structures which in turn affect individual researchers' autonomy, so that the personal incentives to get individual fellowships (versus project funding) may actually be higher for women (unless projects include salary paid for them).

2.3 Evaluation and allocation of limited resources: Selection of proposals and access to jobs

119. A decade ago, the European Commission issued an expert report about the "gender challenge in research funding" (European Commission 2009); data referred to 2007. The report compared several countries and noted that, **in average, proposals with a male IP had 7% more probabilities of being granted** than those with a female IP. There were, however, notable differences across countries and fields. For instance in the health

[4] Although there is evidence on the contrary, showing that research intensity at universities is negatively correlated with the share of female academic faculty [(Long, Allison, and McGinnis 1993), (Hargens 2012)].

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

sciences the majority of countries showed male advantage, whereas in the engineering and technology female researchers got higher success rates in 10 out of 18 countries. In the humanities the balance was slightly in favour of women; in the social sciences 8 countries showed a balance in favour of men and other 8 showed the opposite. However, from differences in the success rates one cannot infer differences in individual productivity, since there was no information in this report to control for other factors such as the institutional affiliation, the quality of the team, previous national funding, etc. that affect the probability that a proposal is granted. From success rates, we can neither infer the existence of discrimination or bias.

120. In the U.S., in order to assess gender differences in research funding, RAND Corporation's researchers (Hosek et al. 2005) analysed three years of data from three federal agencies (NSF, NIH and USD) that accounted for three-fifths of the $43 billion in research funding in FY 2001 and two surveys. **They found no gender differences in federal grant funding outcomes** when they adjusted for other characteristics of applicants, including the researcher's discipline, institution, experience, and past research output. Interestingly, the two exceptions they found referred to NIH. Firstly, they found a gender gap in the average amount of funding that women received relative to their male counterparts, although some caveats were acknowledged: for instance, the gap was smaller when the top 1% grants were excluded, a group in which women were less represented; they also found a gender gap in subsequent application rates.

121. The most recent evidence of the NIH shows that sex differences in the size of grants awarded to comparable first-time female and male PIs exist - also if one only looks at top research institutions. In most grant types, men get more than women. But for R01 grants (the most frequent award for first-time awardees) women received larger grants (Oliveira et al. 2019). In this paper, however, only successful applicants are included; further study of the institutions where inequalities were lowest may provide insights into the reasons for gender imbalances in grant amounts awarded during formative career stages.

122. In this section we review the literature on grant evaluation paying attention firstly to the **processes** and secondly to the **characteristics of the evaluators and applicants**, with a focus on gender.

2.3.1 The process of evaluation

2.3.1.1 Are criteria of assessment gendered? The notion of excellence

123. Academia is a culture of meritocracy, value neutrality, and excellence, while the logic of gender equality is based on values of justice, equal rights, and democracy; Powell (Powell 2018) argues that these logics are incompatible since concepts of meritocracy and excellence are permeated by gendered stereotypes and so the way scientific quality and merit are valued . Although the quest for excellence has been a very prominent idea in the policy discourse, the empirical literature about the **evaluation of excellence** is limited (European Commission 2004). Excellence is not a naturally given variable but socially

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

produced; among the social processes, some argue that who is considered excellent depends on gender relations in the scientific community and in society at large [(Addis, 2019),(Addis, n.d.)]. Additionally, scientific excellence is not a hidden feature of the single scentist; it is a quality of the researcher in relation to others. Addis (Addis 2010) points to some specific issues that emerge in the literature and that are relevant to the issue of **why women may not achieve scientific excellence: homosociability** that favours cooptation of similar people**, gatekeeping** which implies decisions on who enters and who remains outside**, peer evaluation** and what it entails from a gender view point**, and leadership** and the distribution of men and women in leadership positions.

124. Van den Brink and Benschop (van den Brink and Benschop 2012) are more critical and, based on a study of professorial appointments in the Netherlands, argue that **academic excellence is a social construct** that is inherently gendered. Furthermore, they describe that committee members highlighted the decisive role of professional qualifications such as quality and quantity of research, success in obtaining research funding and experience in management in the assessment of candidates, but that despite the gender neutrality of those criteria, their use could also lead to the production or reproduction of inequities, if opportunities to develop and perform in those activities were unequally distributed *ex ante*. Thun (Thun 2019) in a qualitiative case study of the University of Oslo argues that organisations advantage particular groups of academics that correspond to the "ideal type". The focus on excellence may have consequences but work is needed about which consequences and how.

125. At this point, the distinction between **merit and worth** (Lincoln and Guba 1980) seems pertinent. Both merit and worth are aspects of value, but while merit is intrinsic to the individual and can be assessed by the degree to which he/she conforms to certain standards upon which experts agree, worth is extrinsic and is determined by comparing the individual´s value relative to some set of external requirements. Whereas merit criteria are stable, criteria of worth are highly variable depending on context; depending on the instrument, specific programme or call, research funding agencies may put more emphasis on either of those aspects of value. When panels evaluate the adjustment of the candidate and/or the proposal to the host institution, both aspects are being taken into account. Grants that are individual-oriented and portable across institutions will in principle emphasise merit more than worth.

126. In contrast, at the **organisational level, tenure decisions** are the paradigmatic example of evaluations where the candidate must not only show quality, but demonstrate that he/she serves some organisational purpose too (Cruz-Castro and Sanz-Menéndez forthcoming).

127. The **notion of talent** and its relationship with merit and worth also deserves some attention. Qualitative research has shown that, overall, there is large agreement on the notion of scholarly talent among scholars who have experience with identifying talent both in their daily academic work and in the process of grant allocation; the main criteria for assessing talent are publications, study and promotion results, honours degrees, awards,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

grants, and international experience; but talent is also associated to the idea of performing better than expected according to age or experience, and it is not only about professional capital, but also about individual capital (personal traits, ambition, motivation etc) and social capital (networks). Despite this broad agreement, the characteristics ascribed to top talent vary depending on the evaluation context. In grant allocation, a narrower talent definition prevails compared with more general evaluation, and, furthermore, difficulties emerge in the process of panel decision-making, when selection criteria need to be concrete and explicit to enable comparison (van Arensbergen, van der Weijden, and van den Besselaar 2014).

2.3.1.2 Bias in the assessment of similar individuals?

128. In general, academic scientists believe that they function within a meritocratic system that rewards capacity and productivity, and that careers are open to talent. In the allocation of competitive research grants and fellowships, the need to **select among applicants of similar quality** makes the grant selection process liable to subjectivity, arbitrariness, and randomness but the central question of **whether grant peer review is gender-biased** is an empirical one. In our view any serious analysis aiming to claim the existence or non existence of gender bias or gender discrimination should take performance into account, and not only success rates.

129. In Europe, a commentary in Nature by Wenneräs and Wold (Wennerås and Wold 1997) of an analysis of peer-review scores for postdoctoral fellowship applications of the Swedish Medical Research Council in early 1990s, brought the issue of disparities in the evaluation outcomes in RFO to the agenda; latter on some criticisms on the methodological grounds and the statistical analysis were made (Hoff Summers 2009); the study raised the issue of gender bias in the allocation of funding. In their seminal article (Wennerås and Wold 1997), they showed that meritocracy was not the standard, that social relations with committee members played a significant role, and that women needed a substantially higher performance than man to get a grant in the biomedical area and that women needed to have 2,5 more publications than men to get a similar rating. Replicating the study a decade later Sandstrom and Hallsten (Sandström and Hällsten 2008) found nepotism again but not gender bias; other research has found gender bias to take place in some fields more than in others; Brouns (Brouns 2004) for instance reported a negative gender bias in the biology but a positive one in other science fields.

130. According to Bormann et al. (Bornmann, Mutz, and Daniel 2007) some narrative accounts of peer review research have concluded that there is negligible evidence of gender bias in the awarding of grants based on peer review. To the contrary, in their meta-analysis[5] of more than twenty studies they claim evidence of robust gender differences in grant award

[5] The term meta-analysis refers to a statistical approach that combines evidence from different studies to obtain an overall estimate of treatment effects (i.e an independent variable); Meta-analysis allows generalised statements on the strength of the effects, regardless the singularity of individual studies; it is necessary, however, that each study be designed similarly with respect to certain properties (methods, sample…) (Bornmann 2011): 215) .

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

procedures. Even though the estimates of the gender effect vary substantially from study to study, their model estimation shows that all in all, among grant applicants men have statistically significant greater odds of receiving grants than women by about 7%. Although the figure may seem small, even small differences accumulate to multiplicative impacts to women underrepresentation over time.

131. However, extending this meta-analysis, Marsh et al. (Marsh et al. 2009), found no gender differences in grant proposals. They warn that the previous meta-analysis showed a substantial heterogenieity in the effect sizes for the different review processes considered in the analysis, which, in their opinion, compromised the results. Methodologically, they juxtaposed traditional (fixed- and random-effects) and multilevel models, arguing in favour of the advantages of the multilevel approach. Using the latter, the lack of gender effect for grant proposals was robust. Later on, in an empirical analysis of the Austrian Science Fund, Mutz et al. ((R. Mutz, Bornmann, and Daniel 2012) found that there was not discrimination against women in the marks in the review process.

2.3.1.3 Gender effects of Evaluation tools

132. The existence of gender bias in the definition and use of **tools for measuring scientific performance** and assessing scientific excellence is another topic in literature (Addis 2010) .

133. Bibliometrics are used as a proxy for excellence, quality and ability. For some, bibliometrics is gender-blind or gender-insensitive but the use of bibliometrics may be not gender-neutral (Addis 2010); some case studies have argued that the use of bibliometric indicators widen the gender gap in research performance. For instance, Nielsen (M W Nielsen 2017b) finds that in the Danish bibliometric research indicator (BRI), men on average receive more BRI points for their publications than women, and suggests two possible explanations: firstly, women merely comprise 24% of committee members determining the quality classification of communication channels, so women's interest and opinions are less likely to be voices in the discussion of "what counts' as excellent publication chanels; and secondly, the model privileges collaborative research, which disadvantages women due to gender differences in collaborative network relations (M W Nielsen 2017a) (M W Nielsen 2015).

134. Strong advocates of the use of bibliometric criteria to asses scientific merit often base their argument in the gender-neutral character of those tools; however, others point out that (Mathias Wullum Nielsen 2018) "bibliometrics, when used at the individual evaluation level, can serve to perpetuate existing gender inequalities in academia, by providing indisputable and easily measurable proxies for merit that decontextualises scientific achievements and transform different qualities into common metrics".

135. Based on a qualitative study of recruitment at a Danish university, Nielsen (Mathias Wullum Nielsen 2018) is critical about the use of tools such as the h-index that is slanted in favour of researchers who publish in subfields with high citation frequencies, is highly

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

correlated with output and therefore with scientific age, and privileges individual publishing with multiple co-authors, a factor that have been argued to disadvantage women [(Abramo, D'Angelo, and Murgia 2013), (Bozeman and Corley 2004)]. Despite these potential biases, bibliometric measures may function as technologies supporting a managerial narrative of the gender-blind organisation. Notwithstanding the basis of these criticisms, some studies have maybe gone too far in arguing about implications of the use of the h-index (e.g the work of Geraci et al. (Geraci, Balsis, and Busch 2015) about salary differences in psychology).

136. Also regarding e**valuation methods,** Jappelli et al. (Jappelli, Nappi, and Torrini 2017) compiled data on 180,000 research papers evaluated during the Italian national research assessment conducted by the Agency for the Evaluation of Universities and Research Institutes over a 6 year period, and merged this data with information on individual researchers and characteristics of referees. They found a significant gender gap in research evaluation. The gap was reduced once they controlled for researchers' characteristics, such as age and academic rank, but was almost unaffected by research output, co-authorships, and international collaborations. Childbearing and maternity leaves did not account for the remaining gap in research evaluation. The evaluation method (peer review or bibliometric analysis) and the referee mix (whether men or women) did not disadvantage women. Interestingly, the analysis of a random sample of papers evaluated using bibliometric indicators and peer review revealed that bibliometric evaluation proved to be more favourable to women than peer review evaluation.

2.3.1.4 Assessing the research proposals or the candidates?

137. The degree to which the evaluation focuses on **candidates or proposals** and in what order is another relevant aspect. A recent experiment-based paper analysed whether imbalances in men and women research funding stem from evaluations of female research investigators or from their proposed research, finding that gender gaps in grant funding were attributable to less favourable assessments of women as principal investigators, not of the quality of their proposed research (Witteman et al. 2019).

138. Van der Lee and Ellemers (R. van der Lee and Ellemers 2015) found similar results in their study of research funding in the Netherlands, where male applicants to the Veni programme received significantly more competitive "quality of researcher" evaluations (but not "quality of proposal" evaluations) and had significantly higher application success rates than female applicants. Gender disparities were most prevalent in scientific disciplines with the highest number of applications and with equal gender distribution among the applicants (i.e., life sciences and social sciences).

139. However, van den Besselaar et al. (van den Besselaar, Sandström, and Schiffbaenker 2018) in a linguistic analysis of the peer review reports of the ERC Starting grants showed that the performance of the applicant and the content of the proposed project were assessed with the same categories, suggesting that the panellists actually do not make a difference

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

between past performance and promising new research ideas. Their analysis also suggests that panels concentrate in rejecting weak proposals rather than trying to find the best and more promising ones; this is compatible with a strategy of reducing the pool of applications which is also related to the low success rates of these grants.

2.3.1.5 Transparency, clarity of the evaluation criteria and wording of reviews

140. The analysis of **texts and wording in reviews** has produced interesting evidence about the process of evaluation in RFO. Using text mining methodology, Magua et al. (Magua et al. 2017) analysed 51 reviews of grant renewal applications to the NIH awarded to 45 investigators at a U.S. university. Their results showed that male investigators were described as "leaders", "pioneers" in their fields, with "highly innovative" and "highly significant research", whereas female investigators were characterised as having expertise and working in "excellent environments". Applications from men received significantly better priority and significance scores. They conclude that such implicit bias may contribute to sex differences in the award rates of such renewals.

141. Similarly, Rubini and Menegatti (Rubini and Menegatti 2014) investigated the role of language abstraction as a means to discriminate female applicants in academic personnel selection. The level of abstraction of 814 judgments, which were drawn from publicly available archival material, was coded. Results revealed that judgments of female applicants were formulated using negative terms at a more abstract level and positive terms at a more concrete level than those of male applicants. Moreover, this linguistic form of gender discrimination was found only in male committee members.

142. Although there is evidence that **formalisation of evaluation criteria and procedures** influence clarity and transparency of academic assessment, Helgesson and Sjogren (Helgesson and Sjögren 2019) in a case study of a Swedish higher education institution, found that inequality was reproduced through the choice of implictly gendered metrics; they reported how the formalisation of a "good enough" standard, in addition to a standard of "excellence" reinforced the scope for interpretation flexibility among reviewers broadening their discretion to communicate or hide failure, with gendering effects.

143. Not only words, but also **numbers may convey symbolic value**. Quantitative performance ratings are ubiquitous in modern organisations, and even the evaluation scales may have a gendered impact. An interesting experiment conducted by Rivera and Tilcsik (Rivera and Tilcsik 2019) at a large North American university, found that the number of scale points used in faculty teaching evaluations-whether lecturers were rated on a scale of 6 versus a scale of 10-significantly affected the size of the gender gap in evaluations in the most male-dominated fields, so that women fare better in the 6 point scale than in the 10 point one. They suggested that the number of scale points affects the extent to which gender stereotypes of brilliance are expressed in quantitative ratings due to the cultural meaning of numbers.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.3.1.6 Does blinded review reduce or eliminate gender disparities?

144. **Blinded review,** which has been the standard in journal publication peer review, is an increasingly popular approach to reducing bias and increasing diversity in the selection of people and projects. The **effectiveness of blinded peer review has also produced mixed results.** Some studies have analysed the effect of blinding and found that it reduced reviewing bias with regards to personal characteristics of the author, including sex, friendship and/or intellectual conformity with the reviewer [(J. S. Ross et al. 2006), (Goldin and Rouse 2000), (Ross-Hellauer, Deppe, and Schmidt 2017)].

145. However Kolev et al. (Kolev, Fuentes-Medel, and Murray 2019) explore its impact in a study of the innovative research grant proposals submitted to the Gates Foundation from 2008-2017. They find that **despite blinded review, female applicants receive significantly lower scores**, which cannot be explained by reviewer characteristics, proposal topics, or ex-ante measures of applicant quality. By contrast, the gender score gap is no longer significant after controlling for text-based measures of proposals' titles and descriptions (broad versus narrow words), suggesting that, in this case, differing communication styles are a key driver of the gender score gap.

146. In a recent scoping review, Tricco et al. (Tricco et al. 2017) screened more than 5,000 citations and 170 full-text articles in health-related research, concluding that the research on the impact of gender-blinding of grant applications is extremely limited; they identified only one study in which 891 applications for long-term fellowships were included and 47% of the applicants were women. These were scored by 13 peer reviewers (38% were women). The intervention included eliminating references to gender from the applications, letters of recommendations, and interview reports that were sent to the committee members for evaluation. The proportion of successful applications receiving grant funding led by women did not change with gender-blinding, although the number of successful applications that were led by men increased slightly.

2.3.1.7 Negative organisational conditions

147. In an interesting review, Heilman (Heilman 2001) identifies a number of **organisational conditions that may hinder the valuation of women's performance.** Firstly, ambiguity in evaluation criteria: the more vague the judgment criteria, the more easily information can be distorted to fit preconceived ideas about capacity and performance; secondly, the lack of structure in evaluation and decision making processes; thirdly, ambiguity about the source of successful performance: there are some organisational conditions that blur the contribution of individuals, the emphasis on teams is a case in point if group work encourages attributional explanations that limit the degree to which women are seen as responsible for their success; fourthly, ambiguity about the reasons for past success: in this sense, affirmative action or diversity programmes may have unintended consequences in terms of perception of preferential treatment along the career.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

2.3.2 Actors involved in the evaluation: Who evaluates? Who decides?

148. Peer and panel review are the dominant modes of evaluation for research funding, and albeit being the final arbiter of what is valued in academia and in science, the method has been criticised in relation to criteria of r**eliability, fairness and predictive validity [**(Bornmann and Daniel 2005), (Bornmann 2011)].

149. Human decisions are classified as reliable when different persons come to the same or similar conclusions; applied to fellowships or grant awarding processes **reliability** refers to the degree of agreement among reviewers. A process of review of fellowhips or grant applications is **fair** when these are judge solely on the basis of scientific merit, and not on the personal characteristics of the applicants. **Predictive validity** refers to whether the "best" applicants are selected to receive the grants. Assesing the predictive validity of decisions requires an accepted criterion of scientific merit: a conventional approach is to use publications and impact metrics.

150. The rationale for looking closely at evaluation committees lies in the possible existence of the so-called interaction effects. Because attributes of the applicants and attributes of the reviewers are potential sources of bias in peer review, both should be included in the empirical analyses. Sometimes the characteristics of the grant reviewers and dynamics in the panels (gender of reviewer and/or applicant, conflict of interest, panel agreement, etc.) matter to an extent of sufficient magnitude to change application scores from fundable to nonfundable (Tamblyn et al. 2018) .

151. **Peer review is a gate keeper**. Merton (Merton 1973) called "the gate keeper" the fourth major role of a scientist, in addition to researcher, teacher and administrator. Gate keepers are in a key position to control or influence the entry to an area and the access to resources. In the context of research funding, gate-keepers refer both to fund awarding organisations as collective gate-keepers and to individuals who are involved in decision making bodies of such organisations. According to the ETAN Network Report (2000) (ETAN 2000), to a large extent, the gate keepers of research funding in Europe comprise middle-aged male academics, and this predominance applies even in countries, like Finland, where the proportion of women among professors is higher than in other European countries; this is also the case even in fields with strong women representation (European Commission 2009). In a recent paper, Murray et al. (Murray et al. 2019), find evidence of a relationship between the demographics of the gatekeepers (journal reviewers) and authors in determining the outcome of peer review in the biosciences; that is, gatekeepers favored manuscripts from authors of the same gender and from the same country.

2.3.2.1 The evaluation committe composition: Gender and proximity factors

152. The literature on practices and procedures of grant allocation in RFO is limited. On the contrary, much of the literature of **gender bias within committees** refers to academic recruitment and promotion and the lower success rates of female applicants [(van den Brink, Brouns, and Waslander 2006) , (van den Brink 2010)]; it is possible draw relevant

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

insights from this work applicable to grant selection. According to this literature, women researchers are sometimes at disadvantage in academic recruitment due to insufficient network ties and gender biases among evaluators indicating that formal recruitment procedures allow space for mobilising informal, potentially gendered network ties. Moreover, as noted by Van den Brink et al. (van den Brink, Brouns, and Waslander 2006) , the male network is routinely used in selection and decision making as it is based on trust that male candidates have a lower risk level; although female panel members do the same, they are less successful as they remain a minority.

153. In a study of **academic hiring** in Italy, Abramo et al. (Abramo, D'Angelo, and Rosati 2016) argues that among the factors that determine success in a competition for an academic position, the number of the applicant's career years in the same university as the committee member operated in favour of men. Being of the **same gender as the committee's president** was also a factor that favoured male applicants. On the other hand, for female applicants, the presence of a full professor in the same university with the same family name as the candidate was a better predictor of success than for male candidates. However, their results showed that no gender-related differences occured among the candidates who benefit from positive bias, while among those candidates affected by negative bias, the incidence of women was lower than that of men. This finding is not aligned with part of the literature and the authors provide a number of explanations, arguing that if the concentration of male top scientists is greater than for women then one would in fact expect a greater concentration of discrimination against men.

154. Much has been discussed about the need to increase the **presence of women in review panels**. It is often argued that the presence of women in review panels may improve the selection of female-led research projects; but whether the composition of panels indeed matters is an empirical question that has also been more addressed in relation to hiring, promotion and publication than to funding decisions, although some insights can be drawn.

155. By means of an exploitation of a data set on research r**ecruitment processes to entry-level positions** in a leading Italian research centre operating mainly in the STEM fields, Checchi et al. (Checchi, Cicognani, and Kulic 2019) show that **bias against women is attenuated by the presence of a woman in the selection committee**. However, the most important predictor for recruitment in their study is previous connections with the research centre, a mechanism which, due to the lower density of network links with the institute among female candidates, operates as a selection device discriminating against women. A similar result was reported by Sabatier et al. (Sabatier, Carrere, and Mangematin 2006) in a case study of the factors that speed up or slow down the progress of an academic career of French life scientists, where women had to demonstrate higher involvement in different networks in order to be promoted.

156. In contrast, Bagues et al. (Bagues, Sylos-Labini, and Zinovyeva 2017) analysed how a larger presence of female evaluators affected committee decision-making using information on 100,000 applications to associate and full professorships in Italy and Spain.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

These applications were assessed by 8,000 randomly selected evaluators. They found that a **larger number of women in evaluation committees did not increase either the quantity or the quality of female candidates who qualify**. Moreover, information from individual voting reports suggested that female evaluators were not significantly more favourable toward female candidates. Interestingly, male evaluators became less favourable toward female candidates as soon as a female evaluator joined the committee. But to the contrary, Van der Brink et al. (van den Brink, Brouns, and Waslander 2006) found a positive correlation between the success of women in the selection for university professors in the Netherlands and the number of women in the committees.

157. Again, the evidence regarding whether the presence of female reviewers in the panel reduce gender bias is mixed and one can find support for claims in favour or against. For instance, Murray et al. (Murray et al. 2019) reports, in a study of publications in the biosciences, that the acceptance rate for manuscripts with male last authors was higher than for female last authors, and this gender inequity was greatest when the team of reviewers was all male, and that mixed-gender gatekeeper teams lead to more equitable peer review outcomes. But in the area of grant funding, based on reviews of around 15000 grant proposals to the economic programme of the National Science Foundation, Broder (Broder 1993) presented evidence of **significant differences in the reviewing of female and male authors by male and female referees but in the opposite direction**. Even when author quality was controlled by comparing ratings on the same proposal, female reviewers rated female-authored proposals lower than did their male colleagues while no gender differences in the review of male proposals was observed. These results still hold when controls for institutional affiliation were also included.

158. Other empirical evidence about the impact of researcher's gender on peer review of grant proposals found no significant effect; Mutz et al. (R. Mutz, Bornmann, and Daniel 2012) evaluated the grant peer review process at the Austrian Science Fund with respect to gender over 10 years (8,496 research proposals across all disciplines, rated by more than 18,000 reviewers in almost 24,000 reviews); they found that the final decision was not associated with applicant's gender or with any correspondence between gender of applicants and reviewers, but that the approval probability decreased (up to 10%), when there was parity or a majority of women in the group of reviewers, with no significant effect of the gender of the applicant.

159. Likewise, based on 10,023 reviews by 6,233 external assessors of 2,331 proposals from social science, humanities, and science disciplines, Marsh et al.(Marsh, Jayasinghe, and Bond 2011) found, moreover, that these non-effects of gender generalised over reviewer's gender (contrary to a matching hypothesis), discipline, reviewers chosen by the researchers themselves compared to those chosen by the funding agency, and country of the reviewers. It might be the case that the role of gender composition of the committees is confounded with the role of connections.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

160. Two further elements have been shown to be relevant in academic promotion as well as in grant peer review; they might interact with gender and institutional proximity, or have its own independent effect. The first is the **cognitive or intellectual distance** between the knowledge embodied in the research proposals and the evaluators own expertise, Some experimental research work has found [(Boudreau et al. 2016), (Gallo, Sullivan, and Glisson 2016)] that evaluators systematically gave lower scores to research proposals that were closer to their own areas of expertise and to those that were highly novel; in other words, intellectually close applications were reviewed less favourably than distant ones; this was consistent with biases associated with bounded rational evaluation of new ideas and inconsistent with intellectual distance being associated with private interest of evaluators. Interestingly, reviewer and applicant seniority could influence this relationship, suggesting social networks could have subtle influences on reviewer scoring. The main methodological limitations of some of this research are the lack of randomisation of the reviewers' samples, and the self-assesed nature of the reviewer expertise variable.

161. The second element is precisely **the role of social connections**; a study recently reported in Nature news (17th april 2019) examined more than 28,000 reviews from nearly 13,000 Swiss National Science Foundation applications by about 27,000 peer reviewers from all disciplines between 2006 and 2016. The findings were that reviewers nominated by applicants were more likely to give these applicants higher evaluation scores than referees chosen by the SNSF (Severin et al. 2019) in line with Marsh et al. (Marsh, Jayasinghe, and Bond 2011).

2.3.2.2 Cognitive factors and stereotyping

162. One of the sources of gender bias in judgments is **gender stereotyping**. Stereotyping is a cognitive shortcut. When processing information, we tend to condider observations that match our stereotypical expectations as more veridical, reliable and informative than counter-stereotypical observations (Ellemers 2018). Heilman distinguishes between despcritpive and normative stereotypes. Descriptive gender stereotypes promote gender bias because of the negative performance expectations that result from the perception that there is a poor fit between what women are like and the attributes believed necessary for successful performance (Heilman 2012). This lack of fit model emphasises the role of negative performance expectations in gender biased evaluations. According to Heilman (Heilman 2001) information about successful past performance of women can change the perceived lack of fit derived from the descriptive aspect of gender stereotypes, but not the normative or prescriptive aspect of the stereotype, which is related to preconceived ideas about how women should be like.

163. Most of the evidence on the topic of stereotyping is experimental. Heilman (Heilman 2012) acknowledges that in the vast majority of studies on gender stereotypes no differences have been found in the reactions between female and male respondents; the finding is puzzling, and differences may be in the reasons, so that whereas men may have vested interests in keeping a dominant position, women may be gender biased due to social

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

comparison processes (the so-called "queen bee syndrome"[(Derks et al. 2011), (Derks, Van Laar, and Ellemers 2016)]. An alternative, more general explanation is that stereotypes are widespread in society and affect both men and women.

164. In line with the latter explanation, in an experimental study of evaluations conducted among science faculty of research intensive universities to cover a laboratory manager position, Moss-Racussin et al. (Moss-Racusin et al. 2012) reported that women fare worse in evaluations than men even when the applicant gender was randomly assigned to identical CVs; science faculty were also more likely to hire and mentor the male student and to offer him a higher salary. Faculty gender did not affect bias and female faculty did not rate female students higher. Althouth this type of experimental literature is of value, one also needs to critically assess the limitations of some of the studies; for instance, sometimes respondents only evaluate one CV or one application (Eaton et al. 2019) so that the effect of individual scoring differences cannot be controlled for.

165. Reuben et al. (Reuben, Sapienza, and Zingales 2014) studied the e**ffects of stereotypes** with an experimental market were individuals were hired to perform an arithmetic task that on average both genders performed equally well; they found that without any information about the candidate`s sex, both male and female subjects were twice more likely to hire an man than a woman; discrimination survived if the performance was self-reported because men tended to boast about it whereas women underreport it; the discrimination was reduced, but not eliminated by providing full information about previous performance. Implicit stereotypes made employers biased against women less likely to take into account the fact that men on average boast more than women about their future performance, leading to suboptimal hiring choices.

166. Also with an experimental design, Carli et al. (Carli et al. 2016) found that the higher the proportion of women in a scientific field, the more similar the stereotypes in that field were to stereotypes about women; their results were congruent with theories that report incompatibility of female gender stereotypes with stereotypes about high status occupational roles, since women were perceived to lack the qualities needed to be successful scientists. There is some evidence produced in relation to non-academic contexts in which it is shown that men's and women's academic performance have very different consequences in the general labour market. This is the finding of Quadlin et al. (Quadlin 2018) who report that there is penalty for high-achieving women and that this penalty is most concentrated among women who majored in mathematics. High-achieving women may be most readily penalised when they demonstrate achievement in STEM fields where they are underrepresented and expected to perform poorly. Their survey experiment indicates that these gendered patterns may be attributable to employers' shifting standards for men and women job applicants; employers value competence and commitment among men applicants, but they privilege likeability among women applicants, ultimately creating liabilities for high-achieving women.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

167. Another interesting result was reported by Steinpreis et al. (Steinpreis, Anders, and Ritzke 1999) from a small experimental design with U.S. academic psychologists in which they sent CVs ostensibly submitted by men and women candidates for an assistant professorship and for tenure[6]. They found gender bias for both men and women in preference for male job applicants for the assistant professorship but not for the candidates who had already gotten early tenure, suggesting that stage of career might be a relevant factor.

168. Among the theoretical **explanations** that have tried to explain stereotyping in evaluations is "**role congruity theory**" according to which, beliefs about scientists and women tend to be dissimilar whereas beliefs about men and scientists tend to be similar; this incongruity would affect the judgement of female scientists negatively regarding performance in a scientist role. Knobloch-Westerwick and colleagues [(Knobloch-Westerwick, Glynn, and Huge 2013), (Knobloch-Westerwick and Glynn 2013)] tested this idea in an experiment in which participants rated conference abstracts ostensibly authored by women or men, with authors' associations rotated; the abstracts fell into research areas perceived as gender-typed or gender-neutral. The results were that publications from male authors were associated with greater scientific quality, in particular if the topic was male-typed, and that the respondent sex did not influence these patterns.

2.3.3 A divided literature

169. An even impressionistic comparison of the amount of literature that claims that gender bias exists and persists with the literature that claims that it is diminishing or disappearing makes clear that the former is much larger. However, it is important to pay attention to rival explanations of gender gaps and be careful with inferring processes from outcomes. In a recent work Weisshaar (Weisshaar 2017) analyses tenure promotions in a sample of assistant professors in Sociology, Computer Science, and English departments across U.S. research universities, and show that productivity measures account for a portion of the gender gap in tenure, but that a substantial share of the gender gap remains unexplained by productivity or by departments´caracteristics; she concludes that the unexplained variance is due to gendered inequality in the evaluation process. Apart from controlling for rival explanations, it is important not to fall in "sophisticated residualism" (S. Cole and Fiorentine 1991) whereby all unexplained variance is attributable to an unmeasured variable: discrimination[7].

170. Ceci and Williams (Ceci and Williams 2007) provide us with an overview of the empirical evidence up to then about gender bias in science. According to Ceci and Williams (Ceci

[6] In fact the documents recounted the career of a real woman psychologist who had been hired as an assistant professor and had gotten early tenure.

[7] Weisshaar (Weisshaar 2017) seems quite aware of this fallacy and acknowledges that gendered inequality comes from several factors, including evaluators not giving the same recognition to women and men for equal work quality (discrimination) but also from women's self-selection into different career paths before the promotion takes place or inequality in networks and advocacy in a promotion candidate.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

and Williams 2011) **claims that women suffer discrimination in grant and manuscript reviewing and hiring rest on a set of studies undergirding policies and programs geared at remediation**. In their view, more recent and robust evidence fails to support such claims. Their work has been controversial.

171. Based on a review of the past 20 years of data, they suggest that some of those claims are no longer valid, and, if uncritically accepted as current causes of women' lack of progress, can delay understanding of women's underrepresentation. Moreover, Williams and Ceci [(Williams and Ceci 2015), (Ceci and Williams 2015)] based on some national randomised experiments in the U.S. find no evidence to support the idea of preference for male hiring for tenure track associate professorship but rather the contrary; they report a 2:1 preference for women by faculty of both genders and across math-intensive and non-math-intensive fields. They conclude that **differential gender outcomes result from differences in resources attributable to choices,** whether free or constrained, and that such choices could be influenced and better informed through education if resources were so directed. They deduce that the unequal position of women in science would be based on quality differences which is partly the product of own career choices and partly the product of discriminatory arrangements not in science but in society at large. As pointed out by Van den Beseelaar and Sandström (van den Besselaar and Sandström 2016) if this analysis is right, we are back from gender bias to performance differences. But others like Van der Lee and Ellemers (R. van der Lee and Ellemers 2018) link these claims to what they call "individual merit ideology".

172. Although the argument that gender gaps in career advacement are mainly explained by differential performance and previous career choices may be analytically appealing, there is also evidence that the higher we go in the academic hierarchy, the more difficult is to disentangle performance and career differences from other factors that impede women's entry into the most elitist ranks. In a interesting recent paper, Treviño et al. (Treviño et al. 2018) analyse differential appointments by gender to the rank of named professorships in a sample of over 500 management professors at tier one of American research universities. They find that after controlling for research performance and other factors associated with advancement, women are less likely to be awarded named professorships. Women had to score much higher in performance to receive an endowed chair; evidence on previous mobility weighs more heavily on women's chances to secure a named chair than on men's[8]. Moreover, the adverse gender effects were more acute when the endowed chair was awarded to an internal candidate than when they were awarded to external ones; this finding is interesting and consistent with research arguing that women require networks outside their own organisation to advance.

[8] In a qualititative case study of junior faculty research at large R1 US universities, Rivera (Rivera 2017) reports that committee members assumed that women whose partners held academics or high status jobs were not movable, whereas the same did not apply for male applicants whose relationship status were unfrequently discussed, or whose partners were assumed to be movable.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

173. In sum, the literature has yielded heterogeneous results and whereas some show clear effecs of the various potential sources of bias, other find only moderate or weak effects; of course, the lack of common definitions, samples, methods etc. is the most likely explanation of such heterogeneity. In order to test causality of gender or institutional bias, the literature has increasingly introduced measures of performance into the analyses and it is slowly but more and more adopting experimental approaches based on randomised control trials. But the controversy is still evident. As noted by Van der Lee and Ellemers (R. van der Lee and Ellemers 2018) the mixed nature of findings together with the methodological criticisms have led to some reluctance in accepting evidence on the existence of gender bias in academia, when the reality is that different studies focus on different situations and produce different outcomes. Yet in their paper, resistence to evidence against gender bias is made equivalent to resistance to evidence in gender inequality and disparities, and the two are different processes.

2.4 Research grants' impact on productivity and career advancement.

174. To what extent is competitive research funding a key factor in the advancement of research and academic careers by gender and compared to those lacking such funding? Do the evaluation committees for hiring and promotion include the number, size or reputation of grants as part of the decision criteria? The main questions guiding the literature review in this section are: Does peer review in RFOs fulfill its objective of identifying the scientists with the best potential? How much of the advancement in career can be explained by having research grants compared to not having them or to showing high publication performance? Are there differences by gender?

175. The advancement in careers and their explanatory factors is something strongly affected by the institutional context, which makes organisational level studies pertinent. For instance, selection and recruitment in tournaments models are not only based on the qualifications and achievements of applicant but also on who is competing with whom and for what position. In fact, we should distinguish between organisations that give importance to the researchers' capacity of getting competitive funding in terms of career rewards from those which do not. Traditionally, teaching had a more prominent role than research in the valuation of academics. The balance is also affected by the research funding regime in each country.

176. Women **continue to be underrepresented in the higher academic ranks**, and gender differences in promotion to tenure exists after controlling for demographic and productivity characteristics (Ginther and Hayes 2003) . Access to funding is one of the keys to success in academic careers, both for women and for men, providing essential support for research and publications. Indeed, the role of competitive funding is increasing in many European national settings and success in the competition for research funding is now often used as a measure of scientific excellence at both individual and institutional level (Husu and de Cheveigné 2010). As noted by van Arensbergen et al. (van

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Arensbergen, van der Weijden, and van den Besselaar 2014) within recruitment, evaluation and promotion procedures, acquirement of external funding -and especially prestigious career grants- form an important criterion. In their view, skills and dispositions needed to acquire grants have been included in the academic *habitus*. Moreover, qualitative research has documented that competition in academia (especially regarding job security) leads to a high publication and grant pressure which are perceived as too high by early career researchers (Waaijer et al. 2018).

2.4.1 Impact of grants on research performance

177. The proportion of women authors in high profile research journals is substantially lower than the proportion of women receiving competitive grants (Y. A. Shen et al. 2018). So even in the case of positive effect of grants on individual's impact, we should not expect that the speed of both processes is aligned. There are a number of analyses of the **impacts of external funding on individual research performance**. Their focus is often the effects of grants on publications and collaborations, but to a much lesser extent on career. For instance, Jacob and Lefgren (Jacob and Lefgren 2011) find evidence of a limited positive impact of U.S. NIH grants on publications and citations, which they argue may be due to the access of researchers in general to other funding sources; Lee and Bozeman (S. Lee and Bozeman 2005) report a significant effect of receiving a grant on publication, and Gaughan and Bozeman (Gaughan and Bozeman 2002) find such positive effects for public grants but not for private ones.

178. One indirect effect of monitoring productivity after getting grants is to assess the "quality" (predictive validity) of the peer review processes. A few studies have investigated the **quality of peer review for the selection** of young scientists, but they have seldom included an analysis of the subsequent publication output of the applicants; an exception was Melin and Danell (Melin and Danell 2006); they examined the peer review process for the Individual Grant for the Advancement of Research Leaders (INGVAR) of the Swedish Foundation for Strategic Research. Their analyses of the "publication histories" of 40 applicants show – in contrast to the results Bornmann et al. (2008)- only slight mean differences in scientific productivity between approved and rejected applicants. Similar results were reported by van den Besselaar and Leydesdorff (van den Besselaar and Leydesdorff 2009), who evaluated the peer review process of the Council for Social Scientific Research of the Netherlands Organisation for Scientific Research. However, the results of this study focused on highly selected applicants, i.e. besides the approved, only the best rejected applicants. As noted by Bornmann et al. [(Bornmann, Mutz, and Daniel 2008), (Bornmann, Wallon, and Ledin 2008)] large performance differences between accepted and rejected applicants would have been a surprise for these samples.

179. Through the analysis of the Long-term fellowships and the Young Investigators programme of the EMBO, Bornmann et al. (Bornmann, Wallon, and Ledin 2008) investigated the **association between the selection decisions and future scientific**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

performance of the applicants. They found that, although for a part of the applicants, the selection committee did not correctly estimate the applicant's future performance (underestimation of rejected candidates was much higher than overestimation of approved candidates, which is expectable in calls with low success rates), there was a statistically significant association between selection and performance when quantity and impact were used as criteria of achievement.

180. Therefore, the review of the literature does not allow for strong statements about the effects of grants on performance. On the one hand, it is important to distinguish quantity of publications from impact; for instance, based on data on applicants to various Danish and Norwegian grant schemes Langfeld et al. (Langfeldt, Bloch, and Sivertsen 2015), showed higher increases in the number of publications for grant recipients than for rejected applicants, while increases in mean normalised citation rates were not significantly higher for the successful applicants; the grants seemed to have increased productivity, e.g. by helping PIs to add staff to their research teams, but not to have influenced the importance of the research as measured by average citations. However, along with increases in the number of publications also came a greater increase in the number of highly cited papers for grant recipients than for rejected applicants[9].

181. On the other hand, in this type of analysis, one has to acknowledge that it cannot be ruled out that the applicants who received research funding may have published more subsequent to application precisely because they received funding and not necessarily because the committee made the right choice about who received funding. There is circularity to this issue that should be considered in studies investigating grant peer reviews and grants' impact on performance. The endogeneity of research funding and research output and the fact that causality is likely to go both ways have been noted by Sandström (Sandström 2009b) among others.

182. To control in the statistical analysis for the influence of funding on subsequent performance, information is needed on alternative funding of the rejected research and rejected applicants. The applicants' complex landscape of multiple projects and grants makes it difficult to isolate the impact of a single grant. This poses limitations to the use of bibliometrics for measuring the impact of smaller grants (Langfeldt, Bloch, and Sivertsen 2015).

2.4.2 Indirect effects of funding on increasing collaboration networks

183. Assessing the impact of research funding on the performance of researchers is a difficult task, as their grants' output is influenced by a series of factors including seniority, gender, collaboration and geographical location of their host institution.

[9] The literature typically categorises rejection decisions by selection committees as type I errors (falsely drawn approval) or type II errors (falsely drawn rejection). With type I error the committee overestimates the potential of an actually approved candidate, and with type II errors the committee underestimates the potential of an actually rejected candidate. In the empirical literature, the errors are typically detected by studying performance approved and rejected candidates at a later point of time.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

184. Pina et al. (Pina et al. 2019) have analysed publication and citation output of ERC Starting and Advanced grantees and their collaboration networks in a cohort of 355 grantees from life sciences in the years 2007-2009. They show that while senior grantees had overall greater publication output, junior grantees had significantly a greater increase in their number of publications after the award. **The collaboration networks increased** for all grantees, although more for juniors. Differences in performance of grantees before and after the award were not related to gender, although male junior grantees had more publications than female grantees before and after the grant award. Junior grantees located in lower research-performing countries published less and had less diverse collaboration networks than their peers located in higher-performing research countries. The authors concluded that the research environment had greater influence on post-grant award publications than gender, especially for juniors.

185. When analysing the impact of grants on productivity the **institutional level** is relevant as well. In a study of the NIH research project grant funding to institutions, Wahls (Wahls 2018) reports that scientific output is not a linear function of amounts of federal grant support to individual investigators; as funding per investigator increases, beyond a certain point, productivity decreases. Interestingly, prestigious institutions had on average 65% higher grant application success and 50% larger award sizes whereas less prestigious institutions produced 65% more publications and had a 35% higher scientific impact per dollar funding, suggesting that implicit biases and social prestige mechanisms have a powerful impact on grant funding allocation.

2.4.3 Impact of grants on career advancement

186. The literature on the **impact of grants on careers** is limited. Evidence gathered by the Euroepan Commission (European Commission 2014) reveals that Marie Curie fellowships (MC) have definite beneficial impacts on researchers' career prospects. A comparative analysis of Marie Curie fellows and researchers from a control group (CG) shows that MC fellows achieve professorship more frequently than others and are more likely than the CG of leading a team of researchers. The study confirmed the existence of a gender gap in research, which is apparent in many areas such as: less mobility, difficulties in reconciling work and family life, less salary and career progress, etc. Marie Curie Fellowships appear to have some impact in closing the gender gap, for instance with respect to: chances of being appointed associate professor, full professor or principal investigator, resuming an interrupted career, number and quality of publications, and access to international research funds.

187. Positive career effects were also reported by Bloch et al. (Bloch, Graversen, and Pedersen 2014) in their study of applications and awards of grants to independent research projects in Denmark during 2001-2007. They used propensity score matching of successful and rejected applicants and found that the probability of obtaining a full professorship for grant recipients was almost double that for rejected applicants (16% versus 9%); women had

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

higher probability of becoming a full professor (p.91). As the authors rightly point out, whether career advancement comes directly from the receipt of grants, indirectly through increased publication and stronger networks, or via other factors, like recognition or reputation, is difficult to determine.

188. However, in a report of the EMBO long-term fellowships, Klaus and Alamo (Klaus and Alamo 2018) have evaluated the accuracy of the evaluation system firstly, by comparing the level of career progression of the candidates in 2017 with the original award decisions ten years before, and secondly, by exploring the relationship of career progression with indicators derived from the information available to evaluators at the time of application. The results suggest that the peer review system is not substantially better than random selection once a pre-selection of the most promising applicants is performed; findings also show that among other potential sources of uncertainty, the information available at the time of application is not sufficiently predictive of career progression. What they did find was clear differences in career progression between men and women. However, as regards predictive validity of peer review it is also possible to find in the literature evidence in support of such validity, like the one reported by Bornmann and Daniel 2005 with respect to the B.I.F fellowships, although in that paper they only studied successful applicants. In the area of medicine, Holliday et al. (Holliday et al. 2014) studied 82 oncology departments and found that men were more likely to be senior faculty and receive NIH funding. After stratifying for rank, these differences were largely non significant. They also found a systematic gender association, with fewer women achieving senior faculty rank. However, women achieving seniority had productivity metrics comparable to those of male counterparts.

189. **Time to promotion** is a very informative feature of career advancement and it has been found to be longer for female than for male academics in a number of studies (for instance at Canadian universities (P. Stewart, Ornstein, and Drakich 2009)) even controlling for personal attributes such as career breaks and publication history (Ward 2001), although no significant differences have been reported for other countries (Sanz-Menéndez, Cruz-Castro, and Alva 2013).

190. In a longitudinal study of applicants to an early career program of a Social Science Council in the Netherlands, Van den Besselaar and Sandström (van den Besselaar and Sandström 2016) investigate whether gender differences (small or non-existent at the time of application) emerge at a later stage in the career and whether gender career differences have occurred, and if so, if they can be explained by performance differences. They find that, after a decade, the productivity of male researchers has grown faster than of female researchers but the citation impact indicators remain about equal. Furthermore, they observe that although performance data explain to some extent faster careers of male researchers, gender is an important determinant too. They conclude that the process of academic hiring still remains biased.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

191. Dynamics of **time to promotion and career attrition** by gender are very much affected by the national and institutional context and this is why we find differences in the degree of advancement towards greater equality. U.S. research universities constitute perhaps the institutional context for which differences between genders in faculty retention are less and less acute [(Kaminski and Geisler 2012), (Box-Steffensmeier et al. 2015)] although in promotion to full-professor important differences persist (Box-Steffensmeier et al. 2015) .

192. There are very few **longitudinal studies of research funding**. An exception is the study of NIH funding longevity by gender (Hechtman et al. 2018) in which they study funding trajectories over time (n=34,770). A survival analysis demonstrated only slightly lower funding longevity for women and contradicted the common assumption that women experience accelerated attrition compared with men across all career stages. The findings suggest a need to explore women's underrepresentation among initial NIH grantees, as well as their lower rates of new and renewal application submissions.

2.4.4 Impact of grants on the autonomy of researchers versus the management and external elites

193. There is a substantial literature on the changing dynamics of research funding and authority relations within research organistations over the last decades [(Whitley 2011), (Gläser and Velarde 2018), (Gläser and Laudel 2016)] a review of which would exceed the scope of this report. However, it is important to highlight that universities and research institutes can be characterised as autonomous organisations in which the decisions regarding what lines of research to pursue as well as the control of the research process lies mainly with the researchers and research groups, who enjoy a high degree of strategic autonomy (Scott 1965). Scientific standards are established externally by the professional or scientific community and hiring is structured around scientific committees dominated by and even exclusively composed of researchers. Career rewards and promotions are allocated mainly on the basis of the perceived contribution of individual candidates to the field in committee-based processes.

194. The dominance of the professional authority in universities and research institutes is further enhanced if administrative management control over collective resources is limited. This limitation may have two sources: firstly, the formal delegation of the control of scientific infrastructure and other collectives to scientific directors and researchers; secondly, the existence of individual discretion over funding resources obtained externally by researchers themselves, from individual projects or contracts.

195. To study the impact of grants on careers, it is also important to distinguish between types of funding (Cruz-Castro and Sanz-Menéndez 2018). In general, increases in individual project funding for individuals working in universities and research institutues will reinforce the position of researchers *vis à vis* managers and the autonomy of the former. Additionally, when individual researchers succeed in obtaining external funding they may not only maintain or increase their autonomy from managers but also gain leverage to

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

negotiate additional resources such as academic positions or institutional funding to support their agendas, feeding cumulative advantage processes. But even in this type of research settings, where structure places considerable control and discretion in the hands of researchers, larger shares of external competitive individual project funding will also reinforce the foundations of the "republic of science", including the reputational competition, with the consequent shift in the influence of scientific elites over subsequent individual rewards. The extent to which increases in individual project funding will produce careers which are more autonomous, will also depend on who effectively controls the tenure system typical of this type of research organisations.

196. Different forms of individual project funding will also shape the sharing of authority: while curiosity-driven project funding will reinforce the authority of researchers mainly because of asymmetries of information, priority set and oriented public research funding or industry funding will provide external actors with more influence over research agendas.

2.5 Gender equality policies and organisational interventions.

197. Most policy recommendations in this area tend to **focus either on demand** (increasing organisational transparency, formalisation and accountability to reduce bias) **or supply** (interventions to encourage women to apply). However, the supply-side approach (training, networking, mentorship for women) has important limitations if not matched with organisational practices. The review of the literature suggests that effective policies should consider both. A further distinction can be made between policies aimed a changing the process of application and evaluation and those aimed directly at the outcome. The majority of the policies fall on the first type although often the policy and academic discourse about diagnosis and policy solutions is mixed with claims about desired outcomes. As in the rest of the topics, there are problems of clarity in the concepts and definitions surrounding policy action. A recent briefing paper from the project Gender Action states that there is a tendency to restrict gender equality to female participation in research or to ensuring work-life balance, but as stressed by Kalpazidou Schmidt and Cacace (Kalpazidou Schmidt and Cacace 2017) it is important to understand gender equality policies as complex processes, the outcomes of which depend on the interaction of a multiplicity of variables in dynamic contexts leading to the adoption of multidimensional, complexity frames of reference for impact assessment.

198. Gender Equality (GE) policies are perceived as necessary in order to create more opportunities for upward career mobility for women. At the same time, both men and women sometimes argue against these policies due to issues of reverse discrimination and quality loss that they raise, leading sometimes to a dilemma between gender equality on the one hand, and merit and individual advancement on the other. Van der Brink and colleagues [(Van den Brink and Stobbe 2014), (Van den Brink and Benschop 2014)] argue that the "support paradox" provides a discursive tool to counter this dilemma that finds its roots in a strong belief in the meritocracy and blindness for the genderedness of the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

meritocracy and academic careers. Maybe here is it important to bear in mind that biases are a threat to meritocracy too (European Commission 2017).

2.5.1 Design and content of the instruments (calls)

199. From the point of view of the funders, **the design of the calls** may also be an area of intervention. An example is provided by Escobar Alvarez et al. (Alvarez et al. 2019): the Doris Duke Clinical Scientist Development Award (CSDA); in 2016, the request for applications and review guideline documents of this award were revised to clearly articulate attributes held by successful applicants and evaluation criteria that used objective, non-gendered language. These materials were revised to minimise use of words that were thought to be implicitly associated with traditionally masculine traits. For example, the phrase "leadership potential" was changed to "promise to make significant contributions", "importance" to "influence", "innovation" to "originality", and "creativity" to "inventiveness." Magua et al. (Magua et al. 2017) have confirmed these types of gendered associations of certain words in the NIH peer review.

200. This reported case is quite interesting because the funder used the application to encourage institutions to consider gender equity in applicant salaries and raised awareness about the issue. Department chairs were asked to provide the applicant's salary quartile range relative to those at the same faculty rank in the department. Applicants' salary quartiles showed a gender gap between women and men who entered the competition. This question was not used in the application review. Additionally, guidelines were changed to ask recommenders to address the applicants objective research records and avoid references to personal circumstances like gender, age, work-life balance or roles outside the professional setting.

2.5.2 Addressing family factors

201. Parental eligibility **clock stopping policies** have become widespread in RFO calls. These policies can be classified into two main types: policies that only apply to mothers, and policies that are gender-neutral and apply to both sexes. One has to be cautious when advocating for this type of gender-neutral policies because there is evidence of unintended effects. For instance, data gathered at 50 economics departments in a 15 years period in the U.S. shows that gender-neutral tenure clock-stopping policies substantially reduced female tenure rates while substantially increased male tenure rates (Antecol, Bedard, and Stearns 2018); the primary mechanism driving these results appeared to be that men published more in top 5 journals after the implementation of the policies whereas women did not; however, the policy did not reduce the fraction of women who eventually earned the tenure, but increased duration. In 2010, almost a decade ago, the ERC implemented some measures-such as increasing the window of grant eligibility for applicants with children-after which the number of female applicants increased, as did the number of male applicants, so the gap did not narrow (at that point female application rates were 29% for early career grants and 15% for advanced grants). These findings suggest that gender

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

neutral clock stopping policies may not adequately account for the gender-productivity losses associated with having children.

2.5.3 Awarness and information, education and learning from evidence

202. Behavioural sciences have paid much attention to interventions in organisations to mitigate bias but have also signalled the limited empirical evidence of what works in systemic interventions to promote the advancement of a non dominant group at the top of the organisational hierarchy (Vinkenburg 2017). Many diversity interventions assume a link between **educating participants about the existence of bias** and a reduction in biased attitudes and behaviour. Allegedly, scientists are trained to evaluate and interpret evidence without bias or subjectivity; to the extent that research illustrating gender bias in STEM is viewed as convincing, the culture of science can begin to address the bias. Handley et al. (Handley et al. 2015) address precisely the question of whether men and women are equally receptive to this type of evidence by means of a randomised double blind experiment comparing samples of the general public, STEM and non-STEM faculty. Results across experiments showed that men evaluate the gender-bias research less favourably than women, tend to be more vocal about it than women, and, of concern, that this gender difference was especially prominent among STEM faculty. Moreover, Moss-Racussin et al. [(Moss-Racusin, Molenda, and Cramer 2015), (Moss-Racusin et al. 2016)] demonstrate that showing evidence of bias can bring about positive outcomes as well as reactive justifications.

203. As in other parts of the literature, scholarly work about gender equality policies and intervention in organisations is mostly descriptive; however, it is possible to find some recent **experimental papers** using randomised trials which report interventions to increase gender **bias awareness** in academic departments; for instance, Devine et al. (Devine et al. 2017) studied hiring rates of new female faculty and found that the proportion of women hired by intervention departments increased by 18%. Likewise, Sekaquaptewa et al. (Sekaquaptewa et al. 2019) showed that attendance to faculty recruitment workshops led to greater belief in evidence-based descriptions of gender bias, although their approach was not experimental. Given the mixed nature of evidence, we should acknowledge the limitations of bias awareness as a prejudice-reduction tool.

204. Well-oriented interventions should pay attention to the **underlying cause of discrimination**, where it exists. Bohren et al. (Bohren, Imas, and Rosenberg 2019) distinguish between belief-based discrimination and preference-based discrimination (taste-based discrimination in Becker's terms). When initial beliefs favour men, a woman needs to produce higher quality output in order to overcome disparity in beliefs and receive a similar evaluation as a man; according to these authors, this type of discrimination may be mitigated with information about past performance, especially if the evaluators are aware of the women's need to meet a higher standard, in a kind of compensating mechanism. In contrast, if discrimination is caused by a preference against women, then a

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

woman who receives a similar sequence of evaluations will continue to face discrimination in future periods, and discrimination will persist even when quality is perfectly observable. In the paper, it is shown that belief-based partiality decreases when judgement criteria becomes more objective, whereas it remains constant if the partiality is preference-based. What this work highlights is the importance to take into account the dynamic nature of discrimination, and the possibility that learning takes place in successive evaluations.

2.5.4 Addressing gender stereotypes: what organisations can do?

205. As we have mentioned, the literature points to some **ways to deter gender bias** arising from descriptive gender stereotypes (Heilman 2012). The first relates to precluding negative expectations; perceptions of lack of fit can be lessened if gender stereotypes are not activated in a given situation. Organisations can influence activation due to structural features of the workplace. The proportional representation of women in work settings is one such contextual feature. Gender is made salient by numerical scarcity and there is evidence that increased proportional representation of women in the applicant pool favourably affects women's performance evaluations (Sackett, DuBois, and Noe 1991). The second relates to reducing ambiguity in the evaluation process; there are many ways in which organisations can act to lessen or even eliminate ambiguity: collecting and providing evaluators comprehensive job-relevant information can forestall the tendency to use expectations to ''fill in the blanks''; making the criteria for evaluation concrete, and making sure they are used in the same way no matter who is being evaluated, can avoid the possibility of using different criteria for different people (Heilman 2012).

206. Thirdly, obtaining individual performance information when work is done in a group can prevent the attributional rationalisation that causes women to be denied credit for their successes (Heilman 2012). Finally, organisations may reduce bias by increasing the motivation to be accurate through accountability (a mechanism already highlighted by Reskin (Reskin 2003)). It should be noted that making evaluators accountable does not always have the effect of decreasing bias in evaluative judgments; this only occurs if motivations to be accurate due to accountability is compatible with social approval. If organisational culture supports gender bias, biased evaluative judgments may still result; in those cases the motivator is not accuracy but the need for approval of the rest of the panel or the organisation at large.

207. Interventions to prevent gender bias should be adapted and tested in the context of grant peer review to determine if they will have an impact. Prescription-based gender bias is less responsive to contextual interventions or organisational efforts to reduce its effects through the provision of information.

2.5.5 Gender quotas

208. **Gender quotas** are one of the most commonly used policy measures to promote gender equality and in some countries, they are legally mandatory, but as we have seen in section 2.3 the empirical evidence of their effects is not clear-cut. In a recent paper, Voorspoels

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

and Bleijenbergh (Voorspoels and Bleijenbergh 2019) highlight the importance of the implementation practices of the quotas, arguing that a strictly procedural sense-making of gender quotas could undermine women´s representation in university decision making.

209. The European Commission committed to reaching 40% female participation in its advisory structures for Horizon 2020, the current research funding programme; but statistics collected by the European Research Council suggests that quotas are no magic wand to bring about gender equality in research and academia; aggregating data for 2008-2012 the ERC found no correlation between the percentage of women on its evaluation panels and the success rates of female applicants (Vernos 2013). Another concern in ERC funding has been the female slightly lower success rate, but there was no correlation between success rates of female applicants and the gender balance of evaluation panels (Van den Beselaar, forthcomming). Some argue that quotas would place greater demand on the small pool of female scientists who serve on those panels-possibly enough to hinder their career progress (Zinovyeva and Bagues 2015).

210. The mix nature of the evidence presented in section 2.3.2 calls at least for caution when drawing policy implications. As noted by Broder (Broder 1993) institutions, including NSF, have tried to solicit female reviewers for female proposals to avoid potential male bias against women. Her work however, suggests that this type of policy might have lowered women's opportunities rather than raised them and may account for some part of the underrepresentation of women in the senior ranks of the economics profession in U.S..

2.5.6 Gender diversity management

211. **Gender diversity** (Fortunato et al. 2018) has been claimed to lead to **better science** (Mathias Wullum Nielsen et al. 2017). Many policies are oriented to increasing gender diversity in scientific organisations, yet its impact still requires a lot of research and evidence. In relation to academic hiring and based on a dataset combining a survey of department chairs and their performance indicators, Su et al. (Su, Johnson, and Bozeman 2015) investigate the organisational determinants of gender diversity strategies in the STEM fields and find that women chairs prove less likely to pursue a gender diversity strategy, and that more female faculty members hardly increase the likelihood of adopting such a strategy. The authors argue that the findings require care in interpretation because in cases where there are more women, the perceived need for adding women may be lessened. As such, gender diversity strategy may be compensatory in nature. In any case, the study underscores the need for richer theories about recruitment of women STEM faculty. Others like Stweart and Valian (A. J. Stewart and Valian 2018) argue in favour of what the call “an inclusive academy” developing the argument that excellence and diversity –especially gender- are not contradictory.

212. Should **female leadership or gender composition of the teams** be a criterion for grant allocation? Advocates of this type of measures argue that only research teams that show diversity should be funded (O’Connor and Fauve-Chamoux 2016) and that if committees

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

(especially transnational ones) use criteria such as location or country to balance success outcomes, then why should not gender balance be a policy rationale just as valid? But this is a political question, not a scientific one.

213. The need for better theories also applies to the implementation of **affirmative action** and gender equality policies aimed at increasing the representation of women in academic organisations; very broadly, we can find two opposing theories; whereas social contact theory predicts that social prejudice and stereotypes are less likely to flourish when cross-group interactions are high, competition theory links increases in the size of a minority group to increases in the levels of intergroup hostility and conflict. So the question is whether **changes in group proportions** lead to integrative versus competitive dynamics. Although rather outdated, evidence from the 80s showed that, for U.S. academic departments, as the proportion of women in a department grew, turnover among women also increased, confirming the prediction that increases in the relative size of a minority could result in increased intergroup competition increasing the likehood of their leaving the group (Tolbert et al. 1995). This work also suggested, however, that as the proportion of women faculty reached a threshold of about 35%-40%, female turnover began to decline.

214. But the academic career has several stages and the effects of diversity might no be homogenous across different stages. Although the public discourse in science education policy has very much stressed the "leaky pipeline" problem, less attention has focused on experimentally testing solutions to the problem. Dasgupta et al. (Dasgupta, Scircle, and Hunsinger 2015) **report an experiment investigating one solution**: the creation of microenvironments (small groups) in engineering with varying proportions of women to test which environment increased motivation and participation and whether outcomes depend on women's academic stage. Their data suggest that creating small groups with high proportions of women in otherwise male-dominated fields is a way to keep women engaged and aspiring to engineering careers.

215. Some mentoring programmes rest on a similar rationale. For instance, Leenders et al. (Leenders, Bleijenbergh, and Van den Brink 2019) conduct interview and focus groups based qualitative research exploring how the process of transformational change through mentoring programs works to change gendered organisational norms and work practices: 1) by discussing and reflecting upon gendered organisational norms and work practices, 2) by creating new narratives of the "ideal academic", and 3) by experimenting with new work practices. The case study is interesting although the particular university is a 'best practice' one with the highest share of female professors.

216. Experiments have also been developed on this issue. For example, Dennehy and Dasgupta (Dennehy and Dasgupta 2017) addressed the issue providing evidence from a multiyear field experiment demonstrating that female (but not male) mentors protected women's belonging in engineering, self-efficacy, motivation, retention in engineering majors, and postcollege engineering aspirations. Counter to common assumptions, better engineering

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

grades were not associated with more retention or career aspirations in engineering in the first year of college. Notably, increased belonging and self-efficacy were significantly associated with more retention and career aspirations. The benefits of peer mentoring endured long after the intervention had ended, and continued for the first 2 years of college -the window of greatest attrition from science, technology, engineering, and mathematics (STEM) majors. Thus, same-gender peer mentoring for a short period during developmental transition points promotes women's success and retention in engineering, yielding benefits over time.

2.5.7 Changing rules and institutions

217. As this review has shown, some of the underlying dynamics in women applying less and getting fewer research grants are structural. Many of the gender **differences in application behaviour are rooted in** higher education institutions' **employment practices**. Underrepresentation of women in senior positions is a persistent problem in universities worldwide. The opacity of the decision making processes is one of the main problems that scientific institutions face when implementing gender equality policies. The introduction of tenure track in systems previously lacking it may be a strategy to combat this situation assuming that this type of systems are guided by a set of explicit and transparent criteria. In a case study of a Dutch university, Bakker and Jacobs (Bakker and Jacobs 2016) described how chances to being promoted to higher levels were already fairly equal before the introduction of a tenure track system, and improved more for women than for men afterwards. However, they conclude that since promotion rates are small compared to the total amount of staff, current distributions of men and women will exhibit a considerable degree of inertia, unless additional affirmative action is taken; and we could add: or unless promotion rates are increased.

218. In a similar vein, one might think that the introduction of centralised *habilitation* systems in which merit is measured via bibliometric and non bibliometric indicators may yield lower gender gaps than evaluations made at the university level where there may be more discretion and less transparency. Marini and Meschitti (Marini and Meschitti 2018) tested this hypothesis for Italian universities (where, as in Spain, tenure and promotion are obtained in a two step process) and found that among those who obtained the national accreditation and at parity of other conditions and scientific production, men had around 24% more probability to be promoted at the second step, which, in their opinion, revealed a relevant gender discrimination at the local level. In any case, this finding highlights the importance of taking the organisational level into account. The set up of a tenure system might not suffice; as stressed by Fox (Fox 2015), even in systems with long history of university tenure there is variation in the perceived degree of clarity of criteria for tenure and promotion decisions among faculty.

219. Given the limitations of peer review with respect to bias, some authors have made the case of the possible benefits of using **the lottery as an alternative** allocation mechanism that

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

would increase epistemic diversity, fairness and impartiality. Althoug he does not mention gender bias, Roumbanis (Roumbanis 2019) suggests that funding agencies could form a pool of funding applicants with certain qualification levels (a simpler form of peer review) through which proposals of poor quality are excluded and then select randomly from that pool. The most common objection for using a lottery is that it does not take qualifications into account and that the overall quality of granted proposals would be lower. Since there are hardly any funding agencies[10] that use such method, the empirical testing of such claim is yet to be tried. Apart from the alledged potential quality losses, the adoption of lottery mecahisms for grant allocation may imply a kind of moral defeat in giving up the possibility of changing organisations and institutions.

220. From a more positive perspective, Stewart and Valian (A. J. Stewart and Valian 2018) outline a few valuable general principles on how to create institutional structures that enhance desirable changes. Firstly, the need to focus on policies and practices; for them, despite the common temptation to focus policies on education and attitude change (increasing knowledge), changing peoples' beliefs and deep-rooted attitudes is extremely difficult and may even appear as associated to a kind of social engineering that is unlikable to most academics; they recommend to mitigate biases on the ground, through policies and practices that buffer their negative impact, rather than trying to change the biases themselves.

221. Secondly, efforts to produce desired institutional outcomes (e.g. fair decision making in allocation processes) should be grounded in social science knowledge and on the evidence in place about what triggers change and what does not. For instance, transparency about policy, diffusion of information, monitoring impact and demanding accountability, are essential to policy implementation, especially in areas where the key actors need to be engaged. Thirdly, although there may be cases of best practices from which to learn, policies and practices are local and cannot be imported from one organisation to another and expect that they will have the same impact. Finally, the authors call for a pragmatic "change what you can, where you can" principle; radical change is rare in organisations, but incremental and/or partial changes are not; some change, as far as it sets a process in motion, is better than none.

[10] Roumbanis (Roumbanis 2019) mentions the "Explorer Grants" awarded by the Health research Council of New Zeland, and a recent pilot study of the Volkswagen Foundation in Germany called "Experiment!" in which proposals in the same call are divided into two groups after a screening by the Foundation staff (one goes into the peer review track and the into in the the lottery), and the same number of proposals are selected for funding by each method; the objective is to evaluate the outcome of this trial in the future.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

3 IMPLICATIONS FOR THE PROJECT: FINAL REMARKS

222. **Individuals in organisations.** We believe it has been made clear that we cannot address the issues of discrimination and bias if we only analyse the characteristics of individuals and evaluators. We also need to know the features of the organisations, their processes and practices in depth**.**

223. To compare the differences, the segmentation or the discrimination in the allocation of grants to female researchers requires incorporating the study of the granting organisations and their practices, as well as the existing structures of opportunities often derived from previous segregation.

224. Additionally, it is not possible to draw general conclusions about the "disparities or discrimination" (in the allocation of grants) on the basis of a particular RFO. It is also necessary to build a research design that includes a comparative analysis of the operation of the different funding instruments.

225. **Individual actions in organisational contexts**. Who gets what (grants, jobs, etc.) is the **product of the actions of individuals** and the **organisational practices of the different RFO** that constrain and are circumvented by individuals to varying degrees; therefore, we must design our studies to capture variability of such practices within and across organisations and identify the covariates of this diversity (Reskin 2000b) .

226. Although the allocation of grants (and positions/jobs) responds to or depends on the attributes of the organisational processes, in general, research has focused on individual attributes of applicants and reviewers, maybe due to the greater the availability and/or quality of individual data. However, we should not be guided or driven solely by such availability; organisational and structural factors, processes and practices matter and may be determinant of the differences, segregation and discrimination.

3.1.1.1 Understanding the individual factors interacting with organisational behaviour and rules.

227. Research has not been successful in explaining why researchers'gender affects their outcomes, why science domains and ranks are segregated by sex, and why in general men get more research grants and "outearn" women[11]. As gender issues in research attracted more and more attention, scholars started to build universal explanatory arguments: gender-role socialisation, the domestic division of labour, patriarchal impulses, male

[11] Although Kahn and Ginter (Kahn and Ginther 2018) as well as Graham and Smith (Graham and Smith 2005) show that in the U.S., considerable gender pay differences in STEM resulted from women and men being in different fields and different employment sectors. Likewise, a review of the literature on gender wage gaps conducted by Blau and Kahn (Blau and Kahn 2017) concluded that traditional explantations related to occupation and industry effects continued to be very salient, in contrast to human capital explanations that have lost relevance, or in comparison to psychological attributes which, in comparasion, still explain little.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

researchers' responses to competitive threat, etc. Because most readily available data were at the researchers' level, the dominant approach concentrated on individual-level explanations.

228. **Understanding resource inequality (and potential discrimination) requires incorporating RFO and employing organisations into the analysis.** In the same way that the study of employing organisations has confirmed that inequality levels are affected by demographics, leaderhsip, formalisation of personnel practices, recruitment and evaluation methods, external pressures, or disponibility of organisational resources, we can assume that some variables and attributes of RFO and research performing organisations (RPO) are determinant to understand the nature of the outcomes. To study the impact of grants on careers, it will be also important that the GRANteD project distinguishes between types of funding (Cruz-Castro and Sanz-Menéndez 2018)

229. In any case, inequalities (and discrimination) do not come out of the blue**.** As noted by Baron (Baron 1991) it is important to analyse **the organisational context of bias**. Organisational theory provides some general useful insights. Firstly, organisational structures and procedures exhibit considerable inertia; therefore, when analysing RFO it is important to know not only whether gender equality policies are in place but also the organisational history in this respect. Organisational age and founding conditions and missions are relevant aspects in this regard. Secondly, the role of interests should not be overlooked; groups within organisations, including those that populate RFOs and the constellation of interests favouring equal treatment will affect the rate of organisational change. Thirdly, although subject to intertia and path-dependency, organisations respond to environmental pressures and normative isomorphism and they imitate and learn from others in the same organisational field. Finally, it is important to bear in mind the distinction between the formal organisational structures and policies on the one hand, and real practices on the other hand; the relation between the two often takes the form of loose coupling or even decoupling.

230. **Inequalities in grant success and career advancement do not just happen**; they occur through the acts and the failures to act by the people who run and work for organisations. In this context the challenge is to formulate **empirically realistic accounts of how grants are allocated in a cross-sectional study of RFO and how and why careers advance in a cross-sectional study of research organisations**.

231. Ideally, we should develop a research design that allows us to address the existence of segregation or discrimination as an organisational issue, in an appropriate manner. We could envisage/design an evaluation tool (method) to be implemented by two or more agencies (RFOs), and we could introduce the variance in the individual or in the practices.

232. Additionally, we could observe how individual research funding (or its absence) has career effects (or not) and whether and how this has changed over time.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

4 REFERENCES

Abbott, Andrew. 2016. *Processual Sociology*. Chicago ; London: The University of Chicago Press.

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Alessandro Caprasecca. 2009. 'Gender Differences in Research Productivity: A Bibliometric Analysis of the Italian Academic System'. *Scientometrics* 79 (3): 517–39. https://doi.org/10.1007/s11192-007-2046-8.

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Flavia Di Costa. 2018. 'The Effects of Gender, Age and Academic Rank on Research Diversification'. *Scientometrics* 114 (2): 373–87. https://doi.org/10.1007/s11192-017-2529-1.

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Gianluca Murgia. 2013. 'Gender Differences in Research Collaboration'. *Journal of Informetrics* 7 (4): 811–822. https://doi.org/10.1016/j.joi.2013.07.002.

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Francesco Rosati. 2016. 'Gender Bias in Academic Recruitment'. *Scientometrics* 106 (1): 119–41. https://doi.org/10.1007/s11192-015-1783-3.

Addis, Elisabetta. 2010. 'Scientific Excellence'. http://www.eif.gov.cy/mlsi/dl/genderequality.nsf/0/ED761C57A8EAF975C22580F9002C13C0/$file/Scientific%20excellence.pdf.

———. n.d. 'Gender and Scientific Excellence'. C:\Users\hoo\Documents\Citavi 5\Projects\granted_export\Citavi Attachments\Addis - Gender and Scientific Excellence.pdf.

Aguinis, Herman, Young Hun Ji, and Harry Joo. 2018. 'Gender Productivity Gap among Star Performers in STEM and Other Scientific Fields.' *Journal of Applied Psychology* 103 (12): 1283–1306. https://doi.org/10.1037/apl0000331.

Alvarez, Sindy N Escobar, Reshma Jagsi, Stephanie B Abbuhl, Carole J Lee, and Elizabeth R Myers. 2019. 'Promoting Gender Equity in Grant Making: What Can a Funder Do?' *The Lancet* 393 (10171): e9–11. https://doi.org/10.1016/S0140-6736(19)30211-9.

Antecol, Heather, Kelly Bedard, and Jenna Stearns. 2018. 'Equal but Inequitable: Who Benefits from Gender-Neutral Tenure Clock Stopping Policies?' *American Economic Review* 108 (9): 2420–41. https://doi.org/10.1257/aer.20160613.

Arensbergen, Pleun van, Inge van der Weijden, and Peter van den Besselaar. 2012. 'Gender Differences in Scientific Productivity: A Persisting Phenomenon?' *Scientometrics* 93 (3): 857–68. https://doi.org/10.1007/s11192-012-0712-y.

———. 2014. 'Different Views on Scholarly Talent: What Are the Talents We Are Looking for in Science?' *Research Evaluation* 23 (4): 273–84. https://doi.org/10.1093/reseval/rvu015.

Arrow, Kenneth. 1973. 'The Theory of Discrimination'. In *Discrimination in Labor Markets*, edited by O Ashenfelter and A. Rees. Princeton, NJ: Princeton University Press.

Astegiano, Julia, Esther Sebastián-González, and Camila de Toledo Castanho. 2019. 'Unravelling the Gender Productivity Gap in Science: A Meta-Analytical Review'. *Royal Society Open Science* 6 (6): 181566. https://doi.org/10.1098/rsos.181566.

Auspurg, Katrin, and Thomas Hinz. 2010. 'Antragsaktivitäten und Förderchancen von Wissenschaftlerinnen bei Einzelanträgen auf DFG-Einzelförderung im Zeitraum 2005-2008'. C:\Users\hoo\Documents\Citavi 5\Projects\granted_export\Citavi Attachments\Auspurg, Hinz 2010 - Antragsaktivitäten und Förderchancen von Wissenschaftlerinnen.pdf.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Bagues, Manuel, Mauro Sylos-Labini, and Natalia Zinovyeva. 2017. 'Does the Gender Composition of Scientific Committees Matter?' *American Economic Review* 107 (4): 1207–38. https://doi.org/10.1257/aer.20151211.

Bakker, Martha M., and Maarten H. Jacobs. 2016. 'Tenure Track Policy Increases Representation of Women in Senior Academic Positions, but Is Insufficient to Achieve Gender Balance'. *PLoS ONE* 11 (9): 1–15. https://doi.org/10.1371/journal.pone.0163376.

Baron, James N. 1991. 'Organizational Evidence of Ascription in Labor Markets'. In *New Approaches to Economic and Social Analysis of Discrimination*, edited by Richard R. Cornwall and Phanindra V. Wunnava, 113–43. Westport, Conn: Praeger Publishers.

Beaudry, Catherine, and Vincent Larivière. 2016. 'Which Gender Gap? Factors Affecting Researchers' Scientific Impact in Science and Medicine'. *Research Policy* 45 (9): 1790–1817. https://doi.org/10.1016/j.respol.2016.05.009.

Becker, Gary S. 1957. *Human Capital*. https://www.press.uchicago.edu/ucp/books/book/chicago/H/bo3684031.html.

Bertrand, Marianne, and Esther Duflo. 2016. 'Field Experiments on Discrimination'. w22014. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w22014.

Besselaar, Peter van den, and Loet Leydesdorff. 2009. 'Past Performance, Peer Review and Project Selection: A Case Study in the Social and Behavioral Sciences'. *Research Evaluation* 18 (4): 273–88. https://doi.org/10.3152/095820209X475360.

Besselaar, Peter van den, and Ulf Sandström. 2016. 'Gender Differences in Research Performance and Its Impact on Careers: A Longitudinal Case Study'. *Scientometrics* 106 (1): 143–62. https://doi.org/10.1007/s11192-015-1775-3.

———. 2017. 'Vicious Circles of Gender Bias, Lower Positions, and Lower Performance: Gender Differences in Scholarly Productivity and Impact'. *PLOS ONE* 12 (8): e0183301. https://doi.org/10.1371/journal.pone.0183301.

Besselaar, Peter van den, Ulf Sandström, and Hélène Schiffbaenker. 2018. 'Studying Grant Decision-Making: A Linguistic Analysis of Review Reports'. *Scientometrics* 117 (1): 313–29. https://doi.org/10.1007/s11192-018-2848-x.

Bielby, William T. 2000. 'Minimizing Workplace Gender and Racial Bias'. *Contemporary Sociology* 29 (1): 120–29. https://doi.org/10.2307/2654937.

Bielby, William T., and James N. Baron. 1986. 'Men and Women at Work: Sex Segregation and Statistical Discrimination'. *American Journal of Sociology* 91 (4): 759–99. https://www.jstor.org/stable/2779958.

Blake, Margaret, and Ivana La Valle. 2000. 'Who Applies for Research Funding?' Edited by Wellcome Trust. C:\Users\hoo\Documents\Citavi 5\Projects\granted_export\Citavi Attachments\Blake, La Valle 2000 - Who applies for research funding.pdf.

Blau, Francine D., Janet M. Currie, Rachel T. A. Croson, and Donna K. Ginther. 2010. 'Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial'. *The American Economic Review* 100 (2): 348–52. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/27805017.

Blau, Francine D., and Lawrence M. Kahn. 2000. 'Gender Differences in Pay'. *The Journal of Economic Perspectives* 14 (4): 75–99. https://www.jstor.org/stable/2647076.

———. 2017. 'The Gender Wage Gap: Extent, Trends, and Explanations'. *Journal of Economic Literature* 55 (3): 789–865. https://doi.org/10.1257/jel.20160995.

Bloch, Carter, Ebbe Krogh Graversen, and Heidi Skovgaard Pedersen. 2014. 'Competitive Research Grants and Their Impact on Career Performance'. *Minerva* 52 (1): 77–96. https://doi.org/10.1007/s11024-014-9247-0.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Bohren, Aislinn, Alex Oleg Imas, and Michael Rosenberg. 2019. 'The Dynamics of Discrimination: Theory and Evidence'. *American Economic Review*. https://doi.org/10.1257/aer.20171829.

Bornmann, Lutz. 2011. 'Scientific Peer Review'. *Annual Review of Information Science and Technology* 45 (1): 197–245. https://doi.org/10.1002/aris.2011.1440450112.

Bornmann, Lutz, and Hans-Dieter Daniel. 2005. 'Selection of Research Fellowship Recipients by Committee Peer Review. Reliability, Fairness and Predictive Validity of Board of Trustees' Decisions'. *Scientometrics* 63 (2): 297–320. https://doi.org/10.1007/s11192-005-0214-2.

Bornmann, Lutz, Rüdiger Mutz, and Hans-Dieter Daniel. 2007. 'Gender Differences in Grant Peer Review: A Meta-Analysis'. *Journal of Informetrics*, The Hirsch Index, 1 (3): 226–38. https://doi.org/10.1016/j.joi.2007.03.001.

———. 2008. 'How to Detect Indications of Potential Sources of Bias in Peer Review: A Generalized Latent Variable Modeling Approach Exemplified by a Gender Study'. *Journal of Informetrics* 2 (4): 280–87. https://doi.org/10.1016/j.joi.2008.09.003.

Bornmann, Lutz, Gerlind Wallon, and Anna Ledin. 2008. 'Does the Committee Peer Review Select the Best Applicants for Funding? An Investigation of the Selection Process for Two European Molecular Biology Organization Programmes'. *PLOS ONE* 3 (10): e3480. https://doi.org/10.1371/journal.pone.0003480.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and Christoph Riedl. 2016. 'Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science'. *Management Science* 62 (10): 2765–83. https://doi.org/10.1287/mnsc.2015.2285.

Box-Steffensmeier, Janet M., Raphael C. Cunha, Roumen A. Varbanov, Yee Shwen Hoh, Margaret L. Knisley, and Mary Alice Holmes. 2015. 'Survival Analysis of Faculty Retention and Promotion in the Social Sciences by Gender'. *PLOS ONE* 10 (11): e0143093. https://doi.org/10.1371/journal.pone.0143093.

Bozeman, Barry, and Elizabeth Corley. 2004. 'Scientists' Collaboration Strategies: Implications for Scientific and Technical Human Capital'. *Research Policy* 33 (4): 599–616. https://doi.org/10.1016/j.respol.2004.01.008.

Bozeman, Barry, and Mary K. Feeney. 2007. 'Toward a Useful Theory of Mentoring: A Conceptual Analysis and Critique'. *Administration & Society* 39 (6): 719–39. https://doi.org/10.1177/0095399707304119.

Brands, Raina A., and Isabel Fernandez-Mateo. 2017. 'Leaning Out: How Negative Recruitment Experiences Shape Women's Decisions to Compete for Executive Roles'. *Administrative Science Quarterly* 62 (3): 405–42. https://doi.org/10.1177/0001839216682728.

Brink, Marieke van den. 2010. *Behind the Scenes of Science: Gender Practices in the Recruitment and Selection of Professors in the Netherlands*. Amsterdam University Press.

Brink, Marieke van den, and Yvonne Benschop. 2012. 'Gender Practices in the Construction of Academic Excellence: Sheep with Five Legs'. *Organization* 19 (4): 507–24. https://doi.org/10.1177/1350508411414293.

Brink, Marieke van den, Margo Brouns, and Sietske Waslander. 2006. 'Does Excellence Have a Gender?: A National Research Study on Recruitment and Selection Procedures for Professorial Appointments in The Netherlands'. Edited by Liz Doherty. *Employee Relations* 28 (6): 523–39. https://doi.org/10.1108/01425450610704470.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Broder, Ivy E. 1993. 'Review of NSF Economics Proposals: Gender and Institutional Patterns'. *The American Economic Review* 83 (4): 964–70. https://www.jstor.org/stable/2117588.

Brooks, Chris, Evelyn M. Fenton, and James T. Walker. 2014. 'Gender and the Evaluation of Research'. *Research Policy* 43 (6): 990–1001. https://doi.org/10.1016/j.respol.2013.12.005.

Brouns, Margo. 2004. *Gender and Assessment of Scientific Quality*. Gender and Excellence in the Making. Luxembourg.

Caprile, Maria, Elisabetta Addis, Cecilia Castaño, Ineke Klinge, Marina Larios, Danièle Meulders, Jörg Müller, et al. 2012. 'Meta-Analysis of Gender and Science Research: Synthesis Report'. Brussels: European Commission Directorate-General for Research and Innovation. https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/meta-analysis-of-gender-and-science-research-synthesis-report.pdf.

Carli, Linda. L., Laila Alawa, YoonAh Lee, Bei Zhao, and Elaine Kim. 2016. 'Stereotypes About Gender and Science: Women = Scientists'. *Psychology of Women Quarterly* 40 (2016). https://doi.org/10.1177/0361684315622645.

Cech, Erin A., and Mary Blair-Loy. 2019. 'The Changing Career Trajectories of New Parents in STEM'. *Proceedings of the National Academy of Sciences* 116 (10): 4182–87. https://doi.org/10.1073/pnas.1810862116.

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. 'Women in Academic Science: A Changing Landscape'. *Psychological Science in the Public Interest* 15 (3): 75–141. https://doi.org/10.1177/1529100614541236.

Ceci, Stephen J., and Wendy M. Williams. 2007. *Why Aren't More Women in Science?: Top Researchers Debate the Evidence*. 1 edition. American Psychological Association.

———. 2011. 'Understanding Current Causes of Women's Underrepresentation in Science'. *Proceedings of the National Academy of Sciences* 108 (8): 3157–62. https://doi.org/10.1073/pnas.1014871108.

———. 2015. 'Women Have Substantial Advantage in STEM Faculty Hiring, except When Competing against More-Accomplished Men'. *Frontiers in Psychology* 6: 1532. https://doi.org/10.3389/fpsyg.2015.01532.

Checchi, Daniele, Simona Cicognani, and Nevena Kulic. 2019. 'Gender Quotas or Girls' Networks? Evidence from an Italian Research Selection'. *Work, Employment and Society*, January, 0950017018813071. https://doi.org/10.1177/0950017018813071.

Cimpian, A., and S.-J. Leslie. 2015. 'Response to Comment on "Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines"'. *Science* 349 (6246). https://doi.org/10.1126/science.aaa9892.

Cole, J, and H Zuckerman. 1984. 'THE PRODUCTIVITY PUZZLE: PERSISTENCE AND CHANGE IN PATTERNS OF PUBLICATION OF MEN AND WOMEN SCIENTISTS'. *Advances in Motivation and Achievement* 2: 217–58.

Cole, Jonathan R. 1979. *Fair Science: Women in the Scientific Community*. The Free Press.

Cole, Jonathan R., and Stephen Cole. 1973. *Social Stratification in Science*. Y First edition edition. Chicago: University of Chicago Press.

Cole, Jonathan R., and Harriet Zuckerman. 1991. 'Marriage, Motherhood, and Research Performance. 1987'. In *The Outer Circle: Women in the Scientific Community*, edited by Harriet Zuckerman, Jonathan R. Cole, and John T. Bruer, 1992. Yale University Press, 157–70. New York: WW Norton & Co.

Cole, Stephen, and Robert Fiorentine. 1991. 'Discrimination against women in science: The confusion of outcome with process'. In *The Outer Circle: Women in the Scientific*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Community, edited by Harriet Zuckerman, Jonathan R. Cole, and John T. Bruer, 205–26. New Haven & London: Yale University Press.

Conti, A., O. Denas, and F. Visentin. 2014. 'Knowledge Specialization in Ph.D. Student Groups'. *IEEE Transactions on Engineering Management* 61 (1): 52–67. https://doi.org/10.1109/TEM.2013.2283039.

Conti, Annamaria, and Fabiana Visentin. 2015. 'Science and Engineering Ph.D. Students' Career Outcomes, by Gender'. *PLOS ONE* 10 (8): e0133177. https://doi.org/10.1371/journal.pone.0133177.

Corley, Elizabeth, and Monica Gaughan. 2005. 'Scientists' Participation in University Research Centers: What Are the Gender Differences?' *The Journal of Technology Transfer* 30 (4): 371–81. https://doi.org/10.1007/s10961-005-2582-4.

Correll, S.J. 2001. 'Gender and the Career Choice Process: The Role of Biased Self-Assessments'. *American Journal of Sociology* 10 (6): 1691–1730. https://doi.org/10.1086/321299.

———. 2004. 'Constraints into Preferences: Gender, Status, and Emerging Career Aspirations'. *American Sociological Review* 69 (1): 93–113. https://doi.org/10.1177/000312240406900106.

Croson, Rachel, and Uri Gneezy. 2009. 'Gender Differences in Preferences'. *Journal of Economic Literature* 47 (2): 448–74. https://www.jstor.org/stable/27739928.

Cruz-Castro, Laura, and Luis Sanz-Menéndez. 2018. 'Autonomy and Authority in Public Research Organisations: Structure and Funding Factors'. *Minerva* 56 (2): 135–60. https://doi.org/10.1007/s11024-018-9349-1.

Dasgupta, Nilanjana, Melissa McManus Scircle, and Matthew Hunsinger. 2015. 'Female Peers in Small Work Groups Enhance Women's Motivation, Verbal Participation, and Career Aspirations in Engineering'. *Proceedings of the National Academy of Sciences* 112 (16): 4988–93. https://doi.org/10.1073/pnas.1422822112.

De Paola, Maria, Michela Ponzo, and Vincenzo Scoppa. 2017. 'Gender Differences in the Propensity to Apply for Promotion: Evidence from the Italian Scientific Qualification'. *Oxford Economic Papers* 69 (4): 986–1009. https://doi.org/10.1093/oep/gpx023.

Dennehy, Tara C., and Nilanjana Dasgupta. 2017. 'Female Peer Mentors Early in College Increase Women's Positive Academic Experiences and Retention in Engineering'. *Proceedings of the National Academy of Sciences* 114 (23): 5964–69. https://doi.org/10.1073/pnas.1613117114.

Derks, Belle, Naomi Ellemers, Colette van Laar, and Kim de Groot. 2011. 'Do Sexist Organizational Cultures Create the Queen Bee?' *British Journal of Social Psychology* 50 (3): 519–35. https://doi.org/10.1348/014466610X525280.

Derks, Belle, Colette Van Laar, and Naomi Ellemers. 2016. 'The Queen Bee Phenomenon: Why Women Leaders Distance Themselves from Junior Women'. *The Leadership Quarterly*, Special Issue: Gender and Leadership, 27 (3): 456–69. https://doi.org/10.1016/j.leaqua.2015.12.007.

Devine, Patricia G., Patrick S. Forscher, William T.L. Cox, Anna Kaatz, Jennifer Sheridan, and Molly Carnes. 2017. 'A Gender Bias Habit-Breaking Intervention Led to Increased Hiring of Female Faculty in STEMM Departments'. *Journal of Experimental Social Psychology* 73. https://doi.org/10.1016/j.jesp.2017.07.002.

Eagly, A.H., and S.J. Karau. 1991. 'Gender and the Emergence of Leaders: A Meta-Analysis'. *Journal of Personality and Social Psychology* 60 (5): 685–710. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0001561210&partnerID=40&md5=d7d327aea47a02c70689a9914912c506.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Eagly, Alice, and Linda L. Carli. 2007. 'Women and the Labyrinth of Leadership'. *Harvard Business Review*, 1 September 2007. https://hbr.org/2007/09/women-and-the-labyrinth-of-leadership.

Eaton, Asia A., Jessica F. Saunders, Ryan K. Jacobson, and Keon West. 2019. 'How Gender and Race Stereotypes Impact the Advancement of Scholars in STEM: Professors' Biased Evaluations of Physics and Biology Post-Doctoral Candidates'. *Sex Roles*, June. https://doi.org/10.1007/s11199-019-01052-w.

Ellemers, Naomi. 2018. 'Gender Stereotypes'. *Annual Review of Psychology* 69 (1): 275–98. https://doi.org/10.1146/annurev-psych-122216-011719.

elsevier. 2017. 'Gender in the Global Research Landscape Report'. https://www.elsevier.com/research-intelligence/resource-library/gender-report.

Elsevier. 2018. 'Gender in the Global Research Landscape | Research Intelligence | Elsevier'. https://www.elsevier.com/research-intelligence/campaigns/gender-17.

ETAN. 2000. 'Promoting Excellence through Mainstreaming Gender Equality —'. Publication. https://ufm.dk/en/publications/2000/promoting-excellence-through-mainstreaming-gender-equality.

European Commission. 2004. 'Gender and Excellence in the Making.' Website. https://publications.europa.eu/en/publication-detail/-/publication/ce10dfeb-4e36-49ed-a9ec-b2182c8b986f.

———, ed. 2009. *The Gender Challenge in Research Funding: Assessing the European National Scenes ; Report*. EUR 23721. Luxembourg: Office for Official Publ. of the European Communities.

———. 2014. 'Marie Curie Researchers and Their Long-Term Career Development : A Comparative Study : Final Report.' Website. https://publications.europa.eu/en/publication-detail/-/publication/9c2cb7d8-3773-430c-b1c0-db94ec421b01.

———. 2017. 'Implicit Gender Biases During Evaluations: How to Raise Awareness and Change Attitudes? Workshop Report Brussels, 30-31 May 2017'. Workshop report. Brussels: European Commission. http://ec.europa.eu/research/swafs/pdf/pub_gender_equality/report_on_implicit_gender_biases_during_evaluations.pdf.

———. 2019. 'She Figures 2018'. https://doi.org/10.2777/936.

Fernandez-Mateo, Isabel, and Sarah Kaplan. 2018. 'Gender and Organization Science: Introduction to a Virtual Special Issue'. *Organization Science* 29 (6): 1229–36. https://doi.org/10.1287/orsc.2018.1249.

Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, et al. 2018. 'Science of Science'. *Science* 359 (6379): eaao0185. https://doi.org/10.1126/science.aao0185.

Fox, Mary Frank. 2005. 'Gender, Family Characteristics, and Publication Productivity among Scientists'. *Social Studies of Science* 35 (1): 131–50. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/25046632.

———. 2015. 'Gender and Clarity of Evaluation among Academic Scientists in Research Universities'. *Science, Technology, & Human Values* 40 (4): 487–515. https://doi.org/10.1177/0162243914564074.

Fox, Mary Frank, and Catherine A. Faver. 1985. 'Men, Women, and Publication Productivity: Patterns among Social Work Academics'. *The Sociological Quarterly* 26 (4): 537–49. https://www.jstor.org/stable/4106103.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Gallo, Stephen A., Joanne H. Sullivan, and Scott R. Glisson. 2016. 'The Influence of Peer Reviewer Expertise on the Evaluation of Research Funding Applications'. *PLOS ONE* 11 (10): e0165147. https://doi.org/10.1371/journal.pone.0165147.

Gaughan, Monica, and Barry Bozeman. 2002. 'Using Curriculum Vitae to Compare Some Impacts of NSF Research Grants with Research Center Funding'. *Research Evaluation* 11 (1): 17–26. https://doi.org/10.3152/147154402781776952.

———. 2016. 'Using the Prisms of Gender and Rank to Interpret Research Collaboration Power Dynamics'. *Social Studies of Science* 46 (4): 536–58. https://doi.org/10.1177/0306312716652249.

Gender Action. 2019. 'The Role of Funding Agencies in the Promotion of Gender Equality in R&I'. Briefing paper 10. http://genderaction.eu/wp-content/uploads/2019/03/GENDERACTION_PolicyBrief_RFOs-March-8-2019.pdf.

Geraci, Lisa, Steve Balsis, and Alexander J. Busch Busch. 2015. 'Gender and the h Index in Psychology'. *Scientometrics* 105 (3): 2023–2034. https://doi.org/10.1007/s11192-015-1757-5.

Gerber, Alan S., and Donald P. Green. 2011. 'Field Experiments and Natural Experiments'. In *The Oxford Handbook of Political Science*, edited by Box-Steffensmeier, Henry E. Brady, and David Collier, 357–81. Oxford, UNITED KINGDOM: Oxford University Press. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199604456.001.0001/oxfordhb-9780199604456-e-050.

Gino, Francesca, Caroline Ashley Wilmuth, and Alison Wood Brooks. 2015a. 'Compared to Men, Women View Professional Advancement as Equally Attainable, but Less Desirable'. *Proceedings of the National Academy of Sciences* 112 (40): 12354–59. https://doi.org/10.1073/pnas.1502567112.

———. 2015b. 'Compared to Men, Women View Professional Advancement as Equally Attainable, but Less Desirable'. *Proceedings of the National Academy of Sciences* 112 (40): 12354–59. https://doi.org/10.1073/pnas.1502567112.

Ginther, Donna K., and Kathy J. Hayes. 2003. 'Gender Differences in Salary and Promotion for Faculty in the Humanities 1977-95'. *The Journal of Human Resources* 38 (1): 34–73. https://doi.org/10.2307/1558755.

Gläser, Jochen, and Grit Laudel. 2016. 'Governing Science: How Science Policy Shapes Research Content'. *European Journal of Sociology / Archives Européennes de Sociologie* 57 (1): 117–68. https://doi.org/10.1017/S0003975616000047.

Gläser, Jochen, and Kathia Serrano Velarde. 2018. 'Changing Funding Arrangements and the Production of Scientific Knowledge: Introduction to the Special Issue'. *Minerva* 56 (1): 1–10. https://doi.org/10.1007/s11024-018-9344-6.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. 'Performance in Competitive Environments: Gender Differences'. *The Quarterly Journal of Economics* 118 (3): 1049–74. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/25053930.

Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women by Goldin, Claudia: OUP, 9780195050776 - Naomi Symes Books PBFA*. https://www.abebooks.co.uk/9780195050776/Understanding-Gender-Gap-Economic-History-0195050770/plp.

Goldin, Claudia, and Cecilia Rouse. 2000. 'Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians'. *The American Economic Review* 90 (4): 715–41. https://www.jstor.org/stable/117305.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Goulden, M., M. A. Mason, and K. Frasch. 2011. 'Keeping Women in the Science Pipeline'. *The ANNALS of the American Academy of Political and Social Science* 638 (1): 141–62. https://doi.org/10.1177/0002716211416925.

Graham, John W., and Steven A. Smith. 2005. 'Gender Differences in Employment and Earnings in Science and Engineering in the US'. *Economics of Education Review* 24 (3): 341–54. https://doi.org/10.1016/j.econedurev.2004.06.005.

Gross, Edward. 1968. 'Plus Ca Change...? The Sexual Structure of Occupations over Time'. *Social Problems* 16 (2): 198–208. https://doi.org/10.2307/800005.

Gupta, Namrata, C. Kemelgor, S. Fuchs, and Henry Etzkowitz. 2004. 'The `Triple Burden´'. In *Gender and Excellence in the Making*, edited by E.C. European Commission. Gender and Excellence in the Making. Luxembourg.

Handley, Ian M., Elizabeth R. Brown, Corinne A. Moss-Racusin, and Jessi L. Smith. 2015. 'Quality of Evidence Revealing Subtle Gender Biases in Science Is in the Eye of the Beholder'. *Proceedings of the National Academy of Sciences* 112 (43): 13201–6. https://doi.org/10.1073/pnas.1510649112.

Hargens, Lowell L. 2012. 'Academic Labor Markets and Assistant Professors' Employment Outcomes'. *Research in Higher Education* 53 (3): 311–24. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/41475393.

Hechtman, Lisa A., Nathan P. Moore, Claire E. Schulkey, Andrew C. Miklos, Anna Maria Calcagno, Richard Aragon, and Judith H. Greenberg. 2018. 'NIH Funding Longevity by Gender'. *Proceedings of the National Academy of Sciences* 115 (31): 7943–48. https://doi.org/10.1073/pnas.1800615115.

Heilman, Madeline E. 2001. 'Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder'. *Journal of Social Issues* 57 (4): 657–74. https://doi.org/10.1111/0022-4537.00234.

———. 2012. 'Gender Stereotypes and Workplace Bias'. *Research in Organizational Behavior* 32 (January): 113–35. https://doi.org/10.1016/j.riob.2012.11.003.

Helgesson, Karin Svedberg, and Ebba Sjögren. 2019. 'No Finish Line: How Formalization of Academic Assessment Can Undermine Clarity and Increase Secrecy'. *Gender, Work & Organization* 26 (4): 558–81. https://doi.org/10.1111/gwao.12355.

Hoff Summers, Christina. 2009. *The Science on Women and Science*. The AEI Press.

Holliday, E.B., R. Jagsi, L.D. Wilson, M. Choi, C.R. Thomas, and C.D. Fuller. 2014. 'Gender Differences in Publication Productivity, Academic Position, Career Duration, and Funding among U.S. Academic Radiation Oncology Faculty'. *Academic Medicine* 89 (5): 767–73. https://doi.org/10.1097/ACM.0000000000000229.

Holman, Luke, Devi Stuart-Fox, and Cindy E. Hauser. 2018. 'The Gender Gap in Science: How Long until Women Are Equally Represented?' Edited by Cassidy Sugimoto. *PLOS Biology* 16 (4): e2004956. https://doi.org/10.1371/journal.pbio.2004956.

Hosek, Susan D., Amy G. Cox, Bonnie Ghosh-Dastidar, Aaron Kofner, Nishal Ray Ramphal, Jon Scott, and Sandra H. Berry. 2005. 'Gender Differences in Major Federal External Grant Programs'. Product Page. https://www.rand.org/pubs/technical_reports/TR307.html.

Huang, Junming, Alex J Gates, Rob Sinatra, and A-L Barabasi. 2019. 'Historical Comparison of Gender Inequality in Scientific Careers across Countries and Disciplines'. *Arxiv*, no. 1907.04103v1 (July). https://arxiv.org/pdf/1907.04103.pdf.

Hunter, Laura A., and Erin Leahey. 2010. 'Parenting and Research Productivity: New Evidence and Methods'. *Social Studies of Science* 40 (3): 433–51. https://doi.org/10.1177/0306312709358472.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Husu, Liisa, and Suzanne de Cheveigné. 2010. 'Gender and Gatekeeping of Excellence in Research Funding: European Perspectives'. In *GenderChange in Academia: Re-Mapping the Fields of Work, Knowledge, and Politics from a Gender Perspective*, edited by Birgit Riegraf, Brigitte Aulenbacher, Edit Kirsch-Auwärter, and Ursula Müller, 43–59. Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92501-1_4.

Jacob, Brian A., and Lars Lefgren. 2011. 'The Impact of NIH Postdoctoral Training Grants on Scientific Productivity'. *Research Policy* 40 (6): 864–74. https://doi.org/10.1016/j.respol.2011.04.003.

Jappelli, Tullio, Carmela Anna Nappi, and Roberto Torrini. 2017. 'Gender Effects in Research Evaluation'. *Research Policy* 46 (5): 911–24. https://doi.org/10.1016/j.respol.2017.03.002.

Kahn, Shulamit, and Donna Ginther. 2018. 'Women and Science, Technology, Engineering, and Mathematics (STEM)'. *The Oxford Handbook of Women and the Economy*, July. https://doi.org/10.1093/oxfordhb/9780190628963.013.13.

Kahn, Shulamit, and Donna K. Ginther. 2015. 'Are Recent Cohorts of Women with Engineering Bachelors Less Likely to Stay in Engineering?' *Frontiers in Psychology* 6. https://doi.org/10.3389/fpsyg.2015.01144.

Kalev, Alexandra, Erin Kelly, and Frank Dobbin. 2006. 'Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies'. *American Sociological Review* 71 (4): 589–617. https://www.jstor.org/stable/30039011.

Kalpazidou Schmidt, Evanthia, and Marina Cacace. 2017. 'Addressing Gender Inequality in Science: The Multifaceted Challenge of Assessing Impact'. *Research Evaluation* 26 (2): 102–14. https://doi.org/10.1093/reseval/rvx003.

Kaminski, Deborah, and Cheryl Geisler. 2012. 'Survival Analysis of Faculty Retention in Science and Engineering by Gender'. *Science* 335 (6070): 864–66. https://doi.org/10.1126/science.1214844.

Kelly, Kimberly, and Linda Grant. 2012. 'Penalties and Premiums: The Impact of Gender, Marriage, and Parenthood on Faculty Salaries in Science, Engineering and Mathematics (SEM) and Non-SEM Fields'. *Social Studies of Science* 42 (6): 869–96. https://doi.org/10.1177/0306312712457111.

Klaus, Bernd, and David del Alamo. 2018. 'Talent Identification at the Limits of Peer Review: An Analysis of the EMBO Postdoctoral Fellowships Selection Process'. *BioRxiv*, December, 481655. https://doi.org/10.1101/481655.

Knobloch-Westerwick, Silvia, and Carroll J. Glynn. 2013. 'The Matilda Effect—Role Congruity Effects on Scholarly Communication: A Citation Analysis of Communication Research and Journal of Communication Articles'. *Communication Research* 40 (1): 3–26. https://doi.org/10.1177/0093650211418339.

Knobloch-Westerwick, Silvia, Carroll J. Glynn, and Michael Huge. 2013. 'The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest'. *Science Communication* 35 (5): 603–25. https://doi.org/10.1177/1075547012472684.

Kolev, Julian, Yuly Fuentes-Medel, and Fiona Murray. 2019. 'Is Blinded Review Enough? How Gendered Outcomes Arise Even Under Anonymous Evaluation'. Working Paper 25759. National Bureau of Economic Research. https://doi.org/10.3386/w25759.

Kugler, Adriana D, Catherine H Tinsley, and Olga Ukhaneva. 2017. 'Choice of Majors: Are Women Really Different from Men?' Working Paper 23735. National Bureau of Economic Research. https://doi.org/10.3386/w23735.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Lamont, Michele. 2009. *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, Mass.: Harvard University Press.

Langfeldt, Liv, Carter Walter Bloch, and Gunnar Sivertsen. 2015. 'Options and Limitations in Measuring the Impact of Research Grants—Evidence from Denmark and Norway'. *Research Evaluation* 24 (3): 256–70. https://doi.org/10.1093/reseval/rvv012.

Larivière, Vincent, and Cassidy R. Sugimoto. 2017. 'The End of Gender Disparities in Science? If Only It Were True...' CWTS. 2017. https://www.cwts.nl:443/blog?article=n-q2z294.

Larivière, Vincent, Etienne Vignola-Gagné, Christian Villeneuve, Pascal Gélinas, and Yves Gingras. 2011. 'Sex Differences in Research Funding, Productivity and Impact: An Analysis of Québec University Professors'. *Scientometrics* 87 (3): 483–98. https://doi.org/10.1007/s11192-011-0369-y.

Le Feuvre, Nicky. 2018. 'The Gendered Diversification of Academic Career Paths in Comparative Perspective'. In *Gender and Precarious Research Careers*, by Annalisa Murgia, 1st ed., 50–80. Routledge. https://doi.org/10.4324/9781315201245-3.

Leahey, E., J. L. Crockett, and L. A. Hunter. 2008. 'Gendered Academic Careers: Specializing for Success?' *Social Forces* 86 (March): 1273–1309. https://doi.org/10.1353/sof.0.0018.

Leahey, Erin. 2006. 'Gender Differences in Productivity: Research Specialization as a Missing Link'. *Gender & Society* 20 (6): 754–80. https://doi.org/10.1177/0891243206293030.

Lee, Carole J, Cassidy R. Sugimoto, Guo Zhang, and Blaise Cronin. 2013. 'Bias in Peer Review'. *Journal of the American Society for Information Science and Technology* 64 (1): 2–17. https://doi.org/10.1002/asi.22784.

Lee, Romy van der, and Naomi Ellemers. 2015. 'Gender Contributes to Personal Research Funding Success in The Netherlands'. *Proceedings of the National Academy of Sciences* 112 (40): 12349–53. https://doi.org/10.1073/pnas.1510159112.

Lee, Romy van der, and Naomi Ellemers. 2018. 'Perceptions of Gender Inequality in Academia. Reluctance to Let Go of Individual Merit Ideology'. In *Belief Systems and the Perception of Reality*, edited by Bastiaan Rutjens and Mark Brandt, 1 edition. New York: Routledge.

Lee, Sooho, and Barry Bozeman. 2005. 'The Impact of Research Collaboration on Scientific Productivity'. *Social Studies of Science* 35 (5): 673–702. https://doi.org/10.1177/0306312705052359.

Leenders, Joke, Inge L. Bleijenbergh, and Marieke C.L. Van den Brink. 2019. 'Myriad Potential for Mentoring: Understanding the Process of Transformational Change through a Gender Equality Intervention'. *Gender, Work & Organization*, May, gwao.12385. https://doi.org/10.1111/gwao.12385.

Leslie, Sarah-Jane, Andrei Cimpian, Meredith Meyer, and Edward Freeland. 2015. 'Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines'. *Science* 347 (6219): 262–65. https://doi.org/10.1126/science.1261375.

Ley, Timothy J., and Barton H. Hamilton. 2008. 'The Gender Gap in NIH Grant Applications'. *Science* 322 (5907): 1472–74. https://doi.org/10.1126/science.1165878.

Lincoln, Yvonna S., and Egon G. Guba. 1980. 'The Distinction between Merit and Worth in Evaluation'. *Educational Evaluation and Policy Analysis* 2 (4): 61–71. https://doi.org/10.2307/1163674.

Long, J. Scott, Paul D. Allison, and Robert McGinnis. 1993. 'Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity'. *American Sociological Review* 58 (5): 703–22. https://doi.org/10.2307/2096282.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Long, J.Scott. 1992. 'Measures of Sex Differences in Scientific Productivity'. *Social Forces* 71 (1): 159–78. http://search.ebscohost.com/login.aspx?direct=true&db=sih&AN=9301170060&site=ehost-live.

Magua, Wairimu, Xiaojin Zhu, Anupama Bhattacharya, Amarette Filut, Aaron Potvien, Renee Leatherberry, You-Geon Lee, et al. 2017. 'Are Female Applicants Disadvantaged in National Institutes of Health Peer Review? Combining Algorithmic Text Mining and Qualitative Methods to Detect Evaluative Differences in R01 Reviewers' Critiques'. *Journal of Women's Health* 26 (5): 560–70. https://doi.org/10.1089/jwh.2016.6021.

Mairesse, Jacques, and Michele Pezzoni. 2015. 'Does Gender Affect Scientific Productivity ?' *Revue Économique* 66 (1): 65–113. https://doi.org/10.3917/reco.661.0065.

Marini, Giulio, and Viviana Meschitti. 2018. 'The Trench Warfare of Gender Discrimination: Evidence from Academic Promotions to Full Professor in Italy'. *Scientometrics* 115 (2): 989–1006. https://doi.org/10.1007/s11192-018-2696-8.

Marsh, Herbert W., Lutz Bornmann, Rüdiger Mutz, Hans-Dieter Daniel, and Alison O'Mara. 2009. 'Gender Effects in the Peer Reviews of Grant Proposals: A Comprehensive Meta-Analysis Comparing Traditional and Multilevel Approaches'. *Review of Educational Research* 79 (3): 1290–1326. https://doi.org/10.3102/0034654309334143.

Marsh, Herbert W., Upali W. Jayasinghe, and Nigel W. Bond. 2008. 'Improving the Peer-Review Process for Grant Applications: Reliability, Validity, Bias, and Generalizability'. *The American Psychologist* 63 (3): 160–68. https://doi.org/10.1037/0003-066X.63.3.160.

———. 2011. 'Gender Differences in Peer Reviews of Grant Applications: A Substantive-Methodological Synergy in Support of the Null Hypothesis Model'. *Journal of Informetrics* 5 (1): 167–80. https://doi.org/10.1016/j.joi.2010.10.004.

Martell, Richard F., Cynthia G. Emrich, and James Robison-Cox. 2012. 'From Bias to Exclusion: A Multilevel Emergent Theory of Gender Segregation in Organizations'. *Research in Organizational Behavior* 32 (January): 137–62. https://doi.org/10.1016/j.riob.2012.10.001.

Mauleón, Elba, María Bordons, and Charles Oppenheim. 2008. 'The Effect of Gender on Research Staff Success in Life Sciences in the Spanish National Research Council'. *Research Evaluation* 17 (September): 213–25. https://doi.org/10.3152/095820208X331676.

Melin, Göran, and Rickard Danell. 2006. 'The Top Eight Percent: Development of Approved and Rejected Applicants for a Prestigious Grant in Sweden'. *Science & Public Policy (SPP)* 33 (10): 702–12. http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=24327187&site=ehost-live.

Merton, Robert King. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Meyer, Meredith, Andrei Cimpian, and Sarah-Jane Leslie. 2015. 'Women Are Underrepresented in Fields Where Success Is Believed to Require Brilliance'. *Frontiers in Psychology* 6. https://doi.org/10.3389/fpsyg.2015.00235.

Ministerio Ciencia e Innovacion. 2011. 'WHITE PAPER on the Position of Women in Science in Spain'. http://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/UMYC/WhitePaper_Interactive.pdf.

Moss-Racusin, Corinne A., John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2012. 'Science Faculty's Subtle Gender Biases Favor Male Students'.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Proceedings of the National Academy of Sciences 109 (41): 16474–79. https://doi.org/10.1073/pnas.1211286109.

Moss-Racusin, Corinne A., Aneta K. Molenda, and Charlotte R. Cramer. 2015. 'Can Evidence Impact Attitudes? Public Reactions to Evidence of Gender Bias in STEM Fields'. *Psychology of Women Quarterly* 39 (2): 194–209. https://doi.org/10.1177/0361684314565777.

Moss-Racusin, Corinne A., Jojanneke van der Toorn, John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2016. 'A Scientific Diversity Intervention to Reduce Gender Bias in a Sample of Life Scientists'. *CBE Life Sciences Education*. https://doi.org/10.1187/cbe.15-09-0187.

Murray, Dakota, Kyle Siler, Vincent Larivière, Wei Mun Chan, Andrew M. Collings, Jennifer Raymond, and Cassidy R. Sugimoto. 2019. 'Gender and International Diversity Improves Equity in Peer Review'. *BioRxiv*, April, 400515. https://doi.org/10.1101/400515.

Mutz, Diana C. 2011. *Population-Based Survey Experiments (EBook and Paperback)*. http://press.princeton.edu/titles/9620.html.

Mutz, Rüdiger, Lutz Bornmann, and Hans-Dieter Daniel. 2012. 'Does gender matter in grant peer review? An empirical investigation using the example of the Austrian science fund'. *Zeitschrift für Psychologie/Journal of Psychology* 220 (2): 121–29. http://dx.doi.org.ezproxy1.lib.asu.edu/10.1027/2151-2604/a000103.

Nakhaie, M. Reza. 2002. 'Gender Differences in Publication among University Professors in Canada*'. *Canadian Review of Sociology/Revue Canadienne de Sociologie* 39 (2): 151–79. https://doi.org/10.1111/j.1755-618X.2002.tb00615.x.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. 2007. *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. Washington DC: National Academy of Sciences. https://doi.org/10.17226/11741.

National Research Council. 2010. *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*. https://doi.org/10.17226/12062.

Nelson, Julie A. 2014. 'The Power of Stereotyping and Confirmation Bias to Overwhelm Accurate Assessment: The Case of Economics, Gender, and Risk Aversion'. *Journal of Economic Methodology* 21 (3): 211–31. https://doi.org/10.1080/1350178X.2014.939691.

———. 2015. 'Are Women Really More Risk-Averse Than Men? A Re-Analysis of the Literature Using Expanded Methods'. *Journal of Economic Surveys* 29 (3): 566–85. https://doi.org/10.1111/joes.12069.

Neufeld, Jörg, and Stefan Hornbostel. 2012. 'Funding Programmes for Young Scientists—Do the "Best" Apply?' *Research Evaluation* 21 (4): 270–79. https://doi.org/10.1093/reseval/rvs018.

Niederle, Muriel, and Lise Vesterlund. 2007. 'Do Women Shy Away From Competition? Do Men Compete Too Much?' *The Quarterly Journal of Economics* 122 (3): 1067–1101. https://doi.org/10.1162/qjec.122.3.1067.

———. 2011. 'Gender and Competition'. *Annual Review of Economics* 3 (1): 601–30. https://doi.org/10.1146/annurev-economics-111809-125122.

Nielsen, M W. 2015. 'Make Academic Job Advertisements Fair to All'. *Nature* 525 (7570): 427. https://doi.org/10.1038/525427a.

———. 2017a. 'Gender Consequences of a National Performance-Based Funding Model: New Pieces in an Old Puzzle'. *Studies in Higher Education* 42 (6): 1033–1055. https://doi.org/10.1080/03075079.2015.1075197.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

———. 2017b. ‘Scandinavian Approaches to Gender Equality in Academia: A Comparative Study’. *Scandinavian Journal of Educational Research* 61 (3): 295–318. https://doi.org/10.1080/00313831.2016.1147066.

Nielsen, Mathias Wullum. 2018. ‘Scientific Performance Assessments Through a Gender Lens’. *Science & Technology Studies* 31 (1): 2–30. https://doi.org/10.23987/sts.60610.

Nielsen, Mathias Wullum, Sharla Alegria, Love Börjeson, Henry Etzkowitz, Holly J. Falk-Krzesinski, Aparna Joshi, Erin Leahey, Laurel Smith-Doerr, Anita Williams Woolley, and Londa Schiebinger. 2017. ‘Opinion: Gender Diversity Leads to Better Science’. *Proceedings of the National Academy of Sciences* 114 (8): 1740–42. https://doi.org/10.1073/pnas.1700616114.

O’Connor, Pat, and Antoinette Fauve-Chamoux. 2016. ‘European Policies and Research Funding: A Case Study of Gender Inequality and Lack of Diversity in a Nordic Research Programme’. *Policy & Politics* 44 (4): 627–43. https://doi.org/10.1332/030557315X14501227093917.

Oliveira, Diego F. M., Yifang Ma, Teresa K. Woodruff, and Brian Uzzi. 2019. ‘Comparison of National Institutes of Health Grant Amounts to First-Time Male and Female Principal Investigators’. *JAMA* 321 (9): 898–900. https://doi.org/10.1001/jama.2018.21944.

Ooms, Ward, Claudia Werker, and Christian Hopp. 2019. ‘Moving up the Ladder: Heterogeneity Influencing Academic Careers through Research Orientation, Gender, and Mentors’. *Studies in Higher Education* 44 (7): 1268–89. https://doi.org/10.1080/03075079.2018.1434617.

Palomba, Rossella. 2004. ‘Does Gender Matter in Scientific Leadership?’ In *Gender and Excellence in the Making*, edited by European Commission, 121–25. Gender and Excellence in the Making. Luxembourg.

Pezzoni, Michele, Jacques Mairesse, Paula Stephan, and Julia Lane. 2016. ‘Gender and the Publication Output of Graduate Students: A Case Study’. *PLOS ONE* 11 (1): e0145146. https://doi.org/10.1371/journal.pone.0145146.

Phelps, Edmund S. 1972. ‘The Statistical Theory of Racism and Sexism’. *The American Economic Review* 62 (4): 659–61. https://www.jstor.org/stable/1806107.

Pina, David G., Lana Barać, Ivan Buljan, Francisco Grimaldo, and Ana Marušić. 2019. ‘Effects of Seniority, Gender and Geography on the Bibliometric Output and Collaboration Networks of European Research Council (ERC) Grant Recipients’. *PLOS ONE* 14 (2): e0212286. https://doi.org/10.1371/journal.pone.0212286.

Powell, Stina. 2018. ‘Gender Equality in Academia’. *The International Journal of Diversity in Organizations, Communities, and Nations: Annual Review* 18 (1): 19–35. https://doi.org/10.18848/1447-9532/CGP/v18i01/19-35.

Preston, Anne E. 2004. *Leaving Science*. 1 edition. New York: Russell Sage Foundation.

Quadlin, Natasha. 2018. ‘The Mark of a Woman’s Record: Gender and Academic Performance in Hiring’. *American Sociological Review* 83 (2): 331–60. https://doi.org/10.1177/0003122418762291.

Reskin, Barbara F. 1976. ‘Sex Differences in Status Attainment in Science: The Case of the Postdoctoral Fellowship’. *American Sociological Review* 41 (4): 597–612. https://doi.org/10.2307/2094838.

———. 2000a. ‘Getting It Right: Sex and Race Inequality in Work Organizations’. *Annual Review of Sociology* 26 (1): 707–9. https://doi.org/10.1146/annurev.soc.26.1.707.

———. 2000b. ‘The Proximate Causes of Employment Discrimination’. *Contemporary Sociology* 29 (2): 319–28. https://doi.org/10.2307/2654387.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

———. 2003. ‘Including Mechanisms in Our Models of Ascriptive Inequality: 2002 Presidential Address’. *American Sociological Review* 68 (1): 1–21. https://doi.org/10.2307/3088900.

Reskin, Barbara F., and Denise D. Bielby. 2005. ‘A Sociological Perspective on Gender and Career Outcomes’. *The Journal of Economic Perspectives* 19 (1): 71–86. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/4134993.

Reskin, Barbara F., Debra B. McBrier, and Julie A. Kmec. 1999. ‘The Determinants and Consequences of Workplace Sex and Race Composition’. *Annual Review of Sociology* 25 (1): 335–61. https://doi.org/10.1146/annurev.soc.25.1.335.

Reskin, Barbara F., and Debra Branch McBrier. 2000. ‘Why Not Ascription? Organizations’ Employment of Male and Female Managers’. *American Sociological Review* 65 (2): 210–33. https://doi.org/10.2307/2657438.

Reskin, Barbara F., Patricia A. Roos, Katharine M. Donato, Polly A. Phipps, Barbara J. Thomas, Chloe E. Bird, Linda A. Detman, and Thomas Steiger. 1990. *Job Queues, Gender Queues: Explaining Women’s Inroads into Male Occupations*. Temple University Press. http://www.jstor.org/stable/j.ctt14btcj8.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2014. ‘How Stereotypes Impair Women’s Careers in Science’. *Proceedings of the National Academy of Sciences* 111 (12): 4403–8. https://doi.org/10.1073/pnas.1314788111.

Ridgeway, Cecilia L. 2014. ‘Why Status Matters for Inequality’. *American Sociological Review* 79 (1): 1–16. https://www.jstor.org/stable/43187523.

Ridgeway, Cecilia L., and Shelley J. Correll. 2004. ‘Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations’. *Gender and Society* 18 (4): 510–31. https://www.jstor.org/stable/4149448.

Rivera, Lauren A. 2017. ‘When Two Bodies Are (Not) a Problem: Gender and Relationship Status Discrimination in Academic Hiring’. *American Sociological Review* 82 (6): 1111–38. https://doi.org/10.1177/0003122417739294.

Rivera, Lauren A., and András Tilcsik. 2019. ‘Scaling Down Inequality: Rating Scales, Gender Bias, and the Architecture of Evaluation’. *American Sociological Review* 84 (2): 248–74. https://doi.org/10.1177/0003122419833601.

Roos, Patricia A., and Barbara F. Reskin. 1992. ‘Occupational Desegregation in the 1970s: Integration and Economic Equity?’ *Sociological Perspectives* 35 (1): 69–91. https://doi.org/10.2307/1389369.

Rørstad, Kristoffer, and Dag W. Aksnes. 2015. ‘Publication Rate Expressed by Age, Gender and Academic Position – A Large-Scale Analysis of Norwegian Academic Staff’. *Journal of Informetrics* 9 (2): 317–33. https://doi.org/10.1016/j.joi.2015.02.003.

Ross, Erin. n.d. ‘Gender Bias Distorts Peer Review across Fields’. *Nature News*. Accessed 1 May 2019. https://doi.org/10.1038/nature.2017.21685.

Ross, Joseph S., Cary P. Gross, Mayur M. Desai, Yuling Hong, Augustus O. Grant, Stephen R. Daniels, Vladimir C. Hachinski, Raymond J. Gibbons, Timothy J. Gardner, and Harlan M. Krumholz. 2006. ‘Effect of Blinded Peer Review on Abstract Acceptance’. *JAMA* 295 (14): 1675–80. https://doi.org/10.1001/jama.295.14.1675.

Rosser, Sue V. 2004. *The Science Glass Ceiling: Academic Women Scientist and the Struggle to Succeed*. 1st ed. Routledge. http://gen.lib.rus.ec/book/index.php?md5=DFDFB11B0C6833D1CCB89E525E6238DB.

Ross-Hellauer, Tony, Arvid Deppe, and Birgit Schmidt. 2017. ‘Survey on Open Peer Review: Attitudes and Experience amongst Editors, Authors and Reviewers’. *PLOS ONE* 12 (12): e0189311. https://doi.org/10.1371/journal.pone.0189311.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Roumbanis, Lambros. 2019. 'Peer Review or Lottery? A Critical Analysis of Two Different Forms of Decision-Making Mechanisms for Allocation of Research Grants'. *Science, Technology, & Human Values*, January, 0162243918822744. https://doi.org/10.1177/0162243918822744.

Rubini, Monica, and Michela Menegatti. 2014. 'Hindering Women's Careers in Academia: Gender Linguistic Bias in Personnel Selection'. *Journal of Language and Social Psychology* 33 (6): 632–50. https://doi.org/10.1177/0261927X14542436.

Sabatier, M., M. Carrere, and V. Mangematin. 2006. 'Profiles of Academic Activities and Careers: Does Gender Matter? An Analysis Based on French Life Scientist CVs'. *The Journal of Technology Transfer* 31 (May): 311–24. https://doi.org/10.1007/s10961-006-7203-3.

Sackett, Paul R., Cathy L. DuBois, and Ann W. Noe. 1991. 'Tokenism in Performance Evaluation: The Effects of Work Group Representation on Male-Female and White-Black Differences in Performance Ratings'. *Journal of Applied Psychology* 76 (2): 263–67. https://doi.org/10.1037/0021-9010.76.2.263.

Salerno, PE, M Paez-Vacas, JM Guayasamin, and L Stynoski. 2019. 'Male Principal Investigators (Almost) Don't Publish with Women in Ecology and Zoology'. *PLoS ONE* 14 (6): e0218598. https://doi.org/10.1371/journal.pone.0218598.

Sandström, Ulf. 2009a. 'Combining Curriculum Vitae and Bibliometric Analysis: Mobility, Gender and Research Performance'. *Research Evaluation* 18 (2): 135–42. https://doi.org/10.3152/095820209X441790.

———. 2009b. 'Research Quality and Diversity of Funding: A Model for Relating Research Money to Output of Research'. *Scientometrics* 79 (2): 341–49. https://doi.org/10.1007/s11192-009-0422-2.

Sandström, Ulf, and Martin Hällsten. 2008. 'Persistent Nepotism in Peer-Review'. *Scientometrics* 74 (2): 175–89. https://doi.org/10.1007/s11192-008-0211-3.

Sanz-Menéndez, Luis, Laura Cruz-Castro, and Kenedy Alva. 2013. 'Time to Tenure in Spanish Universities: An Event History Analysis'. Edited by Daniele Fanelli. *PLoS ONE* 8 (10): e77028–e77028. https://doi.org/10.1371/journal.pone.0077028.

Scott, W. Richard. 1965. 'Reactions to Supervision in a Heteronomous Professional Organization'. *Administrative Science Quarterly* 10 (1): 65–81. https://doi.org/10.2307/2391650.

Sekaquaptewa, Denise, Koji Takahashi, Janet Malley, Keith Herzog, and Sara Bliss. 2019. 'An Evidence-Based Faculty Recruitment Workshop Influences Departmental Hiring Practice Perceptions among University Faculty'. *Equality, Diversity and Inclusion: An International Journal* 94 (7): 52. https://doi.org/10.1108/EDI-11-2018-0215.

Severin, Anna, João Martins, François Delavy, Anne Jorstad, and Matthias Egger. 2019. 'Potential Bias in Peer Review of Grant Applications at the Swiss National Science Foundation'. e27587v2. PeerJ Inc. https://doi.org/10.7287/peerj.preprints.27587v2.

Shatz, David. 2004. *Peer Review: A Critical Inquiry*. Rowman & Littlefield Publishers. https://rowman.com/ISBN/9780742514348/Peer-Review-A-Critical-Inquiry.

Sheltzer, Jason M., and Joan C. Smith. 2014. 'Elite Male Faculty in the Life Sciences Employ Fewer Women'. *Proceedings of the National Academy of Sciences of the United States of America* 111 (28): 10107–12. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/23805713.

Shen, H.-W., and A.-L. Barabasi. 2014. 'Collective Credit Allocation in Science'. *Proceedings of the National Academy of Sciences* 111 (34): 12325–30. https://doi.org/10.1073/pnas.1401992111.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Shen, Yiqin Alicia, Jason M. Webster, Yuichi Shoda, and Ione Fine. 2018. 'Persistent Underrepresentation of Women's Science in High Profile Journals'. *BioRxiv*, March, 275362. https://doi.org/10.1101/275362.

Sieverding, Monika, Constanze Eib, Andreas B. Neubauer, and Thomas Stahl. 2018. 'Can Lifestyle Preferences Help Explain the Persistent Gender Gap in Academia? The "Mothers Work Less" Hypothesis Supported for German but Not for U.S. Early Career Researchers'. *PLOS ONE* 13 (8): e0202728. https://doi.org/10.1371/journal.pone.0202728.

Smith-Doerr, Laurel. 2004. 'Flexibility and Fairness: Effects of the Network Form of Organization on Gender Equity in Life Science Careers'. *Sociological Perspectives* 47 (1): 25–54. https://doi.org/10.1525/sop.2004.47.1.25.

Sonnert, Gerhard, and Gerald James Holton. 1995. *Who Succeeds in Science?: The Gender Dimension*. Rutgers University Press.

Stack, Steven. 2004. 'Gender, Children and Research Productivity'. *Research in Higher Education* 45 (8): 891–920. http://www.jstor.org.ezproxy1.lib.asu.edu/stable/40197370.

Steinpreis, Rhea E., Katie A. Anders, and Dawn Ritzke. 1999. 'The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study'. *Sex Roles* 41 (7): 509–28. https://doi.org/10.1023/A:1018839203698.

Stephan, Paula E., and Asmaa El-Ganainy. 2007. 'The Entrepreneurial Puzzle: Explaining the Gender Gap'. *The Journal of Technology Transfer* 32 (5): 475–87. https://doi.org/10.1007/s10961-007-9033-3.

Stewart, Abigail J., and Virginia Valian. 2018. *An Inclusive Academy.Achiving Diversity and Excellence*. Cambridge Mass.: The MIT Press. https://mitpress.mit.edu/books/inclusive-academy.

Stewart, Penni, Michael Ornstein, and Janice Drakich. 2009. 'Gender and Promotion at Canadian Universities'. *Canadian Review of Sociology/Revue Canadienne de Sociologie* 46 (February): 59–85. https://doi.org/10.1111/j.1755-618X.2009.01203.x.

Su, Xuhong, Japera Johnson, and Barry Bozeman. 2015. 'Gender Diversity Strategy in Academic Departments: Exploring Organizational Determinants'. *Higher Education* 69 (5): 839–58. https://doi.org/10.1007/s10734-014-9808-z.

Symonds, Matthew R. E., Neil J. Gemmell, Tamsin L. Braisher, Kylie L. Gorringe, and Mark A. Elgar. 2006. 'Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance'. *PLOS ONE* 1 (1): e127. https://doi.org/10.1371/journal.pone.0000127.

Tamblyn, Robyn, Nadyne Girard, Christina J. Qian, and James Hanley. 2018. 'Assessment of Potential Bias in Research Grant Peer Review in Canada'. *CMAJ : Canadian Medical Association Journal* 190 (16): E489–99. https://doi.org/10.1503/cmaj.170901.

Taylor, Susan Washburn, Blakely Fox Fender, and Kimberly Gladden Burke. 2006. 'Unraveling the Academic Productivity of Economists: The Opportunity Costs of Teaching and Service'. *Southern Economic Journal* 72 (4): 846–59. https://doi.org/10.2307/20111856.

Thelwall, Mike. 2018. 'Do Females Create Higher Impact Research? Scopus Citations and Mendeley Readers for Articles from Five Countries'. *Journal of Informetrics* 12 (4): 1031–41. https://doi.org/10.1016/j.joi.2018.08.005.

Thun, Cecilie. 2019. 'Excellent and Gender Equal? Academic Motherhood and "Gender Blindness" in Norwegian Academia'. *Gender, Work & Organization* 0 (0). https://doi.org/10.1111/gwao.12368.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Timmers, Tanya M., Tineke M. Willemsen, and Kea G. Tijdens. 2010. 'Gender Diversity Policies in Universities: A Multi-Perspective Framework of Policy Measures'. *Higher Education* 59 (6): 719–35. https://doi.org/10.1007/s10734-009-9276-z.

Tolbert, Pamela S., and Alice A. Oberfield. 1991. 'Sources of Organizational Demography: Faculty Sex Ratios in Colleges and Universities'. *Sociology of Education* 64 (4): 305–15. https://doi.org/10.2307/2112710.

Tolbert, Pamela S., Tal Simons, Alice Andrews, and Jaehoon Rhee. 1995. 'The Effects of Gender Composition in Academic Departments on Faculty Turnover'. *Industrial and Labor Relations Review* 48 (3): 562–79. https://doi.org/10.2307/2524782.

Treviño, Len J., Luis R. Gomez-Mejia, David B. Balkin, and Franklin G. Mixon. 2018. 'Meritocracies or Masculinities? The Differential Allocation of Named Professorships by Gender in the Academy'. *Journal of Management* 44 (3): 972–1000. https://doi.org/10.1177/0149206315599216.

Tricco, Andrea C., Sonia M. Thomas, Jesmin Antony, Patricia Rios, Reid Robson, Reena Pattani, Marco Ghassemi, et al. 2017. 'Strategies to Prevent or Reduce Gender Bias in Peer Review of Research Grants: A Rapid Scoping Review'. *PLoS ONE* 12 (1): 1–12. https://doi.org/10.1371/journal.pone.0169718.

Van den Brink, Marieke, and Yvonne Benschop. 2014. 'Gender in Academic Networking: The Role of Gatekeepers in Professorial Recruitment'. *Journal of Management Studies* 51 (3). https://doi.org/10.1111/joms.12060.

Van den Brink, Marieke, and Lineke Stobbe. 2014. 'The Support Paradox: Overcoming Dilemmas in Gender Equality Programs'. *Scandinavian Journal of Management* 30 (2): 163–74. https://doi.org/10.1016/j.scaman.2013.07.001.

Vernos, Isabelle. 2013. 'Quotas Are Questionable'. *Nature* 495 (March): 39. http://dx.doi.org/10.1038/495039a.

Vinkenburg, Claartje J. 2017. 'Engaging Gatekeepers, Optimizing Decision Making, and Mitigating Bias: Design Specifications for Systemic Diversity Interventions'. *The Journal of Applied Behavioral Science* 53 (2): 212–34. https://doi.org/10.1177/0021886317703292.

Voorspoels, Jolien, and Inge Bleijenbergh. 2019. 'Implementing Gender Quotas in Academia'. *Equality, Diversity and Inclusion: An International Journal* 55 (3): 359. https://doi.org/10.1108/EDI-12-2017-0281.

Waaijer, Cathelijn J. F., Hans Sonneveld, Simone E. Buitendijk, Cornelis A. van Bochove, and Inge C. M. van der Weijden. 2016. 'The Role of Gender in the Employment, Career Perception and Research Performance of Recent PhD Graduates from Dutch Universities'. Edited by Aurora García-Gallego. *PLOS ONE* 11 (10): e0164784. https://doi.org/10.1371/journal.pone.0164784.

Waaijer, Cathelijn J. F., Christine Teelken, Paul F. Wouters, and Inge C. M. van der Weijden. 2018. 'Competition in Science: Links Between Publication Pressure, Grant Pressure and the Academic Job Market'. *Higher Education Policy* 31 (2): 225–43. https://doi.org/10.1057/s41307-017-0051-y.

Wahls, Wayne P. 2018. 'The NIH Must Reduce Disparities in Funding to Maximize Its Return on Investments from Taxpayers'. *ELife* 7. https://doi.org/10.7554/eLife.34965.

Ward, Melanie E. 2001. 'Gender and Promotion in the Academic Profession'. *Scottish Journal of Political Economy* 48 (3): 283–302. https://doi.org/10.1111/1467-9485.00199.

Weisshaar, Katherine. 2017. 'Publish and Perish? An Assessment of Gender Gaps in Promotion to Tenure in Academia'. *Social Forces* 96 (2): 529–60. https://doi.org/10.1093/sf/sox052.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

Wennerås, Christine, and Agnes Wold. 1997. 'Nepotism and Sexism in Peer-Review'. *Nature* 387 (May): 341–343. https://doi.org/10.1038/387341a0.

White, Harrison C. 1982. 'Fair Science?' Edited by Jonathan R. Cole. *American Journal of Sociology* 87 (4): 951–56. https://www.jstor.org/stable/2778789.

Whitley, Richard. 2011. 'Changing Governance and Authority Relations in the Public Sciences'. *Minerva* 49 (4): 359–385. https://doi.org/10.1007/s11024-011-9182-2.

Williams, Wendy M., and Stephen J. Ceci. 2015. 'National Hiring Experiments Reveal 2:1 Faculty Preference for Women on STEM Tenure Track'. *Proceedings of the National Academy of Sciences* 112 (17): 5360–65. https://doi.org/10.1073/pnas.1418878112.

Witteman, Holly O, Michael Hendricks, Sharon Straus, and Cara Tannenbaum. 2019. 'Are Gender Gaps Due to Evaluations of the Applicant or the Science? A Natural Experiment at a National Funding Agency'. *The Lancet* 393 (10171): 531–40. https://doi.org/10.1016/S0140-6736(18)32611-4.

Wolfinger, Nicholas H., Mary Ann Mason, and Marc Goulden. 2008. 'Problems in the Pipeline: Gender, Marriage, and Fertility in the Ivory Tower'. *The Journal of Higher Education* 79 (4): 388–405. https://doi.org/10.1353/jhe.0.0015.

Xie, Yu, Michael Fang, and Kimberlee Shauman. 2015. 'STEM Education'. *Annual Review of Sociology* 41 (1): 331–57. https://doi.org/10.1146/annurev-soc-071312-145659.

Xie, Yu, and Kimberlee A. Shauman. 1998. 'Sex Differences in Research Productivity: New Evidence about an Old Puzzle'. *American Sociological Review* 63 (6): 847–70. https://doi.org/10.2307/2657505.

———. 2003. *Women in Science : Career Processes and Outcomes*. Cambridge Mass.: Harvard University Press.

Yazilitas, Demet, Jörgen Svensson, Geert Vries, and Sawitri Saharso. 2013. 'Gendered Study Choice'. *Educational Research and Evaluation* 19 (6): 525–45. https://doi.org/10.1080/13803611.2013.803931.

Zeng, Xiao Han T., Jordi Duch, Marta Sales-Pardo, João A. G. Moreira, Filippo Radicchi, Haroldo V. Ribeiro, Teresa K. Woodruff, and Luís A. Nunes Amaral. 2016. 'Differences in Collaboration Patterns across Discipline, Career Stage, and Gender'. *PLOS Biology* 14 (11): e1002573. https://doi.org/10.1371/journal.pbio.1002573.

Zinovyeva, Natalia, and Manuel Bagues. 2015. 'The Role of Connections in Academic Promotions'. *American Economic Journal: Applied Economics* 7 (2): 264–92. https://doi.org/10.1257/app.20120337.

Zuckerman, Harriet. 2001. 'The Careers of Men and Women Scientists: Gender Differences in Career Attainment'. In *Women, Science, and Technology : A Reader in Feminist Science Studies*, edited by Mary Wyer, Mary Barbercheck, Donna Cookmeyer, Hatice Ozturk, and Marta Wayne, 69–78. Routledge. https://doi.org/10.4324/9780203427415.

Zuckerman, Harriet, and Jonathan R. Cole. 1975. 'Women in American Science'. *Minerva* 13 (1): 82–102. https://doi.org/10.1007/BF01096243.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.

HISTORY OF CHANGES		
Version	**Publication date**	**Change**
1.0	09.08.2019	▪ Final draft version
2.0	8.11.2019	▪ Final version

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824574.