

Gender differences in the aims and impacts of research

Lin Zhang^{1,2*}, Gunnar Sivertsen³, Huiying Du⁴, Ying Huang^{2,1}, Wolfgang Glänzel²

¹ School of Information Management, Wuhan University, Wuhan (China)

² Centre for R&D Monitoring (ECOOM) and Department MSI, KU Leuven, Leuven (Belgium)

³ Nordic Institute for Studies in Innovation, Research and Education (NIFU), Oslo (Norway)

⁴ North China University of Water Resources and Electric Power, Zhengzhou (China)

Abstract

This study uses mixed methods – classical citation analysis, altmetric analysis, a survey with researchers as respondents, and text analysis of the abstracts of scientific articles – to investigate gender differences in the aims and impacts of research. We find that male researchers more often value and engage in research mainly aimed at scientific progress, which is more cited. Female researchers more often value and engage in research mainly aimed at contributing to societal progress, which has more abstract views (usage). The gender differences are observed among researchers who work in the same field of research and have the same age and academic position. Our findings have implications for evaluation and funding policies and practices. A critical discussion of how societal engagement versus citation impact is valued, and how funding criteria reflect gender differences, is warranted.

Keywords: Gender; citation impact; altmetrics; usage; societal impact; research performance; research evaluation.

Part 1. General introduction

This study uses mixed methods – classical citation analysis, altmetric analysis, a survey with researchers as respondents, and text analysis of the abstracts of scientific articles – to investigate gender differences in the aims of research projects and their internal scientific and external societal impact. We find that:

- Research mainly aimed at contributing to scientific progress is more often cited in the scientific literature (citation impact) while research mainly aimed at contributing to societal progress is more often read online as available abstracts (usage).
- Male researchers more often value and engage in research mainly aimed at scientific progress. The citation impact of their publications is higher than for female researchers.
- Female researchers more often value and engage in research mainly aimed at contributing to societal progress. The usage of their publications is higher than for male researchers.

The differences are small but significant. We did not expect large differences for two reasons. One is that most of the scientific publications under study were published in collaboration by female and male researchers. We need to distinguish them by the gender of the first author to find the differences. The other reason is that both forms of aims and both forms of impact can be present in one and the same publication.

By investigating the relations between gender, two forms of aims, and two forms of impact, we are contributing to – and connecting for the first time – two classical research agendas in quantitative science

* linzhang1117@whu.edu.cn

studies: studies of gender and research performance, and studies of the relation between the impact of research and its aims and types. Studies of *gender and research performance* were initiated by Harriet Zuckerman and Jonathan R. Cole (1975) and have mainly focused on possible explanations for the lower productivity and/or citation impact of female researchers. Studies of the relation between the *impact of research and its aims and types* were initiated in the 70's by Eugene Garfield (1979) who noted that publications from basic research were more highly cited than publications from applied research.

We are also contributing to a third less developed research agenda, *studies of gender differences in motivations for choosing aims and type of research*. While several studies have focused on gender differences in career choices of discipline or field of research and possible explanations for these differences, we find few previous studies of gender differences in the choice of aims and type of research within the same field.

Following our mixed-methods approach, we will present reviews of previous research, methods, and results in three different sections before we unite them in a final discussion with conclusions at the end. Part 2 is titled *Citation impact and usage*. Here, we present the results of using citation and altmetric indicators. Part 3 is titled *Aims of the research*. It presents the results of a text analysis of almost 1,200 abstracts of publications with which we classify research projects by aims. Part 4 is titled *Motivations for the choice of aims*. Here, we use responses from a large survey among researchers in three different fields of research in five countries to measure to what degree they value scientific progress and societal progress when they express their motivations for doing research and their criteria for valuing good research.

Part 2a. Citation impact and usage: Introduction

Several studies have found that articles published by women are less cited than those published by men (Bendels et al., 2018; Huang et al., 2020). A global and cross-disciplinary bibliometric study concluded that the gender disparity in citation impact holds for both internationally and nationally co-authored publications, but it can to some extent be explained by male dominance in scientific production (Larivière et al., 2013). Studies at the national level came to similar conclusions. Based on a large-scale study of 8,500 Norwegian researchers and more than 37,000 publications covering all areas of research, Aksnes et al. (2011) found that publications by female researchers are less cited but the differences are not large and can be attributed to differences in productivity (see also Larivière & Costas, 2016). Productivity differences can, in turn, be attributed to differences in childcare responsibilities (Kyvik & Teigen, 1996), career trajectories (van den Besselaar & Sandström, 2016), or participation in international collaboration (Aksnes et al., 2019). Studies focused on India (Paswan & Singh, 2020), Canada (Beaudry & Lariviere, 2016), and Germany (Pudovkin et al., 2012) also observed gender disparity in citation impact in medical sciences and other research disciplines. Ghiasi et al. (2015) found that even though female engineers publish in journals with higher Impact Factors, they receive less citations than their male peers. Similar observations have been made in the fields of geography (Rigg, McCarragher, & Krmenec, 2012), international relations (Maliniak, Powers, & Walter, 2013), medical sciences (Nielsen, 2016) and astronomy (Caplar, Tacchella, & Birrer, 2017).

Some other studies did not find gender differences in citation impact (Barrios et al., 2012; Bosquet & Combes, 2013). Similarly, based on publications by active authors during the period 2014–2018, Elsevier's 2020 gender report (Elsevier, 2020) shows that among the countries studied, the average field-

weighted citation impact (FWCI) for men compared to women when assessing all authors, regardless of authorship position, was almost equal in all countries and the EU28. Only when considering the authorship position, the average FWCI of male first authors is higher than that of female first authors (cf. Elsevier, 2020).

Some recent studies report higher citation impact among female authors. A large-scale study conducted by Thelwall et al. (2019) analyzed six million articles published during 1996–2018 from seven large English-speaking nations. After using the mean-normalized logarithmic citation score to compare the citation impact of articles, the results showed that females have a citation advantage over the complete period 1996–2014 and in all seven countries except the USA. This observation holds after a closer analysis at field level (Thelwall, 2020a) and of the most highly cited articles (Thelwall, 2020b). However, it was found that team size can influence gender differences (Thelwall & Sud, 2020).

Studies of gender in relation to impact as measured by altmetric indicators are less abundant. Compared to traditional citation analysis, altmetric analysis covers impact both within and beyond formal scholarly communication, including abstract views, publication downloads, Mendeley reads, social media attention, and other measures of the broader impact of research outputs (Bornmann, Haunschild, & Adams, 2019). Thelwall and Kousha (2014) focused on *Academia.edu*, a social networking website for academics, and investigated profile views of scholars belonging to four disciplines (law, history, computer science, and philosophy). They found a female advantage that is suggestive of general social networking norms in law, history, and computer science. Paul-Hus et al. (2015) showed that a more gender-balanced situation can be found in social media metrics (news, tweets, blogs) than in traditional citation metrics. Based on an analysis of Mendeley readership, Thelwall (2018a) reported a wider audience for female-authored research.

Our results here in part 2 will confirm most of the studies referred above before we proceed to investigate possible explanations for the gender differences in parts 3 and 4.

Part 2b. Citation impact and usage: Data and methods

Our study of citation impact is based on publications and their citations within Web of Science (WoS). For the study of usage of the same publications, we used their DOI (Digital Object Identifier) to match them with *PlumX Metrics*, a set of altmetric indicators provided by Plum Analytics. Within this set, we selected *abstract views* as an indicator of usage (see below).

A limitation of altmetrics is that the indicators measure the broader attention given to *scientific publications* only (Haustein, 2014), not the full range of science-society interactions (Sivertsen & Meijer, 2020). Such interactions need to be studied with the use of data sources reaching beyond written communication (Bornmann, 2013). Still, our study confirms the expectation that altmetrics can supplement traditional citation analysis for the understanding and evaluation of the influences of scientific publications (Bornmann, 2014).

For our purposes, it is necessary to know not only the gender of authors but also to control for other variables such as age, academic position, and field of research. For this reason, we chose to use a data source representing Norway where these variables are available at the individual level for all researchers in the country's public sector. We use this national data source in combination with two international data sources that provide citation indicators and altmetric indicators:

1. *The Norwegian Science Index (NSI)* is a subset of the Current Research Information System in Norway (Cristin). It has complete coverage since 2011 of all peer-reviewed scientific and scholarly publication outputs from Norwegian institutions. The database also records the gender and year of birth of all researchers affiliated with Norwegian research organizations. For these persons, there is no need of disambiguation of author names. The metadata for publications can be matched to the two other databases.
2. Representing *Web of Science (WoS)*, we use the *National Citation Report for Norway (NCR, 1981-2018)*, a data set provided by Clarivate Analytics with a representation of all articles in WoS with minimum one address in Norway, and their accumulated citation counts. This data set has the same 254 subject categories that are used to assign journals in WoS, and the same indicators that are present in the Web of Science-based bibliometric product *InCites* from Clarivate Analytics. NCR has all the basic data for each publication as well.
3. *The PlumX Metrics*, which provides insights into the ways people interact with individual pieces of research output in the online environment. It categorizes metrics into five separate categories: Citations, Usage, Captures, Mentions, and Social Media. Metadata for publications can be matched to WoS via the Digital Object Identifier (DOI).

As a first step, we selected 26,976 *journal articles* in NCR from 2011-2017 by using three criteria: Firstly, at least one of the authors were affiliated with one of Norway's four largest universities (Oslo, Bergen, Trondheim, Tromsø). The time and resources to perform research are equal and comparable among these four institutions. Secondly, the publications could be linked to altmetric indicators in *PlumX Metrics* through their DOI number in NCR or NSI. Data for each publication was retrieved from *PlumX Metrics* by the end of February 2020. The third criterium was to include only publications with Norwegian first authors (hereafter: *1st authors*) that can be identified as persons in NSI. Most of the publications in our data set are multi-authored with a representation of both female and male researchers. As in other gender studies using bibliometric data (e.g. Thelwall, 2018a; 2018b), we chose to focus on 1st authors because they often take most of the responsibility for the specific project. However, this may not be the case when authors are presented in alphabetical order. We checked: Only five percent of the articles with three or more authors in our data set present them in alphabetical order.

The third selection criterium resulted in a total of 11,725 *identifiable persons* as the 1st authors. Among these, 6,678 (57%) are men and 5,047 (43%) are women. Table 1 shows how their publications are distributed by gender, age, and major area of research. In general, 40.3 percent of the publications are by female 1st authors. The classification of publications in major areas of research was done according to the journal-based classification of publications in NSI, which consists of four major areas and 84 subfields. Publications are assigned to only one major area each. This classification per publication is only used in part 2 of our study where we focus on citation impact versus usage. Here, normalization of the measure of impact is important at the level of publications. In parts 3 and 4 below, where we focus on the researchers, all publications by a specific 1st author are classified within only one area of research. If the publications by a 1st author are in different areas of research, we use their organizational affiliation to determine the classification (e.g. a department of sociology would assign the researcher's publications to the category Social Sciences). In most cases, however, the area of research of a publication is the same as the area of research of its 1st author.

As in a previous study (Zhang & Sivertsen, 2017), we included age and academic position in the first round of investigations reported below. This time, we found that the age and the academic position of the 1st authors do not differ much in their influence on the citation and usage indicators. They seem to represent experience and seniority in research in the same way. We favoured age as an independent variable over academic position for two reasons. Age is often used to explain differences in altmetric studies. Moreover, it is possible to study age more dynamically than position in our data set. We know when each researcher was born and can relate this information to the year of the publication (for example, if a researcher was born in 1980, then his/her age was 35 years when publishing in 2015). As seen in Table 1, which shows the number of publications under study by area of research and the gender and age group of the 1st author, we have subdivided publications into two age groups (21–45, 46+) according to the age of the 1st author when publishing.

Table 1. The number of publications under study by area of research and the gender and age group of the 1st author

Age group	Health Sciences		Natural Sciences and Engineering		Social Sciences		Humanities	
	Female	Male	Female	Male	Female	Male	Female	Male
21-45	4579	3666	3027	7118	457	577	167	214
46+	1933	1991	302	1803	281	593	112	156
Totals	6512	5657	3329	8921	738	1170	279	370
Gender balance	54%	46%	27%	73%	39%	61%	43%	57%

To study the relations between the groups of publications as shown in Table 1 and different forms of impact, we started with four forms of impact and then continued focusing on two of them, citation impact and usage. We will, however, start the analysis below by showing the results for all four of them:

1. The ‘Mendeley’ indicator is based on the number of readers a paper has had in Mendeley.
2. The ‘SocialMedia’ indicator represents the number of times a publication is referred to in Twitter or Facebook. The two frequencies are summed up.
3. The ‘Usage’ indicator represents the frequency of abstract views.
4. The ‘WoSCit’ indicator represents field-normalized citations within Web of Science

Among these, the first three are altmetric indicators that we selected after observing that some indicators among the *PlumX Metrics* have too low indications of activity to be reliable and valid for meaningful statistical analysis. We also decided to exclude the ‘full-text view’ indicator of usage since it heavily depends on full-text availability and can be biased towards open-access publications. Our indicator is thereby based on recorded *abstract views* only.

The WoSCit indicator of scientific impact could be constructed from citation indicators in the NCR database. We used citations within WoS, normalized by WoS subject category and year, and expressed as percentiles after ranking the articles from the most to the least cited. Journals in more than one subject category were assigned to the category with the highest average citations per article. The percentile indicator in NCR is compared to all articles in the world in the same subfield and year.

We adopted the same *percentile* method for the indicators based on PlumX data. Here, the basis for the normalization is 80,466 articles in NCR with at least one author affiliated with the four largest Norwegian universities. This larger dataset of articles provided reference values for normalization of the altmetric indicators in our primary dataset of articles by identified 1st authors. We normalized by the year and major area of research of the *publication*, using the NSI classification with four major areas and 84 subfields mentioned above. Within each major area of research, we selected and tagged the top 10 percent publications according to impact in each of the three categories ‘Mendeley’, ‘Usage’ and ‘Social Media’.

We use the percentile-rank (PR) method (Bornmann et al., 2013) because citation distributions are known to be highly skewed (Seglen, 1992). In large datasets, the mean number of citations per publication is typically 20 points over the median, and one tenth of the publications will receive almost half of all citations (Albarrán et al., 2011). A large proportion of the publications will seldom or never be cited. A similar skewness has been observed for altmetric indicators (e.g., Chi et al., 2019; Glänzel and Chi, 2020). We find the same for our three altmetric indicators. Table 2 demonstrates this skewness by using 3,908 publications in the Health Sciences and the year 2014 as an example. As we can see, mentions of scientific publications in social media are particularly skewed among publications. The observed skewness of all four indicators is one of our reasons for choosing percentile-based impact indicators. Comparing with averages would give more weight to extreme scores. Also because of the observed skewness, we have chosen a dichotomous rather than a continuous measure. Table 2 shows that there is little variance among publications scoring below the top 10 percent. The 10 percent indicator was chosen because it captures our observations in Table 2 and because it is widely used and known, e.g. from the CWTS Leiden Ranking of universities (leidenranking.com).

Table 2. Skewness of impact according to different indicators for articles in Health Sciences, publication year of 2014

	WoSCit	Mendeley	Usage	SocialMedia
Maximum score among top 10 percent	201	1137	7990	950
Minimum score among top 10 percent	10	182	977	13
Percentage publications scoring zero	4.2 %	2.8 %	2.2 %	44.6 %

The percentile standardization values were calculated for each indicator to normalize the impact of articles from different subject fields and publication year. The lower the value of the percentile-rank is, the higher impact has the publication. For example, articles with a percentile (in WoSCit) equal to or lower than 10% indicate that they are among the top 10% most cited in their subject field and publication year. The same applies to the other indicators. Field normalization of the impact indicators allows us to make comparisons even across areas of research. For the usage indicator, we especially checked whether gender differences were consistent also at subfield level within the same area of research.

For each of the four indicators, we measure the share of publications by female 1st authors versus male 1st authors among the top 10 percent high impact publications compared to (divided by) the shares among *all* publications (high impact or not). Values above, below, or equal to 1 will thereby show

whether the share of publications by female 1st authors versus male 1st authors is over, below, or equal to its general share in all publications. We will give an example in connection to Figure 1 below.

The analysis will be performed in three steps, firstly for all areas of research combined, then for age groups in all areas combined, and thirdly for age groups within each area of research. Since the publications in our data are placed in one category only, we can aggregate without duplicating.

Before we start, Table 3 presents the Pearson correlation between the four impact indicators under study. The strongest correlation was found between WoSCit and Mendeley, while the weakest correlation is between WoSCit and Usage. This observation is by and large in line with those made by Glänzel and Chi (2020) and indicative of why we found usage to be the most distinctive and interesting altmetric indicator in our results.

Table 3. Pearson correlation between four indicators for the 26,976 documents

	Percentile WoSCit	Percentile Usage	Percentile Mendeley	Percentile SocialMedia
Percentile WoSCit	1	0.078**	0.447**	0.106**
Percentile Usage		1	0.339**	0.162**
Percentile Mendeley			1	0.300**
Percentile Social Media				1

We have used the chi-square to test the significance of the gender differences observed in our results in the following. These tests were conducted using SPSS 24.0.

Part 2c. Citation impact and usage: Results

Figure 1 presents the results for each of the four different indicators in all areas combined. Taking *usage* as example, the share of publications by female 1st authors among top impact publications is 51.8%, which is clearly higher than the general share of publications by female 1st authors among all publications, which is 40.2%. The corresponding value on the vertical axis is therefore 1.29 ($\approx 0.518/0.402$).

In general, publications by male 1st authors are relatively more frequent among publications with high *citation impact* while publications by female 1st authors are relatively more frequent among publications with high *usage*. Publications by female 1st authors also perform slightly better by the Mendeley and Social Media indicators. Although these findings seem to confirm earlier studies by Thelwall (2018a; 2018b), our results from using the two latter altmetric indicators did not pass the significance test. Hence, in the following, we will concentrate on the indicators of citation impact and usage.

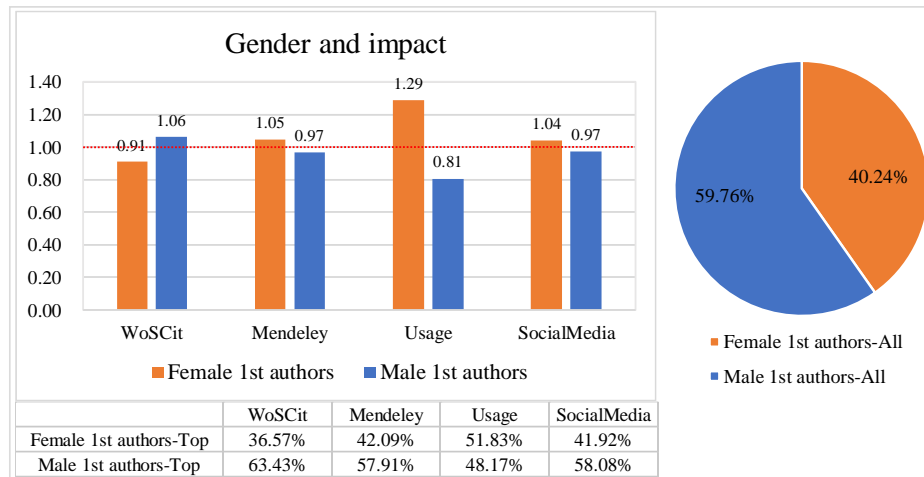


Figure 1. Gender and four forms of impact. The shares in top 10% impact publications are compared to (divided by) the general shares in all publications (to the right). Values around 1.00 indicate the degree to which the share in top 10% impact publications is higher/lower than in all publications.

In combination with gender, Figure 2 shows the influence on of the age of the 1st author when publishing the citation impact and usage of publications. The general results in Figure 1 are confirmed within each age group: Publications by male 1st authors are overrepresented among publications with high impact while publications by female 1st authors are overrepresented among publications with usage. Citation impact and usage are higher for publications with older 1st authors. Younger researchers do not score higher on the altmetric indicator. Most remarkable is the high usage impact for publications by female 1st authors in the older age group.

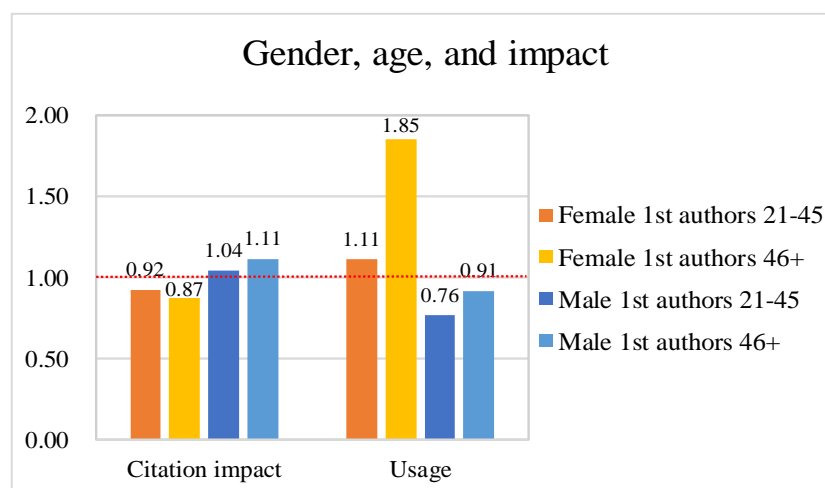


Figure 2. Gender, age, and citation impact versus usage. See Figure 1 for explanation of the indicator.

As seen in Table 1, the gender balance in the shares of publications differ among the areas of research, reflecting for example that there are relatively more female 1st authors in the Health Sciences than in the Natural Sciences and Engineering. Different career choices regarding field of research might explain gender differences in the aims and impacts of research. We checked area of research as an influencing

factor. The results are shown in Figure 3 where we keep the age groups (as in Figure 2) in the analysis per main area of research.

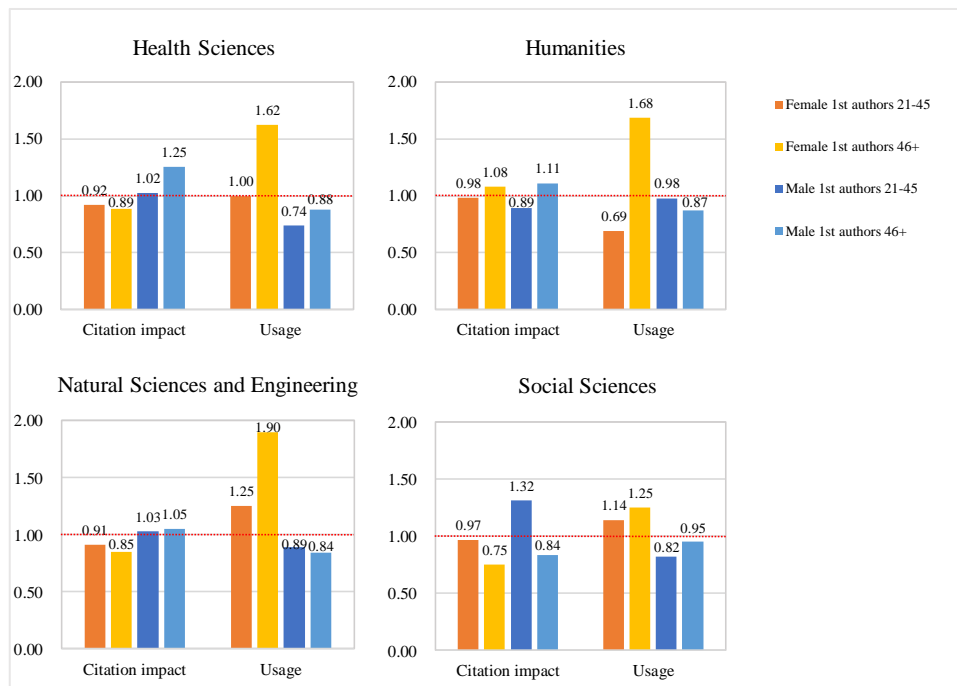


Figure 3. Gender, age and citation impact versus usage by area of research. See Figure 1 for explanation of the indicator.

The main gender differences in citation impact and usage prevail in all four areas of research. The results are mainly consistent in the Health Sciences and the Natural Sciences and Engineering where most of the publications under study belong (see Table 1 above). The most notable difference from the general pattern is the higher citation impact of the group of younger researchers in the Social Sciences.

We also checked the situation at the subfield level. Even in most of the subfields with relatively few female 1st authors, we find the general pattern of female overrepresentation in publications with high usage impact and male overrepresentation in publications with high citation impact. Examples in our data of *male dominated subfields with higher usage impact of publications by female 1st authors* are: Civil Engineering, Economics, Electronics, Energy, History, Philosophy, Physics, Sports sciences, Marine technology, and Theology and religion. One of the few exceptions is Business and Finance, where publications by male 1st authors have slightly higher citation *and* usage impact than female 1st authors.

Our main and most significant result from the study of citation impact and usage is that publications by male 1st authors are relatively more frequent among publications with high citation impact, while publications by female 1st authors are relatively more frequent among publications with high usage. The age of 1st authors when publishing or the field of research of the publications could not explain the difference. We now return to our hypothesis that the differences can be explained by the aims of the research.

Part 3a. Aims of the research: Introduction

In our review of the literature in part 2a, we referred to several possible explanations for observed gender differences in the impact of publications as measured by citation indicators or altmetric indicators. Our hypotheses for the investigations in this part of the study are not meant to exclude other possible explanations, only to supplement them with a new idea:

- Male researchers are overrepresented as 1st authors of publications with high citation impact because they more often engage in research mainly aimed at *scientific progress*.
- Female researchers are overrepresented as 1st authors of publications with high usage because they more often engage in research mainly aimed at *societal progress*.

We arrived at these hypotheses after reading a random sample of abstracts of 100 publications among those under study. After having excluded field of research as an explanatory factor in the analysis in part 2 above, we expected to find differences in research topics within the field of research. Instead, we found indications of a possible difference at a more general level: *the aims of the research*. Without excluding the possibility that there are gender differences in chosen topics within fields of research, we established the above hypotheses.

These hypotheses brought us to a second classical strand of relevant research, classification of publications in large datasets by their contents. There is a rich literature and an array of methods in quantitative science studies on how to classify large datasets of scientific publications by their contents. The methods vary from simple journal classifications and keyword counts to advanced use of citation relations and trained machine learning. To classify publications by their *aims*, not their contents, is still demanding. Similar general aims are not easy to identify among publications by using topical words, shared references or standard phrases.

There is, however, a strand of research that has tried to deal with a similar problem. Francis Narin and colleagues (1976) created a four-level classification of the journals covered by the *Web of Science* at that time which represented different degrees of *basic research versus applied research*. Following Narin's idea, Boyack et al. (2014) used the classification scheme and developed it further to be applied at article level independently of journals. They used trained machine learning to read title, abstract words and cited references for the classification of millions of articles and classified them according to the same four levels of research. Following up, Donner and Schmoch (2020) applied the model to measure differences in citation impact. They got clearer results by reducing the model to only two levels, basic or applied, and found that articles from basic research are cited more frequently than articles from applied research. The study thereby confirmed the original observation made by Garfield (1979) and several subsequent studies focusing on specific areas of research, e.g. clinical research (van Eck et al., 2013).

Valuable for our study is the awareness that these previous studies create towards the relation between types and aims of research on the one hand and citation impact on the other. We deviate from the previous studies by not using the distinction between basic and applied research. Instead, we distinguish between research mainly aimed at scientific progress and research mainly aimed at societal progress. We do so because the traditional basic/applied distinction is not in accordance with our understanding of how research is performed in science-society interactions and because our own distinction is easier to define

operationally for classifying the aims of research projects as expressed in the abstracts of scientific publications. Explaining these two reasons will lead to the operational definitions of our new distinction.

The classification of research in basic and applied types, although still widely used in research policy to describe the roles of organizations and how resources are spent, originates from an understanding of the workflow in science that was influential more than half a century ago. The distinction was first defined in 1963 in the first edition of OECD's *Frascati Manual* with guidelines for collecting and reporting data on R&D activities and spending. Following the tradition since 1963, the latest edition (OECD, 2015) defines not only two, but three general types of research and development:

Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without any particular application or use in view.

Applied research is original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific, practical aim or objective.

Experimental development is systematic work, drawing on knowledge gained from research and practical experience and producing additional knowledge, which is directed to producing new products or processes or to improving existing products or processes.

The third general type is an indication that OECD, according to its original aim of stimulating economic progress and world trade, has been focusing mainly on economic and industrial development when defining research. As shown by Godin (2006), the definitions are based on the so-called *linear model* for understanding the contribution of research to economic and industrial development. This model postulates that innovation starts with basic research, enters the phase of applied research and development, and ends with production and diffusion.

Stokes (1997) has also criticized the linear model as expressed in the OECD definitions of R&D, arguing that knowledge development and utilisation in practice is a complex interactive process and that many of the most famous scientists have been motivated by both practical contributions and theoretical understanding simultaneously. To illustrate the possible combination of aims, Stokes created 'Pasteur's Quadrant' with distinctions between 'Bohr-type' aims of research (to advance scientific understanding), 'Edison-type' aims of research (applied research driven by market needs), and 'Pasteur-type' aims of research (combination of scientific interest and considerations of use).

More generally, the linear model for understanding innovation and research practice has been challenged by models of interaction and interdependence between science and society, such as the Triple Helix model (Etzkowitz & Leydesdorff, 1995) and the 'Mode 2' model (Gibbons et al., 1994). Models for evaluating societal outcomes of research also stress the interaction, such as the 'Payback Framework' (Klautzer et al., 2011), the concept of 'Productive Interactions' (Spaapen & van Drooge, 2011; de Jong et al, 2014), and the 'Contributions Approach' (Morton, 2015).

The linear model is still influential, however, e.g. when governments justify spending on academic research and when universities explain their purposes and roles in society. On an empirical basis, Gulbrandsen & Kyvik (2010) have shown that even though most academic staff members at Norwegian universities are able to use the OECD distinctions when describing their own activities, they practice a mix of types of research covering all categories. This finding supports the results that we will present

below. We find a mix of aims of research in the published abstracts by academic staff from the four Norwegian universities that are supposed to have the distinct role of mainly performing basic research within a larger higher education sector that also includes universities of applied sciences.

Society is absent in the OECD definitions, perhaps because society is taken for granted in the perspective of economic and industrial development. In contrast, society is present in more recent paradigms for research policy, evaluation, and funding where societal impact, interaction and responsibility has become much more explicit terms. The definition of *societal impact* by the Research Excellence Framework (REF) of the United Kingdom, which needed to serve the evaluation of universities in all areas of research, is wide enough to cover all possible outcomes of all types of research. The REF defines societal impact as ‘an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia.’ (re.ukri.org/research/ref-impact). The only limitation of this definition is that it mentions effects, changes and benefits as if they are ‘accidental’ and ‘extraordinary’ not related to normal science-society interactions guided by the specific purposes of organized research contributions within fields of research (Sivertsen & Meijer, 2020).

Rather than distinguishing between main types of research, as OECD does, or main effects of research, as in the REF definition above, we find it useful to distinguish between *main aims of research* and include a societal perspective in the distinction. We are not claiming that research mainly aimed at scientific progress cannot be useful for society in the future. Our distinction is related to the aims as expressed in the publication. Our second reason for not using the basic/applied distinction is that our distinction and its definitions need to be operational in relation to the *genre* of abstracts for research publications.

Research processes can easily be initiated and performed without clear ideas about aims and possible implications, but as the process enters the phase of presentation in a publication, there is a need for explicit statements that relate the results to a body of established knowledge and explain how knowledge is advanced. Abstracts are usually required to shortly present the aims of a study and the implications of the findings along with methods and results. We will give some examples of how the requirements are formulated. These examples helped us in developing our *classification criteria* for the text analysis.

The International Committee of Medical Journal Editors (ICMJE) requires:

The abstract should provide the context or background for the study and should state the study's purpose, basic procedures (selection of study participants, settings, measurements, analytical methods), main findings (giving specific effect sizes and their statistical and clinical significance, if possible), and principal conclusions.

Here, the purpose of the research is explicitly asked for. Aims or purposes are not asked for explicitly in the requirements for abstracts by the American Psychological Association (representing publishing practices in psychology and the social sciences). However, the purpose is indirectly present in the requirements to present the “problem under investigation or research question(s)”, a “clearly stated hypothesis or hypotheses”, and, at the end, the “implications (i.e., why this study is important, applications of the results or findings)”.

PLOS One, a journal representing all areas of research, explicitly recommends including purpose and implications in the abstract. In the beginning, “Objectives or Aims – What is the study and why did you

do it?”, and at the end: “Tell the reader why your findings matter, and what this could mean for the ‘bigger picture’ of this area of research.” The journal *Nature* recommends that the structure of the abstract parallels the Introduction and Conclusion sections in the full paper:

Accordingly, you can think of an abstract as having two distinct parts — motivation and outcome — even if it is typeset as a single paragraph. For the first part, follow the same structure as the Introduction section of the paper: State the context, the need, the task, and the object of the document. For the second part, mention your findings (the what) and, especially, your conclusion (the so what — that is, the interpretation of your findings); if appropriate, end with perspectives, as in the Conclusion section of your paper.

In our study, we could see such guidelines practiced when recording the aims of the research in a large number of abstracts. If the aims were not explicitly expressed in *the first part* of the abstract, they could be inferred from the expected implications of the research *at the end*. If this was not helpful, which happened in rare cases, the abstract was only a presentation of the contents of the article. Such abstracts are called *indicative* in the literature, as opposed to *informative abstracts*, which always cover the objectives, methods, results, conclusions, and implications (Brkić et al., 2003). The few indicative abstracts in our data occurred in the humanities. We consulted the title, full text, and the aims of the journal to determine the main aims of the research.

On this basis, we define *the aims* of research operationally as those that can be inferred from the presentation in the publication, most often at the beginning and at the end of the abstract.

The *aim of contributing to scientific progress* is then defined by its presentation in the publication: Statements of the aims and implications of the research refer to the advancement of knowledge in relation to previous research and/or potential new knowledge. External use of knowledge is not mentioned.

The *aim of contributing to societal progress* is defined by its presentation in the publication: Statements of the aims and implications of the research refer explicitly to external usefulness. The aim of contributing to scientific progress may also be stated, but not necessarily.

In essence, for the distinction, there needs to be a statement about external usefulness in the abstract or in other parts of the publication. Combining the two aims in one research publication is possible, as seen in the definitions above and as confirmed by our results. We thereby confirm the idea of ‘Pasteur’s Quadrant’ (Stokes, 1997) but will not use its terminology.

Part 3b. Aims of the research: Data and methods

As mentioned above, it is demanding to classify publications by their *aims* with the methods developed for contents classification in quantitative science studies. Similar aims are not easy to identify among publications by using topical words, shared references, or standard phrases. We have discussed using trained machine learning with experts in the field and concluded negatively about this option. Boyack et al. (2014) provide an example of scientific terms in the medical field urology that they use to distinguish the four levels of research. We reviewed the example and have so far concluded that we could not extract similar information concerning the aims of research from scientific terminology in our data set.

As explained in part 2b above, most publications under study are seldom cited or exposed to abstract views. We chose to focus on publications with the clearest indications of citation impact or usage or both. To create a manageable dataset for reading, we trimmed the indicators to an even higher level of impact than those with 10 percent highest scores and arrived at a sample of 1,193 abstracts. Of these, 460 abstracts (38.6 percent) represent the most highly used publications, 369 abstracts (30.9 percent) represent the most highly cited publications, and 364 abstracts (30.5 percent) represent publications with top scores on both indicators.

For these publications, we produced a ‘blinded’ data file with linkable publication ID to the original dataset and no other information than the title, abstract, and DOI for link to full text if needed. Two researchers from our team read all abstracts independently and coded them according to the definitions above: contributing to scientific progress or societal progress or both. Mentions of external use or not was the primary criterium, and we used the ‘both’ alternative whenever both main aims were expressed. The phase of independent blinded reading resulted in a 74 percent agreement in coding. Still using blinded data, the two readers discussed the remaining 26 percent together and concluded with only one coding alternative. We will give three examples of how we coded the abstracts. Instead of presenting evident solutions, all of them will show how we reasoned in cases of doubt.

Abstract example 1

Current knowledge about relationships between leadership and workplace safety is based mainly on cross-sectional studies focusing on constructive forms of leadership. We suggest that this one-sided attention to constructive leadership and the lack of temporal research designs have restrained our understanding of: 1) the impact of both constructive and destructive forms of leadership on safety, 2) whether and how leadership is related to safety overtime, and 3) potential bidirectional associations between leadership and safety. To substantiate these claims empirically, time-lagged relationships between constructive-, laissez-faire-, and tyrannical leadership and psychological safety climate were examined among 683 employees from the offshore petroleum industry. We found that associations with psychological safety climate were dependent upon the types of leadership examined. A bidirectional relationship was established between leadership and psychological safety climate. The findings support the importance of a multidimensional approach and a temporal design in research on leadership and safety.

We coded example 1 as mainly contributing to scientific progress. This is not evident because the study makes use of large-scale empirical data from working life, and the topic of leadership in industry is of high societal interest. However, looking more closely at the first and last sentences, where the aims and implications usually are expressed in abstracts, these sentences reveal that the main purpose here is to advance methods and knowledge within the field.

Abstract example 2

This study investigated language function associated with behavior problems, focusing on pragmatics. Scores on the Children's Communication Checklist Second Edition (CCC-2) in a group of 40 adolescents (12-15 years) identified with externalizing behavior problems (BP) in childhood was compared to the CCC-2 scores in a typically developing comparison group (n = 37). Behavioral, emotional and language problems were assessed by the Strengths and Difficulties Questionnaire (SDQ) and 4 language items, when the children in the BP group were 7-9 years (T1). They were then assessed with the SDQ and the

CCC-2 when they were 12-15 years (T2). The BP group obtained poorer scores on 9/10 subscales on the CCC-2, and 70% showed language impairments in the clinical range. Language, emotional and peer problems at T1 were strongly correlated with pragmatic language impairments in adolescence. The findings indicate that assessment of language, especially pragmatics, is vital for follow-up and treatment of behavioral problems in children and adolescents.

We coded example 2 as mainly contributing to societal progress. This purpose is not evident in the first sentence, which seems to imply a general purpose of advancing knowledge. However, the study is based on diagnostic tools and the last sentence clearly indicates that the research is meant to contribute to improve treatment methods for children with behavioural problems.

Abstract example 3

Understanding the forces underpinning female genital mutilation/ cutting (FGM/C) is a necessary first step to prevent the continuation of a practice that is associated with health complications and human rights violations. To this end, a systematic review of 21 studies was conducted. Based on this review, the authors reveal six key factors that underpin FGM/C: cultural tradition, sexual morals, marriageability, religion, health benefits, and male sexual enjoyment. There were four key factors perceived to hinder FGM/C: health consequences, it is not a religious requirement, it is illegal, and the host society discourse rejects FGM/C. The results show that FGM/C appears to be a tradition in transition.

We coded example 3 as contributing to both scientific and societal progress. Independently, our two readers had selected each of the two other alternatives. There was a coding disagreement. After further discussion, we found that both types of aims were present: On the one hand, FGM/C is recognized internationally as a violation of the human rights of girls and its prevention is high on the agenda of the World Health Organization. The aims of the study are clearly related to societal progress. However, the combination of aims is expressed in the first sentence: Understanding the underpinning forces is a necessary first step. The study also aims to advance knowledge.

After having coded all 1,193 abstracts with only one of the three alternatives, we were ready to uncover how our classification configured with the two forms of impact and the gender of the 1st authors.

Part 3c. Aims of the research: Results

Before coding according to aims, the 1,193 abstracts in focus here were already distributed among the impact categories with 369 abstracts (30,9 percent) representing the most *highly cited* publications, 460 abstracts (38,6 percent) representing the most *highly used* publications, and 364 abstracts (30,5 percent) representing publications with top scores on *both* indicators.

After coding according to aims independently of the distribution above, the 1,193 abstracts were distributed among the categories of aims with 366 abstracts (30,7 percent) coded as mainly aimed at *scientific progress*, 502 abstracts (42,1 percent) coded as mainly aimed at *societal progress*, and 325 abstracts (27,2 percent) coded as having *both aims*.

According to our hypotheses, highly cited publications should be overrepresented among publications classified as aimed at scientific progress while highly used publications should be overrepresented in among publications classified as aimed at societal progress. This is confirmed in the results presented in Figure 4.

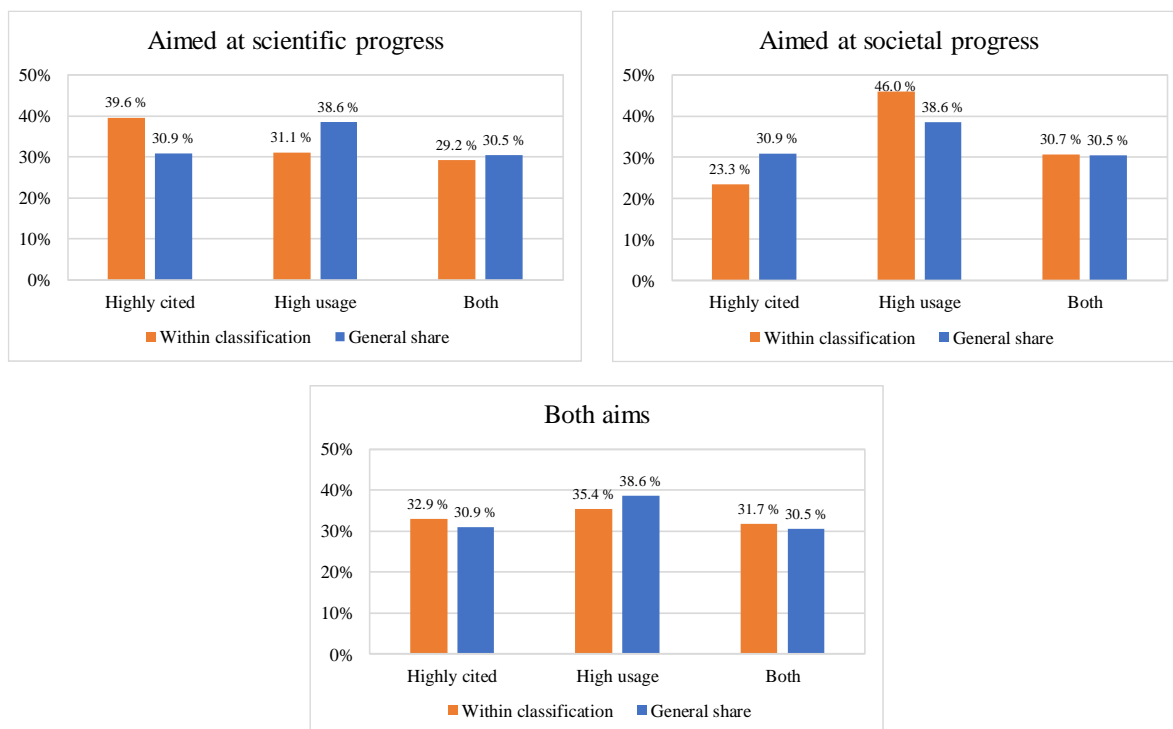


Figure 4. Within each of three classes of aims (scientific, societal, both) the share of publications by form of impact (highly cited, highly used, both) is compared to the general share in all classes of aims.

Among publications classified as aimed at scientific progress, 39.6 percent are highly cited while the general share of the highly cited publications among the 1,193 publications is 30.9 percent. Among publications classified as aimed at societal progress, 46.0 percent are highly used while the general share of the highly used publications is 38.6 percent. The same pattern of overrepresentation is not found among publications with both aims, which also confirms our hypothesis.

There are only 31 publications from the Humanities in the sample of 1,193 publications. We did not find the same pattern of overrepresentation in the Humanities, where citation indicators are less applicable. The pattern of overrepresentation is present in the other areas of research: Health Sciences, Natural Sciences and Engineering, and Social Sciences.

So far, we conclude that we have clear, but not strong indications of a relationship between the aims of research and the forms of impact. This relationship is independent of the gender perspective and deserves to be further investigated in studies comparing citation impact and altmetric indicators.

We now return to the main hypotheses of our study:

- Male researchers are overrepresented as 1st authors of publications with high citation impact because they more often engage in research mainly aimed at scientific progress.
- Female researchers are overrepresented as 1st authors of publications with high usage because they more often engage in research mainly aimed at societal progress.

Among the 1,193 abstracts under study in this part, 530 (44.4 percent) represent publications by female 1st authors while 663 (55.6 percent) represent publications by male 1st authors. The share of publications with female 1st authors is higher in this small dataset of publications with very high impact than among

all publications with high impact (43.1 percent in the dataset studied in part 2) and within the total dataset of 26,976 journal articles under study (40.3 percent). If only citation impact was considered, the shares of publications with female 1st authors would have been lower among those with the highest impact. It is due to the addition of usage as a form of impact that there is a relative overrepresentation of publications by female 1st authors among the 1,193 publications with very high impact.

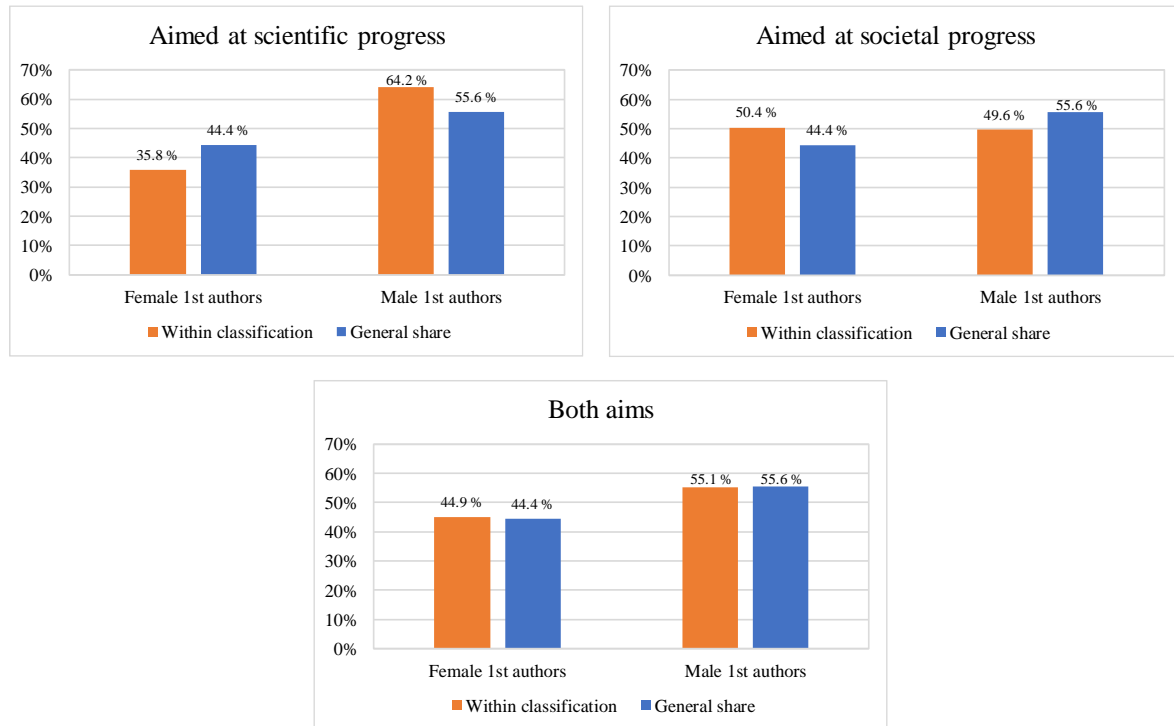


Figure 5. Within each of three classes of aims (scientific, societal, both) the share of publications by female versus male 1st authors is compared to their general share in all classes of aims.

Figure 5 shows that, compared to the general shares of 44.4 percent versus 55.6 percent, publications by male 1st authors are overrepresented among publications classified as mainly aimed at scientific progress while publications by female 1st authors are overrepresented among publications classified as mainly aimed at societal progress. Among publications classified as having both aims, the shares are almost the same as the general shares, which is an indication that the other results are reliable with regard to the hypotheses. We found the same patterns of overrepresentation in all areas of research, but with varying degrees.

Again, the indications that our hypotheses are confirmed are clear, but not strong. Female researchers more often engage in research aimed at societal progress while male researchers more often engage in research aimed at scientific progress.

We will now include the impact categories in the analysis as we return to the main aim of this part of the study, to investigate possible explanations for why publications by male 1st authors are relatively more often cited and publications by female 1st authors are relatively more often used. We have already seen female 1st authors engage relatively more in research that we have classified as aimed at societal progress. The indications were clear, but not strong, implying that there are highly cited publications among those classified as mainly aimed at societal progress, and, vice versa, that there are highly used

publications among those mainly aimed at scientific progress. We now turn the perspective and analyse the distribution of gender and aims within each impact category. Figure 6 shows that male 1st authors have their larger shares in publications aiming at scientific progress in *all impact categories* (highly cited, highly used, both high) while female 1st authors have their larger shares in articles aiming at societal progress in *all impact categories*.

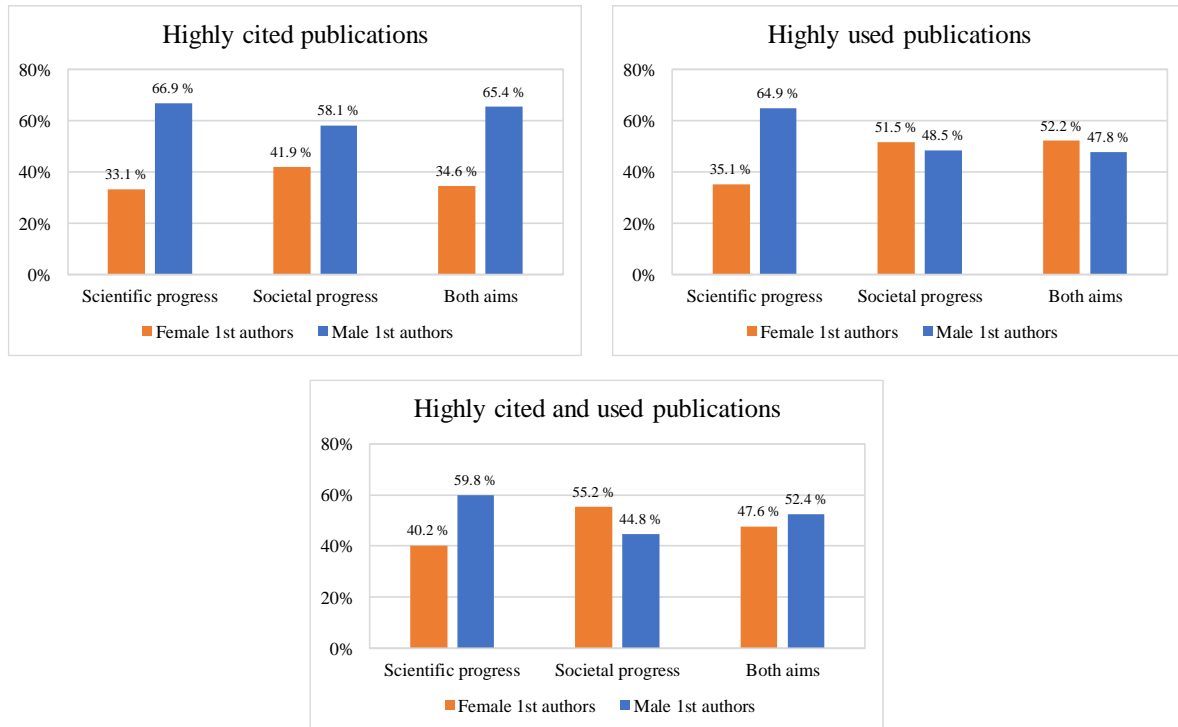


Figure 6. The share of publications by female versus male 1st authors within each class of aims (scientific, societal, both) is compared within each form of impact (highly cited, highly used, both).

The consistent gender differences across all impact categories indicate that gender and aims of research are probably *directly related*. Even among the highly cited publications, female 1st authors have their larger shares in publications aiming at societal progress. And even among the highly used publications, male 1st authors have their larger shares in publications aiming at scientific progress.

We have already shown that publications with mainly scientific aims are more often highly cited and that publications with mainly societal aims are more often highly used. We hereby suggest that differences in the aims of research can be one of the explanations for gender differences in impact.

Part 4a. Motivations for the choice of aims: Introduction

In this last part with empirical investigations, we will use responses from a large survey among researchers working in three different fields of research in five countries to study possible differences in motivations that might explain the observed gender differences in impact (part 2) and in aims (part 3).

While several studies have focused on gender differences in the career choices of discipline and field of research and focused on possible explanations for these differences (Thelwall et al., 2020), we find little

previous research into gender differences in the choice of aims and type of research within the same discipline or field. One main reason for this lack is that studies of researchers' motivations to engage with society are rare in themselves and have been dominated by innovation studies focusing on how researchers engage with industry (Perkmann et al., 2013). Gender is not an issue in these studies. An increased policy emphasis on societal relevance and impact in recent years has generated an interest in studying broader interactions with society (D'Este et al., 2018; Gunn & Mintrom, 2016; Holbrook, 2010). Some of these studies have even included the gender variable for control without analysing it (Jensen et al., 2008; van de Burgwal et al., 2019).

Closest to our research interest is a study by Ooms et al. (2019) who find that researchers have better chances of becoming full professors when “bridging between the quest for fundamental understanding and socio-economically relevant applications of their research, i.e. when they conduct predominantly Pasteur research.” Contrary to our results, this study finds that female researchers engage less in research that combines the two main aims of scientific and societal progress. The authors conclude that (along with lower productivity) the lack of ‘Pasteur-type’ research among females may explain inferior career success. The study is limited to a sample of 248 academics at two leading European universities of technology.

Recently, Chubb & Derrick (2020) found that gendered notions of societal impact in research emerged as a significant theme from two independent data sets based on interviews with researchers. Our results in this part will show that there is also a gendered difference in how researchers report their motivations for performing research and what they consider the best research to engage in.

Part 4b. Motivations for the choice of aims: Data and methods

As members of the Centre for Research Quality and Policy Impact Studies (R-QUEST, r-quest.no), where researchers from five countries join efforts to understand research assessments, standards and practices in different fields of research, we have taken part in designing and analysing a large international survey among researchers where some of the questions asked are relevant for the present study. The survey was conducted in 2017-2018 among researchers in cardiology, economics, and physics in five European countries: Denmark, the Netherlands, Norway, Sweden, and the United Kingdom. We collected 2,587 responses with an overall 28.7% response rate. Of these, 445 researchers (17.2 percent) represented cardiology, 624 (24.1 percent) represented economics, and 1,518 (58.7 percent) represented physics. Most of the respondents informed about their gender in the questionnaire. Among these, there are 501 female researchers (21.9 percent) and 1,786 male researchers (78.1 percent).

The main purpose of the survey was to study how notions of research quality vary by field of research, country, and conditions for performing research, but the survey also included questions about motivations for performing research and about engagement with society. We focus on two such questions here.

One main question in the survey was: *Which of the following motivates/inspires you to do research?* Respondents were asked to rate eight possible answers on a scale of five alternatives from 1=not important to 5=very important:

- Curiosity/scientific discovery/understanding the world
- Application/practical aims/creating a better society
- Inspiration from my colleagues (locally and/or internationally)
- Contribute to the standing of my research unit/group
- Progress in my career (e.g. tenure/permanent position, higher salary, more interesting/independent work)
- This is what I do for a living/the job I am most competent for
- Inspiration from students/young talents
- Inspiration from users (outside science)

In our analysis, we estimated the percentage of respondents by gender who rated an answering alternative as very important (5) or important (4).

Another main question in the survey was: *Think about the research you consider to be the best in your specific field/specialty. Why do you consider this the best research? You may select more than one option.* Here, the respondents only needed to decide between select or not select. The ten alternatives were:

- Has answered/solved key questions/challenges in the field
- Has changed the way research is done in the field
- Has been a centre of discussion in the research field
- Has changed the key theoretical framework of the field
- Has benefited society
- Has enabled more reliable or precise research results
- Was published in a journal with a high impact factor
- Has attracted many citations
- Has drawn much attention in the larger society
- Is what all students/prospective researchers need to read

In our analysis, we estimated the percentage of respondents by gender who selected each of the alternatives.

Part 4c. Motivations for the choice of aims: Results

The results from the first question are shown in Figure 7: *Which of the following motivates/inspires you to do research?* Female and male researchers rate the importance of most of the alternatives alike. Scientific progress is rated as most important and with no gender difference (95 percent for both). However, there are four indications of gender differences. Two of them are less relevant for our study: Female researchers rate collegial inspiration and career progress as more important motivations. The other two indications are directly relevant. Female researchers regard the aim of societal progress as a more important motivation (70.0 percent versus 61.6 percent). Inspiration from external users is also more important for female researchers (31.6 percent versus 20.6 percent). These indications are particularly clear in cardiology and economics (not shown in the figure).

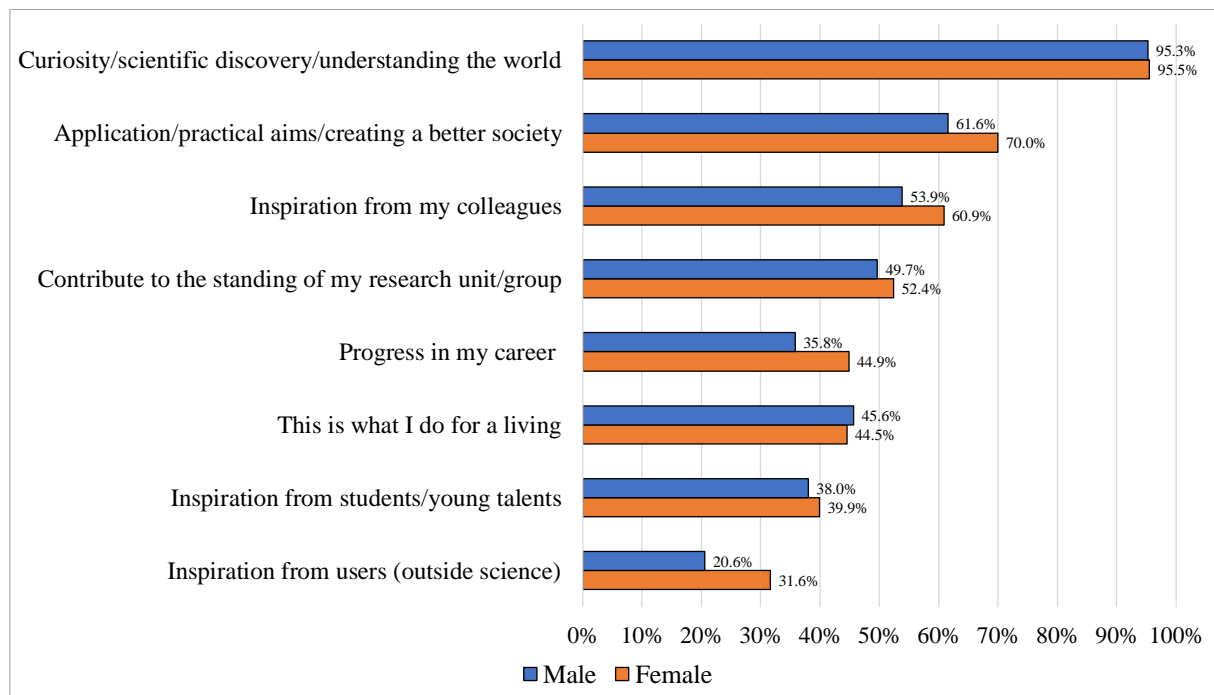


Figure 7. Motivations for performing research. Answers to the question: Which of the following motivates/inspires you to do research? Percentage within each gender who rates an alternative as very important (5) or important (4) on a scale from not important (1) to very important (5).

The results from the second question are shown in Figure 8: *Think about the research you consider to be the best in your specific field/specialty. Why do you consider this the best research? You may select more than one option.* The clearest gender difference is to what degree the researchers consider contributing to society as a characteristic for the best research. This alternative is selected by 32.9 percent of the female researchers as opposed to 25.8 percent of the male researchers. Here, the indications of gender differences are clearest in economics and physics (not shown in the figure).

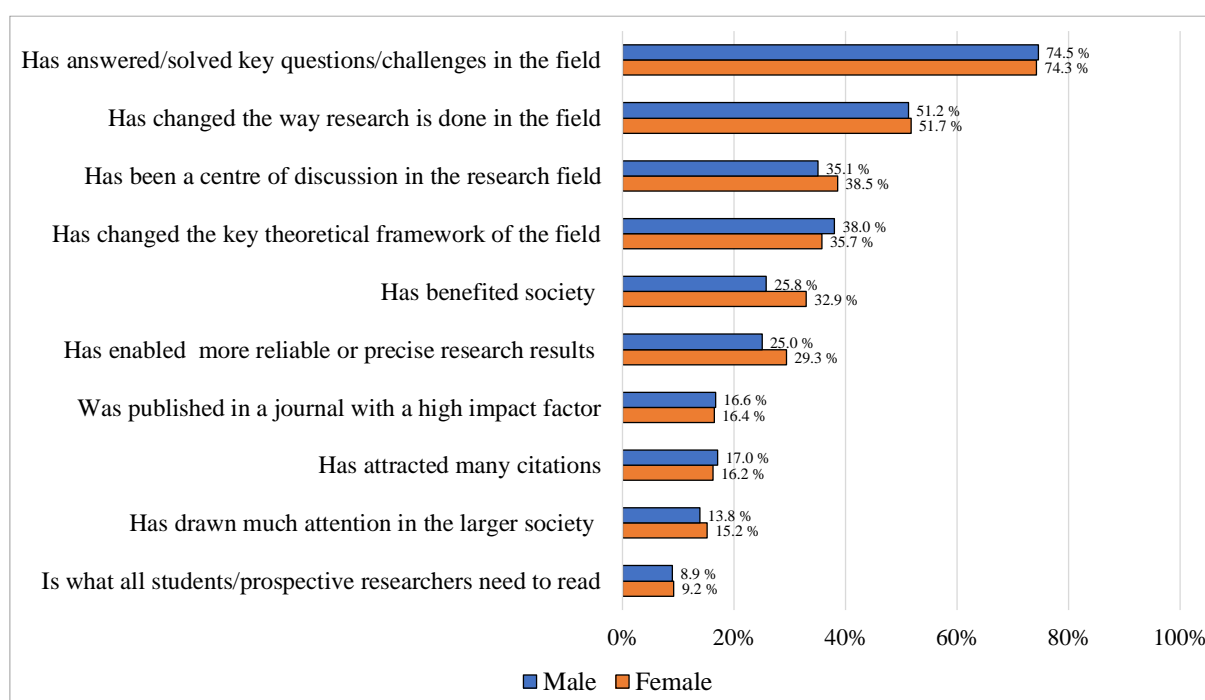


Figure 8. Perceived characteristics of best research. Percentage of respondents by gender who selected the alternative.

The results from both questions clearly indicate that scientific progress is highly valued among the respondents, and there is no gender difference in this respect. On the other hand, there are clear, but not strong indications that female researchers are more motivated than male researchers to engage in research aimed at societal progress. Female researchers also more often rate contributing to societal progress as a characteristic of the best research. Adding motivation to the factors investigated in the previous parts of this study, we can conclude as follows:

- Male researchers are overrepresented as 1st authors of publications with high citation impact because they more often value and engage in research mainly aimed at scientific progress.
- Female researchers are overrepresented as 1st authors of publications with high usage because they more often value and engage in research mainly aimed at societal progress.

As stated before, these explanations for gender differences in impact do not exclude other possible explanations. It is also important to note that female and male researchers value scientific progress just as highly. The difference is in how societal progress is valued, and in the relative frequency of engagement in research with this aim.

Part 5. Discussion and conclusions

We have used mixed methods – classical citation analysis, altmetric analysis, a survey with researchers as respondents, and text analysis of the abstracts of scientific articles – to investigate gender differences in the aims and impacts of research. We started by observing gender differences in two forms of impact of scientific publications, citation impact and abstract views, the latter interpreted as usage. We found that publications by male 1st authors are more often highly cited. This finding is in accordance with several other studies. We also found that publications by female 1st authors more often attract abstract

views. This finding is in accordance with some other studies using altmetric indicators to study broader impact of research.

We went on to seek explanations for the observed differences. Inspired by a literature that has demonstrated that basic research is more highly cited than applied research, and also by a literature that has challenged the ‘linear model’ and the basic/applied dichotomy by focusing on mutual science-society interactions, we rephrased the basic/applied dichotomy by distinguishing between three variants of aims of the research: mainly aiming at scientific progress, mainly aiming at societal progress, or both.

We operationalized the distinctions to serve the classification of almost 1,200 abstracts after reading them. The abstracts were selected as those with the highest citation impact or the highest usage or both among the publications under study. With clear but not strong indications it was confirmed that research mainly aimed at contributing to scientific progress is more often cited in the scientific literature (citation impact) while research mainly aimed at contributing to societal progress is more often read online as available abstracts (usage). This finding is independent of the gender perspective and new to the literature.

The relationship between aims and impacts of research was also found to be a possible explanation for the initially observed gender differences in forms of impact. The results from the text analysis of abstracts was supplemented by a questionnaire with response from more than 2,000 researchers in five countries in Northern Europe where we could see similar gender differences in how researchers are motivated and how they value the best research. We found a possible explanation for the gender differences in impacts which is also new to the literature:

- Male researchers more often value and engage in research mainly aimed at scientific progress, which is more highly cited.
- Female researchers more often value and engage in research mainly aimed at contributing to societal progress, which has higher usage.

Throughout, the observed gender differences are small but significant. For two reasons, we could not expect large differences. One is that most of the scientific publications under study were published in collaboration by female and male researchers. We need to distinguish them by the gender of the 1st author to find the differences. The other reason is that both forms of aims and both forms of impact can be present in one and the same publication. Around one third of the publications under study has the combination of high citation impact and high usage. We also classified around one third of the abstracts as expressing both major aims, i.e. scientific progress *and* societal progress. Furthermore, we could see from our survey results that female and male researchers attribute value to scientific progress at the same rate. The difference is only in how they value the additional aim of societal progress. On this background, which also clarifies the limitations of our study, we conclude that the observed gender differences are still significant and worthwhile following up or challenging in future studies of gender and research performance.

Our findings also have implications for research evaluation and funding and for gender policy in the research system. In a study that inspired our own, referring to their own confirmation that basic research is more cited than applied research, the authors say: “There is no epistemic justification for rating basic

research higher than applied research” (Donner & Schmoch, 2020). Our study raises a similar question about what type of research is valued the most.

Our distinctions between scientific progress and societal progress as major aims, as well as the distinction between citation impact and usage, are relevant for a more recent contest between two influential research policies and their evaluation criteria. Although seemingly in harmony (top research performance is best for society), the so-called Excellence Policy with a focus on the citation impact of international publications, and the Responsible Research and Innovation (RRI) policy with a focus on societal interaction and presence in local media, are not aligned but in conflict (D’Este et al., 2018; Bandola-Gill, 2019; Zhang & Sivertsen, 2020; Sivertsen, 2021). Referring to our results, it might be tempting to place female researchers on one side of this conflict. However, we showed that the difference is only in how female researchers attribute value to and engage in the additional aim of societal progress. Female researchers are not opposed to scientific progress, they value it and engage in it. Nevertheless, a critical discussion of how dominating evaluation criteria treat societal engagement and reflect gender differences is warranted.

Declaration of Competing Interests

The first author (Lin Zhang) is an associate-editor of *Scientometrics*, and the last author (Wolfgang Glänzel) is the editor-in-chief of *Scientometrics*.

Acknowledgements

The present study is an extended version of a paper presented at the 17th International Conference on Scientometrics and Informetrics, Rome (Italy), 2–5 September 2019 (Zhang et al. 2019). The authors acknowledge support from the National Natural Science Foundation of China (Grant No. 71974150; 72004169; 71573085), the Major Project of National Social Science Foundation of China (Grant No. 19ZDA349) and by the Research Council of Norway, Grant 256223, FORINNPOL. The authors are very grateful for valuable advices from the two anonymous reviewers.

References

- Aksnes, D.W., Rørstad, K., Piro, F., Sivertsen, G. (2011). Are female researchers less cited? A large-scale study of Norwegian scientists, *Journal of the Association for Information Science and Technology*, 62(4): 628-636.
- Aksnes, D. W., Piro, F. N. & Rørstad, K. (2019). Gender gaps in international research collaboration: a bibliometric approach. *Scientometrics*, 120, 1-28.
- Albarrán, P., Crespo, J., Ortuño, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, 88(2), 385-397.
- Bandola-Gill, J. (2019). Between relevance and excellence? Research impact agenda and the production of policy knowledge. *Science and Public Policy*, 46(6), 895–905.
- Barrios, M., Villarroya, A., Borrego, Á. (2013). Scientific production in psychology: a gender analysis. *Scientometrics*, 95(1): 15-23.
- Beaudry, C., Larivière, V. (2016). Which gender gap? Factors affecting researchers' scientific impact in science and medicine. *Research Policy*, 45(9), 1790-1817.
- Bendels, M. H., Müller, R., Brueggmann, D., & Groneberg, D. A. (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PLoS ONE*, 13(1), e0189136.
- Bornmann, L. (2013.) What is societal impact of research and how can it be assessed? A literature survey. *Journal of the American Association for Information Science and Technology*, 64(2): 217-233.
- Bornmann L, Leydesdorff L, Mutz R. (2013). The use of percentiles and percentile rank classes in the analysis

- of bibliometric data: Opportunities and limits. *Journal of Informetrics*, 7(1):158-165.
- Bornmann, L. (2014.) Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895-903.
- Bornmann, L., Haunschild, R., Adams, J. (2019). Do altmetrics assess societal impact in a comparable way to case studies? An empirical test of the convergent validity of altmetrics based on data from the UK research excellence framework (REF). *Journal of Informetrics*, 13(1), 325-340.
- Boyack, K. W., Patek, M., Ungar, L. H., Yoon, P., & Klavans, R. (2014). Classification of individual articles from all of science by research level. *Journal of Informetrics*, 8(1), 1–12.
- Bosquet, C., Combes, P.P. (2013). Are academics who publish more also more cited? Individual determinants of publication and citation records. *Scientometrics*, 97(3), 831–857
- Brkić, S., Vuconevic, M., Docik, Z. (2003). Title, abstract, key words and references in biomedical articles. *Archive of Oncology*, 11(3), 207-9
- Caplar, N., Tacchella, S., Birrer, S. (2017). Quantitative Evaluation of Gender Bias in Astronomical Publications from Citation Counts. *Nature Astronomy*, 1(6), 0141.
- Chi, P.-S., Gorraiz, J., & Glänzel, W. (2019), Comparing capture, usage and citation indicators: an altmetric analysis of journal papers in chemistry disciplines. *Scientometrics*, 120(3), 1461–1473.
- Chubb, J., & Derrick, G. E. (2020). The impact a-gender: gendered orientations towards research impact and its evaluation. *Palgrave Communications*, 6(1), 1-11.
- D’Este, P., Ramos-Vielba, I., Woolley, R., and Amara, N. (2018). "How do researchers generate scientific and societal impacts? Toward an analytical and operational framework." *Science and Public Policy*, 45(6), 752–763.
- de Jong, S., Barker, K., Cox, D., Sveinsdottir, T., & Van den Besselaar, P. (2014). Understanding societal impact through productive interactions: ICT research as a case. *Research Evaluation*, 23(2), 89-102.
- Donner, P., Schmoch, U. 2020. The implicit preference of bibliometrics for basic research. *Scientometrics*, 124(2), 1411-1419.
- Elsevier (2020). The Researcher Journey Through a Gender Lens. Retrieved from <https://www.elsevier.com/research-intelligence/resource-library/gender-report-2020>
- Etzkowitz, H., Leydesdorff, L. (1995). The Triple Helix -- University-Industry-Government Relations: A Laboratory for Knowledge Based Economic Development., 14(1), 14-19.
- Garfield, E. (1979). *Citation indexing—Its theory and application in science, technology and humanities*. New York: Wiley.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*, London: Sage.
- Ghiasi, G., Larivière, V., & Sugimoto, C. R. (2015). On the Compliance of Female Engineers with a Gendered Scientific System. *PLoS ONE*, 10(12), e0145931.
- Glänzel, W., & Chi, P.S. (2020), The big challenge of Scientometrics 2.0: Exploring the broader impact of scientific research in public health. *Scientometrics*, 125(2), 1011-1031.
- Godin, B. (2006). The Linear Model of Innovation: The Historical Construction of an Analytical Framework. *Science, Technology, & Human Values*, 31(6), 639-667.
- Gulbrandsen, G., Kyvik, S. (2010). Are the concepts basic research, applied research and experimental development still useful? An empirical investigation among Norwegian academics. *Science and Public Policy*, 37(5), 343–353.
- Gunn, A., & Mintrom, M. (2016). Higher Education Policy Change in Europe: Academic Research Funding and the Impact Agenda. *European Education*, 48(4), 241–257.
- Haustein, S. (2014). Readership metrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond bibliometrics: harnessing multi-dimensional indicators of performance* (pp. 327-344). Cambridge, MA, USA: MIT Press.
- Holbrook, J. B. (2010). The use of societal impacts considerations in grant proposal peer review: a comparison of five models. *Technology & Innovation*, 12(3), 213–224.
- Huang, J., Gates, A. J., Sinatra, R., & Barabási, AL. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *The Proceedings of the National Academy of Sciences*, 117(9), 4609-4616.
- Jensen, P., Rouquier, J. B., Kreimer, P., & Croissant, Y. (2008). Scientists who engage with society perform better academically. *Science and public policy*, 35(7), 527–541.
- Klautzer, L., Hanney, S., Nason, E., Rubin, J., Grant, J., & Wooding, S. (2011). Assessing policy and practice impacts of social science research: the application of the Payback Framework to assess the Future of

- Work programme. *Research Evaluation*, 20(3), 201–209.
- Kyvik, S., & Teigen, M. (1996). Child Care, Research Collaboration, and Gender Differences in Scientific Productivity. *Science, Technology, & Human Values*, Vol 21(1), 54–71.
- Larivière, V., & Costas, R. (2016). How many is too many? On the relationship between research productivity and impact. *PLoS ONE* 11(9), e0162709.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., Sugimoto, C.R. (2013). Bibliometrics: Global Gender Disparities in Science. *Nature*, 504(7479), 211–213.
- Maliniak, D., Powers, R., & Walter, B. F. (2013). The Gender Citation Gap in International Relations. *International Organization*, 67(04), 889–922.
- Morton, S. (2015). Progressing research impact assessment: A “contributions” approach. *Research Evaluation*, 24(4), 405–419.
- Narin, F., Pinski, G., & Gee, H. H. (1976). Structure of the biomedical literature. *Journal of the American Society for Information Science*, 27(1), 25–45.
- Nielsen, M. W. (2016). Gender inequality and research performance: moving beyond individual-meritocratic explanations of academic advancement. *Studies in Higher Education*, 41(11), 2044–2060.
- OECD. (2015). *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development. The Measurement of Scientific, Technological and Innovation Activities*. Paris: OECD Publishing.
- Ooms, W., Werker, C., & Hopp, C. (2019). Moving up the ladder: heterogeneity influencing academic careers through research orientation, gender, and mentors. *Studies in Higher Education*, 44(7), 1268–1289.
- Paswan, J., & Singh, V. K. (2020). Gender and research publishing analyzed through the lenses of discipline, institution types, impact and international collaboration: a case study from India. *Scientometrics*, 123, 497–515.
- Paul-Hus, A., Sugimoto, C. R., Haustein, S., & Larivière, V. (2015). *Is there a gender gap in social media metrics?* Paper presented at the 15th International Conference on Scientometrics & Informetrics, Istanbul, Turkey.
- Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Brostrom, A., D’Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Krabel, S., Kitson, M., Llerena, P., Lissoni, F., Salter, A., & Sobrero, M. (2013). Academic engagement and commercialisation: A review of the literature on university-industry relations. *Research Policy*, 42(2), 423–442.
- Pudovkin, A., Kretschmer, H., Stegmann, J., & Garfield, E. (2012). Research evaluation. Part I: productivity and citedness of a German medical research institution. *Scientometrics*, 93(1), 3–16.
- Rigg, L. S., McCarragher, S., & KrMaleec, A. (2012). Authorship, Collaboration, and Gender: Fifteen Years of Publication Productivity in Selected Geography Journals. *Professional Geographer*, 64(4), 491–502.
- Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628–638.
- Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: An empirical analysis of the potential. *Scientometrics*, 91(2), 567–575.
- Sivertsen, G., & Meijer, I. (2020). Normal versus extraordinary societal impact: how to understand, evaluate, and improve research activities in their relations to society? *Research Evaluation*, 29(1), 66–70.
- Sivertsen, G. (2021). Publishing in the social sciences and its representation in research evaluation and funding systems. In: Engels, T. C. E. & Kulczycki, E. (Eds.). *Handbook on Research Assessment in the Social Sciences*. Edward Elgar, forthcoming (ISBN 9781800372542).
- Spaapen, J., & van Drooge, L. (2011). Introducing “productive interactions” in social impact assessment. *Research Evaluation*, 20(3), 211–218.
- Stokes, D.E. 1997. *Pasteur’s Quadrant: Basic Science and Technological Innovation*. Washington, DC: The Brookings Institution.
- Thelwall, M., & Kousha, K. (2014). Academia.edu: Social network or Academic Network? *Journal of the American Society for Information Science and Technology*, 65(4), 721–73.
- Thelwall, M. (2018a). Do females create higher impact research? Scopus citations and Mendeley readers for articles from five countries. *Journal of Informetrics*, 12(4), 1031–1041.
- Thelwall, M. (2018b). Does Female-authored Research have More Educational Impact than Male-authored Research? Evidence from Mendeley. *Journal of Altmetrics*, 1(1), 3.
- Thelwall, M., Bailey, C., Tobin, C., & Bradshaw, N. A. (2019). Gender differences in research areas, methods and topics: Can people and thing orientations explain the results? *Journal of Informetrics*, 13(1), 149–169.

- Thelwall, M. (2020a). Gender differences in citation impact for 27 fields and 6 English speaking countries 1996-2014. *Quantitative Science Studies*, 1(2), 599–617.
- Thelwall, M. & Sud, P. (2020). Greater female first author citation advantages do not associate with reduced or reducing gender disparities in academia. *Quantitative Science Studies*, 1(3), 1283–1297.
- Thelwall, M. (2020b). Female citation impact superiority 1996-2018 in six out of seven English-speaking nations. *Journal of the Association for Information Science and Technology*, 71(8), 979–990.
- Thelwall, M., Abdoli, M., Lebiedziewicz, A., Bailey, Carol. (2020). Gender disparities in UK research publishing: Differences between fields, methods and topics. *Profesional de la información*, 29(4), e290415.
- van de Burgwal, L.H.M, Hendrikse, R., Claassen, E. (2019). Aiming for impact: Differential effect of motivational drivers on effort and performance in knowledge valorisation. *Science and Public Policy*, 46(5), 747–762.
- van den Besselaar, P., & Sandström, U. (2016). Gender differences in research performance and its impact on careers: a longitudinal case study. *Scientometrics*, 106(1), 143-162.
- van Eck, N. J., Waltman, L., van Raan, A. F., Klautz, R. J., & Peul, W. C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLoS ONE*, 8(4), e62395.
- Zhang, L., Sivertsen, G. (2017). Productivity versus citation impact: A study of persons, not just authors. In *Proceedings of the 16th International Conference on Scientometrics & Informetrics, Wuhan, China, 2017*(pp. 970-975).
- Zhang, L., & Sivertsen, G., (2020). The New Research Assessment Reform in China and Its Implementation. *Scholarly Assessment Reports*, 2(1), p.3.
- Zhang, L., Du, H., Huang, Y., Glänzel, W., Sivertsen, G. (2019). Gender, age, and broader impact: A study of persons, not just authors. In *Proceedings of the 17th International Conference on Scientometrics & Informetrics, Vol. II* (pp. 1888-1893).
- Zuckerman, H., Cole, J.R. (1975). Women in American science. *Minerva*, 13, 82–102.