



# Evaluating gender equality effects in research and innovation systems

Susanne Bühler<sup>1</sup> · Evanthia Kalpazidou Schmidt<sup>2</sup> · Rachel Palmén<sup>3</sup> · Sybille Reidl<sup>4</sup>

Received: 16 December 2019 / Published online: 13 July 2020  
© The Author(s) 2020

## Abstract

Despite the fact that the topic of “women in research and innovation” has been on the agenda for decades and numerous measures have been implemented at both national and supranational levels to improve gender equality in research and innovation systems, it is still unclear which measures and under which conditions these measures are most effective. Even less research has been carried out on the effects of better representation of women in terms of research and innovation. This paper is based on the application of an innovative evaluation framework, which encompasses complexity and theory of change approaches and aims at exploring the link between interventions and their subsequent effects to two case studies. We discuss two major German flagship programmes aiming at increasing the participation of female researchers in the science system, the “Women Professorship Programme” and the “Pact for Research and Innovation”. Through the two programmes, we tested and validated the evaluation framework and its indicators. As part of the validation process, a theory of change has been developed for each of the programmes. The theory-based evaluation approach helped not only to identify gender equality impacts but also broader effects on research and innovation that might have otherwise remained undetected. We studied the effects of the two programmes on the number of women in leadership positions and analysed whether an increase in the proportion of women leaders influences publication patterns. Although linear linkages are challenging to establish due to the complexity of the process, the findings suggest that the flagship programmes have contributed not only to higher shares of women researchers but also to improved female publication and citation rates. There are clear benefits for Germany in terms of scientific results from the increased proportion of female researchers in research and innovation.

**Keywords** Gender equality · Flagship programmes · Evaluation · Bibliometrics · Theory of change · Research and innovation

---

✉ Susanne Bühler  
susanne.buehrer@isi.fraunhofer.de

Extended author information available on the last page of the article

## Introduction

With the rise of evidence-based policy making (e.g. Nutley et al. 2002; Solesbury 2001; Sanderson 2002), expectations have grown regarding the use of scientific evidence in decision-making in the policy realm. Scholars show that assessments of evidence-based policy-making have been increasing alongside the growing interest for impact assessment (Reale et al. 2014). At the same time, establishing causal relationships between policy interventions and observed effects poses a theoretical challenge, and entails empirical and methodological problems, as linear relations between interventions and impacts are difficult to identify (Halpern 2014; Reale et al. 2017). As regards the effects of gender equality policies, scholars call attention to the lack of studies and evidence, and the simplification of approaches in impact assessment of policies (Timmers et al. 2010; Kalpazidou Schmidt and Cacace 2017).

Despite many efforts undertaken in the past, there is no comprehensive and rigorous analytical framework to consider the numerous variables in gender equality interventions. Various projects funded by the European Commission such as PRAGES (Practising Gender Equality in Science), GARCIA (Gendering the Academy and Research: combating Career Instability and Asymmetries), GEDII (Gender Diversity Impact—improving research and innovation through gender diversity), GENERA (Gender Equality Network in the European Research Area), GenPORT (An internet portal for sharing knowledge and inspiring collaborative action on gender and science), GenSET (Increasing Capacity for Implementing Gender Action Plans in Science), STAGES (Structural Transformation to Achieve Gender Equality in Science) and GENOVATE (Transforming Organisational Culture for Gender Equality in Research and Innovation) have explored the gender equality (GE) dimension with different foci, encompassing varying levels of applications, from basic research to developing various supporting mechanisms. While these previous projects and subsequent studies have illustrated a number of evaluation approaches, concepts, etc. and provided examples of measuring different kinds of impacts, a clear understanding of the linkages between gender equality-related policy initiatives and interventions (inputs) and results (outcomes and impacts) is still not available. There is hence limited knowledge about how effective GE interventions have been and little is known about the effects of these interventions in research and innovation. In order to address these challenges, an evaluation framework has been developed within an H2020 project,<sup>1</sup> aiming at mapping the mechanisms that mediate the relationship between gender equality inputs and the expected effects not only in terms of gender equality itself, but also on research and innovation (R&I). The evaluation framework<sup>2</sup> provides the theory and tools to analyse how gender equality-related interventions contribute to the achievement of set objectives on gender equality<sup>3</sup> and how those achievements affect the desired outcomes of research and innovation.

---

<sup>1</sup> <https://www.efforti.eu/>.

<sup>2</sup> For a more detailed presentation of the evaluation framework, see Kalpazidou Schmidt and Graversen (2020).

<sup>3</sup> In the European Research Area three objectives are pursued: More women in research, more women in decision making positions and the integration of the gender dimension in research and education.

## Theoretical framework

Evaluating policy interventions is a complex process as no linear link between interventions and effects can be easily established (Cartwright and Hardie 2012; Dahler-Larsen 2012). Our theoretical framework has its point of departure in complexity theory and adopts a theory of change perspective in studying the link between GE interventions and effects in research and innovation (Kalpazidou Schmidt and Graversen 2020).

According to the notion of complexity, gender equality interventions are embedded in the complex systems that they form part of and involve multiple variables that interact in non-linear ways to produce effects. These systems respond to the policy interventions, adapt and produce new conditions (Halpern 2014). Evaluating complex interventions poses thus great challenges as new conditions are constantly emerging (Rogers 2008). Reale et al. (2017) assert that it is problematic to speak about attribution, as it is difficult to determine to what extent the interventions are the exclusive or most important cause of the measured effects. Therefore, in complex systems, impact cannot be directly attributed to a particular intervention but has to be conceptualised by means of evaluative approaches that pursue intervention contributions to achieve impact (Kalpazidou Schmidt and Cacace 2017, 2019).

One way to address the complex challenges discussed above is the theory-based impact evaluation approach (Kalpazidou Schmidt and Graversen 2020). Theories of change may be used as models of how change is expected to occur or how change has come about (Mayne and Johnson 2015). It is similar to the logic model, often used in development interventions, which convey a scheme, program, or project in a brief, visual format (McLaughlin and Jordan 2004; Knowlton and Phillips 2012) but explicitly includes a reflection on assumptions. Mayne and Johnson (2015) assert that a theory of change develops a framework to highlight how an intervention will perform. Thus, a tested and verified theory of change can be the point of departure for assessing the contribution of the intervention or programme to the measured effects. We build on this approach in line with a growing strand of research that discusses how the theory of change can contribute to the evaluation and understanding of policy interventions (Funnell and Rogers 2011; Rogers 2008).

According to Vogel (2012, 3) a theory of change approach is “an outcomes-based approach which applies critical thinking to the design, implementation and evaluation of initiatives and programmes intended to support change in their contexts”. In theory-based impact evaluation, causality is defined as a problem of contribution, not attribution. “Why and how” questions are typically asked instead of “how things would have been without” as counter-factual approaches do (Döring and Bortz 2016). The goal is to answer the “why it works” question by identifying the theory of change (“how things should logically work to produce the desired change”) behind a policy intervention or program and assessing its success by comparing theory with actual implementation in a certain context. Articulating assumptions about the links between interventions and their effects to make explicit the mechanisms producing change are key to the development of a theory of change (Van Belle et al. 2010).

Theory of change and gender equality approaches can enrich each other. The added value of combining gender sensitive perspectives and theory of change is that both address change and seek to articulate how change occurs, adhering to the non-linearity and contribution paradigm. A theory of change is characterized by reflexivity, where key assumptions linked to a programme are made explicit with the aim to be verified or challenged through empirical testing. Gender transformative work encompasses

questioning assumptions behind gender roles as well as relations, it can also help to unpick gendered assumptions that may form the basis of the design of various interventions (HIVOS 2014). There are often many assumptions about gender relations that essentialise both women and men. In the field of R&I—these include gender differences in terms of different publication behaviours, networking patterns, research topics, leadership styles etc. Making these assumptions explicit means that interventions need to be justified on the basis of evidence. Reflexivity is also central to gender sensitive approaches, as is context. Gender perspectives and reflexive approaches are sought after in the literature that could help understand how to evaluate policies from a gender perspective (Bustelo 2017).

According to the adopted evaluation framework, gender equality interventions are embedded in a complex context in which a large number of variables interact with each other and thus ultimately determine the effectiveness and impact of a programme (Kalpazidou Schmidt and Graversen 2020). The actual results of GE policies depend both on policy effectiveness and on other contextual variables. Contextual factors are organizational structures and cultures, as well as national and regional structures, capabilities and policies. The application of gender sensitive and theory-based impact evaluation approaches allows us to take these different levels of influences on policy effectiveness—mechanisms and context—systematically into account.

This theoretical perspective, based on theory of change and complexity, has important implications for gender equality policy and practice. It implies that the effects of interventions are largely expected in terms of contributions to change, improved conditions to foster change, as well as an increased probability that change can happen. Therefore, to better understand the effects of the programmes discussed below, we need to move away from deterministic models expecting the programmes to lead to change over a relatively short period of time (and in a linear logic). Instead, we need to adopt a reflexive and probabilistic approach demonstrating contributions to change under different contextual conditions and over longer periods of time based on clearly formulated assumptions about the impact of interventions in specific contexts, developing theories of change for each of the cases we study (Kalpazidou Schmidt and Graversen 2020).

## Methodology

We carried out a comprehensive desk research drawing on already developed and applied indicators in gender equality interventions and R&I research (RIO Observatory, OECD STI Scoreboard etc.), as well on recent studies on RRI indicators (Ravn et al. 2015a, b; European Commission 2015a) to develop the evaluation framework and create the preliminary list of indicators. First, we identified the most relevant indicators according to a systematic literature review. Second, based on the review and “smart practice” examples implemented in different organisations and contexts, we clustered the indicators into different categories, dimensions and sub-dimensions, according to an evaluation logic model based on the theory of change. The indicators have also been differentiated in terms of input, output, outcome and impact. Finally, for each of these aspects, the indicators have been categorised at micro/individual or team level, meso/organisational level and macro/policy or country level (Kalpazidou Schmidt et al. 2018).

The evaluation framework and the indicators are based on an extensive literature review and the collection and review of “smart practices” implemented in Europe and beyond. The identification of “smart practices” was based on an assessment of practices that are

**Table 1** Examples for the relationship between gender equality and benefit areas

|   |   |
|---|---|
| <p><b>GE and scientific benefits</b><br/>                     Interdisciplinarity and thematic diversity<br/>                     Higher share of publications<br/>                     Higher share of citations<br/>                     Scientific excellence<br/>                     Better dissemination of research results<br/>                     Social responsiveness</p> | <p><b>GE and economic benefits</b><br/>                     Increased creativity and organizational innovation<br/>                     Better strategic decision-making<br/>                     Increased overall competitiveness<br/>                     Better financial performance<br/>                     Positive employment effects and job satisfaction<br/>                     More effective recruiting and retention<br/>                     Increased organizational attractiveness, brand image and reputation<br/>                     Better networking and access to customers and markets<br/>                     Stronger adherence to ethics and rules of conduct</p> |
| <p><b>GE and environmental benefits</b><br/>                     More sustainability initiatives<br/>                     Higher environmental consciousness in consumption<br/>                     More eco-innovations</p>   | <p><b>GE and societal benefits</b><br/>                     Combating gender discrimination through symbolic commitment to equality<br/>                     Empowerment and confidence<br/>                     More corporate social responsibility<br/>                     More supportive and philanthropic behaviour</p>  |

Developed from Bühner and Yorulmaz (2019)

relevant, effective and efficient in the context that they operate in terms of their quality of both evaluation and measurement (Kalpazidou Schmidt et al. 2018). “Smart practice” examples evaluated various different measures in terms of scope and length: some constituted large national programmes with a long-term perspective, while others were of a more limited character. The selection of “smart practices” was based on the following criteria (1) the quality of the implemented measures, and (2) the impact of the measures. The quality of the measures was assessed based on the parameters of relevance, effectiveness, efficiency, and sustainability of the interventions, while the impact of the measures was assessed in relation to its subjective/objective dimension. Subjective impact refers to the satisfaction of beneficiaries and the ability of the intervention to promote consensus among the key stakeholders, which is a precondition for impact. Objective impact involves the effects of the intervention on the organization but—in line with the contribution, not attribution approach of the complexity theory—it also includes a “probabilistic stance”, i.e. “the creation of conditions pointing to the activation of further change processes” (Kalpazidou Schmidt and Cacace 2017, 106).

In addition, we used the existing evidence on the beneficial impacts of gender equality on different areas, integrating the respective indicators into the developed evaluation framework, in order to illustrate the wider areas where effects can be recorded (Table 1).

The links between gender diversity and the selected types of benefits mentioned above can be traced back to greater diversity of thinking, differences in values and norms, the activation of underused human capital as well as different collaboration styles. Power asymmetries are important as well. In the field of science, this often leads to the fact that women tend to publish in new and emerging fields of science, not yet dominated by male colleagues. Those new fields are often interdisciplinary in nature and typically associated with higher shares of citations (Okamura 2019, Wang et al. 2015). From a business perspective, the presence of diversity balances biases, which thereby contributes to the generation of alternative perspectives and experiences for exploring new problems. Thus, diverse teams are much more likely to consider and implement alternative approaches and have innovative ideas. Through their diverse viewpoints and ideas, diverse teams

develop ideas and solutions that are more creative and more inclusive, which often leads faster to results and innovative products (Cosley et al. 2015). Furthermore, diversity in the workforce makes it possible to better adapt to different customer groups or markets. By giving the customer groups interlocutors who “speak the same language”, the company is also able to address and win new customers. The presence of female directors signals to stakeholders that the company is committed to the advancement of women, which is interpreted as a socially responsible action (Bilimoria and Wheeler 2000). Finally, female employees can bring in their particular competencies to cope with necessary changes induced by eco-innovations (shift in firms’ organizational goals, practices and routines due to their complexity and systemic character) through emphasizing teamwork and cohesion (Horbach and Jacob 2017). The link between environmental impacts and gender diversity is derived from the observation that pro-environmental behavior is associated with femininity (Brough et al. 2016). According to recent research, women have more environmentally focused values (Civitas 2020). Risk perception is a further key mediating concept: Difference in environmental concern is due to differentially perceived vulnerability to risk in terms of health and safety as well as social and economic threats (Xiao and McCright 2012). Women are more likely to have a higher recognition of health issues and more highly developed risk perceptions, often acting on their internalised health and environment orientation (Schultz and Stieess 2009). However, it is worth mentioning that collaboration within diverse teams is challenging as well and the performance of such diverse teams depends on various factors like gendered competency expectations, hierarchy in teams (Rommes 2014), level of team deference, scale of empathy, acknowledgement of gender issues in the team and perception of rebalancing power, to mention just a few.

In accordance with the theory of change approach, we revisited, validated and refined the initial configurations regarding the two cases. Thus, theories of change were developed for these interventions in order to provide examples of the mechanisms behind the specific programmes and a framework for understanding the multi-faceted character of the interventions (Kalpazidou Schmidt and Graversen 2020).

As already mentioned, the evaluation framework developed based on complexity and theory of change approach was followed up by the case studies that have been conducted to validate and further improve the evaluation framework. Case studies as a method have been used extensively in evaluation research. Yin (1994, 13) defines a case study inquiry as one that “Investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident.” Therefore, the case study method lends itself to research where contextual factors are highly pertinent to the phenomenon of study (ibid). The multiple case study work shed light on those factors and mechanisms that shape and influence the effects of gender equality interventions on research and innovation outputs. It also set out to explain how the national science system influences the intervention in terms of the main contextual elements as well as the main agendas, strategies, and policies that shape the intervention.

Using the evaluation framework, the case studies examined whether two of the major German flagship programmes to increase the participation of female researchers in the German science system, the “Women Professorship Programme” and the “Pact for Research and Innovation”, have actually increased the number of women, especially in leadership positions. In a second step, based on literature and desk research as well as bibliometric analysis using Scopus, we analysed whether such an increase influences the publication patterns of authors with German affiliation.

In the following section, we focus on two cases of the German flagship programmes to increase the participation of female researchers in the science system, the “Women

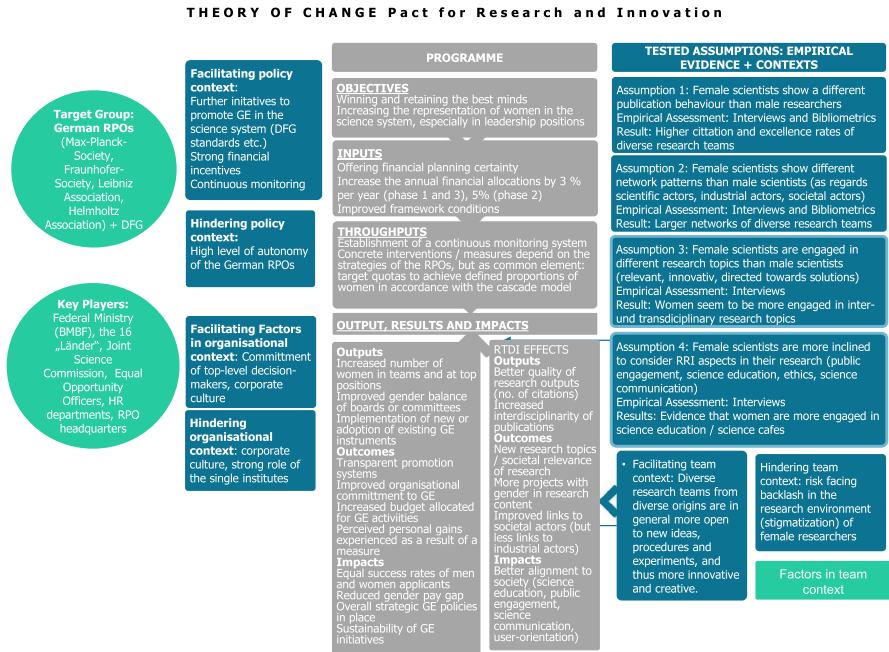
Professorship Programme” and the “Pact for Research and Innovation”, and discuss their effects and possible benefits in terms of more women researchers and publications. For each of the selected programmes, we present the Theory of Change (ToC) and describe not only the programme objectives, inputs and throughputs, but also the target group, the central actors, as well as promoting and hindering contextual factors at policy, organisational and team level. Each ToC also presents the expected or already occurred effects in the form of short-term outputs, medium-term outcomes and longer-term impacts. Based on the impact assumptions also explicitly made in the ToC, which go beyond the actual GE effects, the ToC also contains references to R&I impacts.

## The case studies

The two cases were selected because they constitute the two most important examples in the German science system to overcome its rather poor performance in terms of gender equality. In Germany the central policy context factors include freedom of science, research and teaching as a fundamental right pursuant to Article 5 of the Basic Law. Accordingly, universities as well as RPOs have comparatively high autonomy. Going along with this structure, actors in the science system have committed to gender equality goals, and there are positive incentives, but there are only a few legally binding measures. This can be assumed one of the reasons why gender equality in academia is improving, but this process is slow. Women in academia have lower positions on average compared to men; they are also more often working in precarious positions than men. Compared to the European average, approximately twice as many female researchers in Germany have precarious working contracts (European Commission 2015b, 104). Gender inequality in R&I in Germany is connected to the working time and attendance culture. The working time culture in academia means that even researchers in full-time positions regularly work overtime (Eurostat 2016). Male researchers work more overtime than female researchers (Eurostat 2016), which may lead to further career advantages for men. The high workload and attendance culture that characterises scientific careers is less attractive for women than for men (Niessen et al. 2010) and thus may constitute a potential barrier to gender equality in science.

The first of our cases, the “Pact for Research and Innovation”, is a flagship project in Germany, for several reasons: its size (3% growth rate of the institutional funding for each of the RPOs, during phase 2 even 5%), long duration and the functional mechanism (linking overall policy objective including GE with institutional funding). The intervention is rather unique as it connects institutional funding with overall policy objectives like improved gender equality (GE). The second case, the “Women’s Professorship Programme”, was selected for the following reasons: its size (150 million euros for each of the first two phases, 200 million euros for the third phase), its duration (since 2008) and the smart combination of goals: not only promoting GE and GE in leadership positions, but also structural change at German Higher Education Institutions). Finally, this case shows a rather innovative functional mechanism through relying on competition, and setting incentives to develop convincing gender equality plans.

The “Pact for Research and Innovation” started in 2006. Meanwhile, the start of a fourth phase has been decided, extending the program until 2030. The current “Pact for Research and Innovation” (2016–2020) has two predecessors. The first phase lasted from 2005 to 2010, the second phase from 2011 to 2015 and the third phase runs from

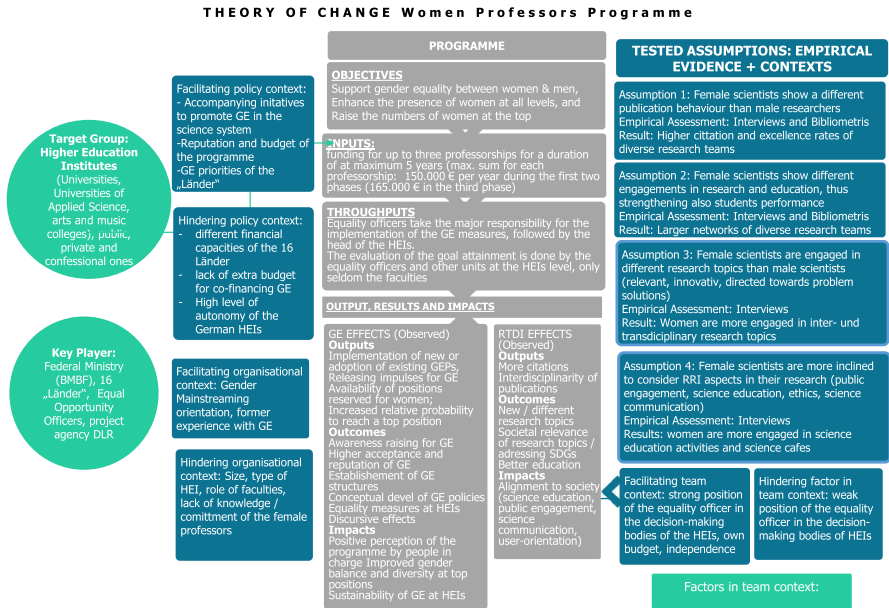


**Fig. 1** Theory of change for the “Pact for Research and Innovation”

2016 to 2020. The overriding goal of the concerted action by the federal government and the states is to strengthen the competitiveness of the German research system. It addresses the German Research Foundation (DFG) as the most important source of third-party funds in Germany and the publicly funded non-university research institutions Fraunhofer-Gesellschaft (FhG), Max Plank Society (MPG), Helmholtz Association (HGF) and Leibniz Association (WGL). The Pact obliges the research organisations to comply with several negotiated targets. The organisations themselves are responsible for the progress towards these targets and must document this in an annual monitoring report. In return for their compliance, the organisations’ budgets receive an annual boost of currently 3%. Furthermore, the government guarantees them sufficient autonomy and flexibility in budgeting, human resources and construction, public procurement and participation rights. In the context of the “Pact for Research and Innovation”, the research organisations also set targets for the share of women at different hierarchical levels, applying the logic of the cascade principle (BMBF 2016; GWK 2016). Figure 1 shows the ToC for the “Pact for Research and Innovation”, as derived from the theory-of-change approach. Beside the effect on improved shares of women in research teams and decision-making positions, we expect different types of research outputs such as a better quality of research, operationalized as number of citations.

The “Women’s Professorship Programme” is a national initiative that addresses higher education institutions (HEIs) in Germany. Not only universities but also universities for applied sciences (“Fachhochschulen”) and art and/or music colleges are eligible for funding. The eligibility criteria are different for HEIs applying for funding for the first time, and

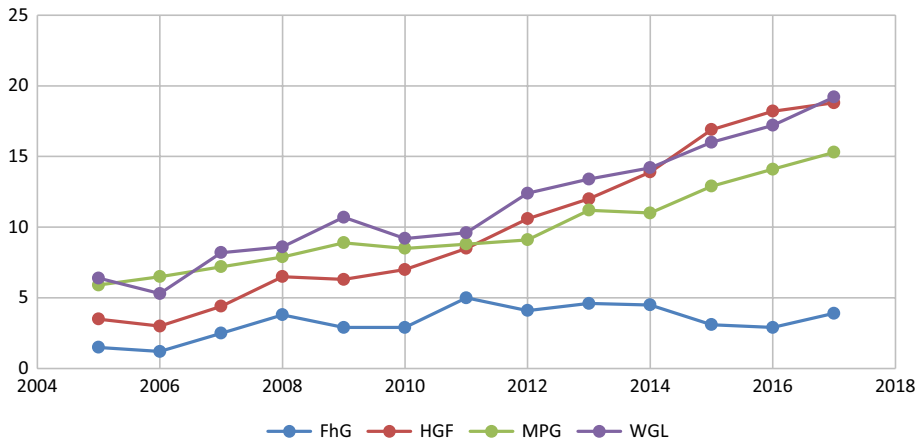




**Fig. 2** Theory of change for the “Women’s Professorship Programme”

those that have already participated in one of the precursor stages. If HEIs apply for funding a second or even third time, they have to describe the success and/or failure of previous gender equality measures and the lessons learned for the future. They also have to indicate the evaluation approaches planned for continuous monitoring. Furthermore, they have to describe how they intend to anchor their GE interventions in a sustainable way (BMBF 2018, Bekanntmachung). The programme grants funding to universities for initial appointments of women to tenured professorships at the rank of a full professor (W2 and W3 positions). Submitting a promising and tailored gender equality plan (and in later stages, providing evidence for its successful implementation) is the prerequisite to receive funding (BMBF 2018). The programme offers primarily financial resources: Each HEI with an approved gender equality plan (GEP) can receive funding for up to three professorships for a duration of 5 years maximum. The maximum sum for each professorship is 150,000 euros per year during the first two phases and 165,000 euros in the third phase. In round three, a maximum of 10 HEIs with top scores in the appraisal can receive further funding for a fourth professorship. The Women Professorship Programme (“Professorinnenprogramm”) was launched in March 2008 when the funding guidelines were announced. In June 2012, the Joint Science Conference (Gemeinsame Wissenschaftskonferenz, GWK) decided to continue the programme (“Professorinnenprogramm II”). The third round (“Professorinnenprogramm III”) was launched in February 2018 (BMBF 2018, Bekanntmachung). Figure 2 illustrates the ToC for the second German case. The main expected output in the area of gender equality is the increased probability for women to reach a top position and in the area of research an increase in the number of citations.

In the case studies, we focused on the development of the female researchers in leadership positions among the largest German research performing organisations, on the one hand and at the universities, on the other. As mentioned in the ToCs, one of the main

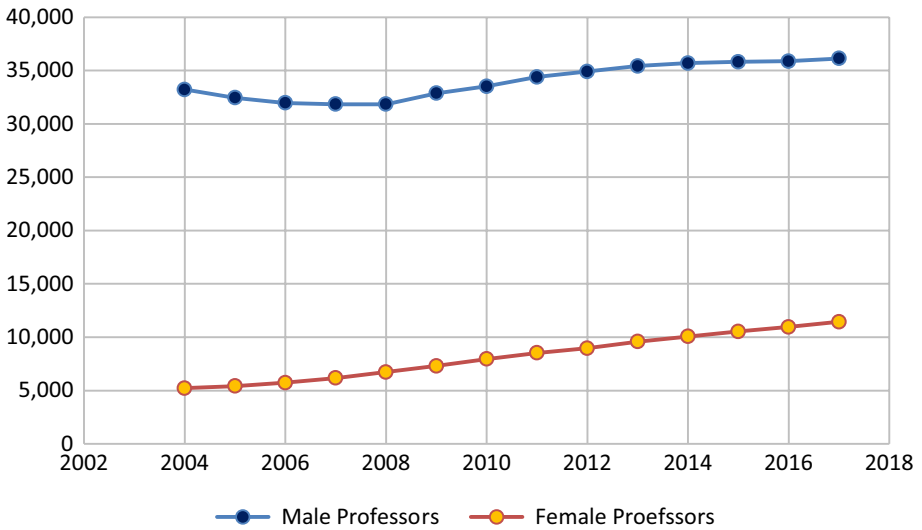


**Fig. 3** Share of women in leading positions at German non-university research institutions 2005–2017. *Source:* GWK (2012, 2018)

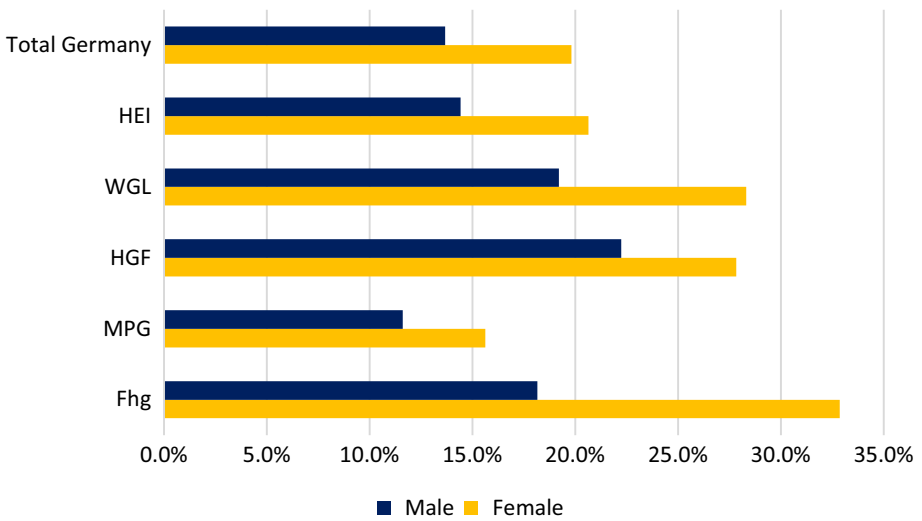
objectives of the “Women Professorship Programme” is to enhance the presence of women at all levels and to raise the numbers of women at the top of universities (see Fig. 2). The “Pact for Research and Innovation” (see Fig. 1) requires an increased representation of women in the science system, especially in leadership positions, as well. Furthermore, we looked at the publications and excellence rates of female researchers compared to their male colleagues. This was done because we expected, as listed in the assumptions displayed in the ToC figures, effects on the publication output if the research teams become more diverse. The following Fig. 3 shows the share of women in leading positions in the four big German research performing organisations. Due to an organisation-specific definition of three “management levels”, a comparison of the total figures between the Fraunhofer-Gesellschaft (FhG), the Helmholtz Association of German Research Centres (HGF), the Max Planck Society (MPG) and the Leibniz Association (WGL) is only possible to a limited extent. However, the share of women in management positions in science rose from a total of 2.0% in 2005 to 17.8% in 2017. What is also noticeable is that the increase occurred after the establishment of the “Pact for Research and Innovation” in 2005 (see also Bühner and Frietsch 2020).

Figure 4 shows the number of full-time professors at German universities between 2004 and 2017. 36,126 male full-time professors and 11,442 female full-time professors were employed at German universities in 2017. Here, too, it can be pointed out that the number of women professors has risen more dramatically than the total number of professorships. A particularly dynamic development can also be observed here that coincides with the launch of the women professorship programme in 2008.

According to the above, there is a significant increase of women in research (see Figs. 3, 4). Also the number of women (co)authors of scientific publications grew during the same period (see Fig. 5). Furthermore, a look at the most frequent quality indicators (also to be understood as longer-term impact), such as citations and excellence, reveals that the rates are roughly the same for male and female authors (see Bühner and Frietsch 2020). This finding is widely confirmed in the academic literature as well: women typically publish less frequently overall, but with higher quality as measured by citations (Campbell et al. 2013; Tower et al. 2011; Powell et al. 2009). These results enable us to show that more



**Fig. 4** Number of male and female professors 1999–2017. *Source:* <https://de.statista.com/statistik/daten/studie/160365/umfrage/professoren-und-professorinnen-an-deutschen-hochschulen>



**Fig. 5** Average annual growth rate in publications by women and men in Germany between 2005 and 2016. *Source:* Elsevier—Scopus; Fraunhofer ISI calculations.

**Table 2** Excellence rate of female and male authors in Germany. *Source:* Elsevier—Scopus; Fraunhofer ISI calculations

|              | 2005 (%) | 2006 (%) | 2007 (%) | 2008 (%) | 2009 (%) | 2010 (%) | 2011 (%) | 2012 (%) | 2013 (%) | 2014 (%) |
|--------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| FhG female   | 81       | 78       | 79       | 75       | 75       | 79       | 74       | 72       | 74       | 71       |
| FhG male     | 67       | 65       | 69       | 63       | 64       | 64       | 60       | 62       | 62       | 60       |
| MPG female   | 76       | 74       | 78       | 74       | 74       | 75       | 73       | 70       | 71       | 70       |
| MPG male     | 58       | 58       | 59       | 58       | 60       | 58       | 59       | 59       | 57       | 57       |
| HGF female   | 65       | 66       | 65       | 65       | 66       | 66       | 66       | 64       | 64       | 63       |
| HGF male     | 56       | 54       | 52       | 53       | 53       | 52       | 52       | 53       | 52       | 52       |
| WGL female   | 67       | 68       | 62       | 64       | 68       | 63       | 62       | 62       | 62       | 64       |
| WGL male     | 54       | 52       | 50       | 48       | 49       | 49       | 48       | 48       | 47       | 48       |
| HEI female   | 67       | 66       | 66       | 65       | 65       | 65       | 64       | 64       | 64       | 63       |
| HEI male     | 49       | 48       | 48       | 48       | 48       | 48       | 48       | 47       | 47       | 47       |
| Total female | 67       | 66       | 65       | 65       | 65       | 64       | 63       | 63       | 63       | 62       |
| Total male   | 51       | 50       | 50       | 49       | 49       | 49       | 49       | 49       | 48       | 48       |

women in the science system not only bring about a “gain in justice”, but also a concrete scientific benefit.

The excellence rates are measured as the shares of publications of an institution that belong to the top 10% of the most highly cited articles in their particular fields. Our analysis shows that women’s excellence rates are higher than that of men on average, for both the university and the non-university sector, as illustrated in Table 2 below (adopted from Bühner and Frietsch 2020).

## Discussion and conclusions

It is obvious that our evaluation approach reflects many of the limitations that are characteristic of the evaluation field. A key issue is that there are many factors intervening in the social space, which means that impact is indirect and long-term (Martin 2011; Reale et al. 2014). Thus, as pointed out in the discussion of our theoretical point of departure, it has not been our intention to establish a direct linear link between the flagship programmes and the results in terms of publications and excellence rates but to show the contributory nature of these kind of programmes to achieve gender equality effects.

Second, despite the fact that we compare our results to the results of other studies to validate them, we are aware that only quantitative methods are not enough in order to study complex phenomena as such as the flagship programmes, and their effects. Adhering to the gender-sensitive and complexity line of research, we acknowledge that further research is required to complement traditional quantitative measures with a gender-sensitive and reflexive design that requires a combination of quantitative and qualitative methodologies (Bustelo 2017; Espinosa 2013).

Based on a thorough analysis of the relevant knowledge in gender equality, evaluation as well as science and innovation research and the structured analysis of smart practice examples, we developed an evaluation framework with indicators for the assessment of GE interventions. Through the case studies of the flagship programmes, we tested and validated our evaluation framework and indicators. As part of the validation process, a theory of change has been developed for each of the programmes. The theories of change developed for the cases analysed herewith have shown that going a step further when investigating the effects of GE measures and looking at research outputs in terms of research and innovation, is a very innovative and promising approach. Concretely, the theory-based evaluation approach helped us not only to identify GE impact but also broader effects on research and innovation that might have remained undetected otherwise (compared for example to GESIS (2017) that focused on GE effects only).

There are clear gaps in the literature as regards the link between gender sensitive approaches and evaluation, and evaluations from a gender perspective are limited (Espinosa 2013). By combining a gender perspective with a theory of change approach, we position change in context and offer an in-depth reflection on assumptions of change. This approach has the potential to effectively address some of the challenges that evaluation of gender equality programmes has faced. It may support organisations and programme owners to develop evidence-based strategies. It may also support better monitoring to generate more elaborated data, moving away from only quantitative data towards evaluating gender disparities and advances articulated in theories of change.

Wrapping up, in accordance with our complexity, contribution and theory based approach we could identify broader impacts beyond pure gender equality effects. The cases

presented above show that, not least due to the two large national GE interventions, the role of women academics in the German publication landscape has changed significantly over the past 15 years and there has been a clear increase in the number of (co-)publications by female authors. Furthermore, although the overall number of women has also increased significantly since the introduction of the flagship promotional programmes of the “Women Professorship Programme” and “Pact for Research and Innovation”, it has not risen to the same extent as women’s participation in scientific publications. Thus, we can show that GE programmes may have broad positive effects on science as well (cf. Bühner and Frietsch 2020).

In line with our evaluation framework and approach, we have aimed at investigating the linkage between GE programmes and research and innovation in terms of contribution to change and in a probabilistic way, i.e. whether the programmes have created the conditions so that change can take place - and not in terms of attribution. We have thus good reasons to claim that the programmes have contributed not only to the higher shares of women within the research performing organisations, but—as a long-term impact—also to improved female publication patterns and citation rates.

**Acknowledgements** Open Access funding provided by Projekt DEAL. The paper is based on the H2020 funded project EFFORTI (Grant Agreement No. 710470). We thank two anonymous reviewers for suggestions that improved the article.

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

## References

- Bilimoria, D., & Wheeler, J. (2000). Women corporate directors: Current research and future directions. *Current Research Issues*, 2, 138–163.
- BMBF: Federal Ministry of Education and Science (Bundesministerium für Bildung und Forschung). (2016). *Pakt für Forschung und Innovation*. Retrieved March 13, 2018, from <https://www.bmbf.de/de/pakt-fuer-forschung-und-innovation-546.html>.
- BMBF: Federal Ministry of Education and Science (Bundesministerium für Bildung und Forschung). (2018). *Bekanntmachung. Richtlinie zur Umsetzung des Professorinnenprogramms des Bundes und der Länder zur Förderung der Gleichstellung von Frauen und Männern in Wissenschaft und Forschung an deutschen Hochschulen—Professorinnenprogramm III*. Bundesanzeiger vom 21.02.2018 <https://www.bmbf.de/foerderungen/bekanntmachung-1600.html> vom 13.03.2018.
- Brough, A. R., Wilkie, J. E., Jingjing, M., Isaac, M. S., & David, G. (2016). Is eco-friendly unmanly? The Green-Feminine Stereotype and its effect on sustainable consumption. *Journal of Consumer Research*, 43(4), 567–582.
- Bühner, S., & Frietsch, R. (2020). How do public investments in gender equality initiatives and publication patterns interrelate? The case of Germany. *Evaluation and Program Planning*, 79, 101769.
- Bühner, S., & Yorulmaz, M. (2019). *The manifold benefits of gender equality and (responsible) research & innovation*. Retrieved March 30, 2020, from [https://www.effort.eu/sites/default/files/2019-10/Effort\\_i\\_Overview\\_Performance\\_Effects\\_GenderDiversity.pdf](https://www.effort.eu/sites/default/files/2019-10/Effort_i_Overview_Performance_Effects_GenderDiversity.pdf).
- Bustelo, M. (2017). Evaluation from a gender+perspective as a key element for (re)gendering the policy-making process. *Journal of Women, Politics & Policy*, 38(1), 84–101.


- Campbell, L. G., Mehtani, S., Dozier, M. E., & Rinehart, J. (2013). Gender-heterogeneous working groups produce higher quality science. *PLoS ONE*, 8(10), e79147.
- Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford: Oxford University Press.
- Civitas. (2020). Smart choices for cities. Gender equality and mobility: mind the gap! Cleaner and better transport in cities (CIVITAS), 2020.
- Cosley, D., Forte, A., Cioffi, L., & McDonald, D. (Eds.). (2015). Proceedings of the 18th ACM conference on computer supported cooperative work & social computing. The 18th ACM conference. Vancouver, BC, Canada. New York, NY: ACM.
- Dahler-Larsen, P. (2012). *The evaluation society*. Stanford: Stanford University Press.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und evaluation in den sozial- und humanwissenschaftlichen*. Berlin: Springer.
- Espinosa, J. F. (2013). Towards a gender sensitive evaluation? Practices and challenges in international development evaluation. *Evaluation*, 19(2), 171–182.
- European Commission (Ed.). (2015a). *Indicators for promoting and monitoring Responsible Research and Innovation*. Brussels.
- European Commission. (2015b). She figures 2015: Gender in research and innovation. Brüssel. Retrieved September 30, 2017, from [https://ec.europa.eu/research/swafs/pdf/pub\\_gender\\_equality/she\\_figures\\_2015-final.pdf](https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf).
- European Union. (2016). *European innovation scoreboard 2016*. Brussels: European Union.
- European Union (Ed.). *Research innovation observatory*. <https://rio.jrc.ec.europa.eu/en/stats/key-indicators>.
- Eurostat. (2016). Average number of actual weekly hours of work in main job, by sex, professional status, full-time/part-time and occupation (hours). Retrieved from [http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa\\_ewhais&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_ewhais&lang=en).
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco: Jossey-Bass.
- GARCIA. *Gendering the academy and research: combating career instability and asymmetries*. <http://garciaproject.eu/>.
- GEDII. *Gender diversity impact—Improving research and innovation through gender diversity*. <https://www.gedii.eu/>.
- GENERA. *Gender equality network in the European research area*. <https://genera-project.com/>.
- GenPORT. (2017). *E-discussion: Gender balance in decision-making in research organisations*. Retrieved from <http://www.genderportal.eu/group/e-discussion-gender-balance-decision-making-research-organisations>.
- GESIS. (2017). *Evaluation des Professorinnenprogramms des Bundes und der Länder: Zweite Programmphase und Gesamtevaluation. Abschlussbericht. Januar 2017*. Beauftragt vom Bundesministerium für Bildung und Forschung. Download: [http://www.gwk-bonn.de/fileadmin/Papers/Evaluation\\_des\\_Professorinnenprogramms-Bericht\\_Januar\\_2017.pdf](http://www.gwk-bonn.de/fileadmin/Papers/Evaluation_des_Professorinnenprogramms-Bericht_Januar_2017.pdf). am 13.03.2017.
- GWK: Joint Science Conference (Gemeinsame Wissenschaftskonferenz). (2012). Pakt für Forschung und Innovation. Monitoring-Bericht 2012. Heft 28. <http://www.gwk-bonn.de/themen/wissenschaftspakte/pakt-fuer-forschung-und-innovation/>.
- GWK: Joint Science Conference (Gemeinsame Wissenschaftskonferenz). (2016). *Chancengleichheit in Wissenschaft und Forschung. 20. Fortschreibung des Datenmaterials (2014/2015) zu Frauen in Hochschulen und außerhochschulischen Forschungseinrichtungen*. Retrieved from <http://www.gwk-bonn.de/fileadmin/Papers/GWK-Heft-50-Chancengleichheit.pdf>.
- GWK: Joint Science Conference (Gemeinsame Wissenschaftskonferenz). (2018). *Pakt für Forschung und Innovation. Monitoring-Bericht 2018*. Heft 58. <http://www.gwk-bonn.de/themen/wissenschaftspakte/pakt-fuer-forschung-und-innovation/>.
- Halpern, D. F. (2014). It's complicated—In fact, it's complex: Explaining the gender gap in academic achievement in science and mathematics. *Psychological Science in the Public Interest*, 15, 72–74.
- HIVOS. (2014). Gender and theories of change: 4th E-discussion June 2014, end note.
- Horbach, J., & Jacob, J. (2017). The relevance of personal characteristics and gender diversity for (eco-) innovation activities at the firm-level. Results from a linked employer-employee database in Germany. *Business Strategy and The Environment*, 27(7), 924–934.
- Kalpaizidou Schmidt, E., & Graversen, E. K. (2020). Developing a conceptual evaluation framework for gender equality interventions in research and innovation. *Evaluation and Program Planning*, 79, 101750.
- Kalpaizidou Schmidt, E., & Cacace, M. (2017). Addressing gender inequality in science. The multifaceted challenge of assessing impact. *Research Evaluation*. <https://doi.org/10.1093/reseval/rvx003>.

- Kalpaizidou Schmidt, E., Bühner, S., Schraudner, M., Reidl, S. Müller, J., et al. (2018). *A conceptual evaluation framework for promoting gender equality in research and innovation. Toolbox I—A synthesis report. EFFORTI—Deliverable 3.3.*
- Knowlton, L. W., & Phillips, C. C. (2012). *The logic model guidebook: Better strategies for great results.* Thousand Oaks: Sage.
- McLaughlin, J. A., & Jordan, G. B. (2004). Using logic models. *Handbook of Practical Program Evaluation*, 2, 7–32.
- Martin, B. (2011). The research excellence framework and the impact agenda: are we creating a Frankenstein monster? *Research Evaluation*, 20(3), 247–254.
- Mayne, J., & Johnson, N. (2015). Using theories of change in the agriculture for nutrition and health CGIAR research program. *Evaluation*, 21(4), 407–428.
- Niessen, C., Sonnentag, S., Neff, A., & Unger, D. (2010). Ressourcen und Belastungen von Doppelkarrierepaaren in der Wissenschaft—eine arbeitspsychologische Perspektive. In E. Gramespacher, J. Funk, & I. Rothäusler (Eds.), *Dual Career Couples an Hochschulen Zwischen Wissenschaft, Praxis und Politik* (pp. 75–88). Barbara Budrich: Opladen.
- Nutley, S., Huw, D., & Walter, I. (2002). *Evidence based policy and practice: Cross sector lessons from the UK.* ESRC UK Centre for Evidence Based Policy and Practice. Working paper 9. London.
- OECD. (2014). *Science, technology and industry outlook 2014.* Paris: OECD Publishing.
- OECD. (2015). *Science, technology and industry scoreboard 2015. Innovation for growth and society.* Paris: OECD Publishing.
- Okamura, K. (2019). Interdisciplinarity revisited: evidence for research impact and dynamism. *Palgrave Commun.* <https://doi.org/10.1057/s41599-019-0352-4>.
- Powell, A., Hassan, T. M., Dainty, A., & Carter, C. (2009). Note. Exploring gender differences in construction research: a European perspective. *Construction Management and Economics*, 27(9), 803–807.
- PRAGES. *Practicing gender equality in science.* <http://www.pragesdatabase.eu/>.
- Ravn, T., Nielsen, M. W., & Mejlgard, N. (2015a). *Synthesis report on existing indicators across RRI dimensions.* Progress report D3.1. Monitoring the evolution and benefits of responsible research and innovation (MoRRI). [www.technopolis-group.com/morri/](http://www.technopolis-group.com/morri/).
- Ravn, T., Nielsen, M. W., & Mejlgard, N. (2015b). *Metrics and indicators of responsible research and innovation.* Progress report D3.2. Monitoring the evolution and benefits of responsible research and innovation (MoRRI). [www.technopolis-group.com/morri/](http://www.technopolis-group.com/morri/).
- Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., et al. (2017). A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. *Research Evaluation*, 27(4), 298–308.
- Reale, E., Nedeva, M., Thomas, D., & Primeri, E. (2014). Evaluation through impact: A different viewpoint. *Fteval Journal*, 39, 36–41.
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48.
- Rommens, E. (2014). Feminist interventions in the design process. In W. Ernst & I. Horwath (Eds.), *Gender in science and technology interdisciplinary approaches* (pp. 41–55). Bielefeld: Transcript Verlag.
- Sanderson, I. (2002). Evaluation, policy learning, and evidence based policy making. *Public Administration*, 80(1), 1–22.
- Schultz, I., & Stiess, I. (2009). *Gender aspects of sustainable consumption strategies and instruments.* EUP-OPP work package 1, deliverable 1.1. Frankfurt/Main: Institute for Social-Ecological Research (ISOE).
- Solesbury, W. (2001). *Evidence based policy: Whence it came and where it's going.* In: ESRC UK Centre for Evidence Based Policy and Practice (Ed.). Working paper 1. London.
- STAGES. Structural transformation to achieve gender equality in science.
- Timmers, T. M., Willemsen, T. M., & Tjijdens, K. G. (2010). Gender diversity policies in universities: A multi-perspective framework of policy measures. *Higher Education*, 59, 719–735.
- Tower, G., Plummer, J., & Ridgewell, B. (2011). A multidisciplinary study of gender-based research productivity in the world's best journals. *Journal of Diversity Management*, 2(4), 23–32.
- Yin, R. (1994). *Case study research. Design and methods* (2nd ed.). London: Sage Publications.
- Van Belle, S. B., Marchal, B., Dubourg, D., & Kegels, G. (2010). How to develop a theory driven evaluation design? Lessons learned from an adolescent sexual and reproductive health programme in West Africa. *BMC Public Health*, 10, 741.
- Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development.* London: Review Report.
- Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PLoS ONE*, 10(5), e0127298. <https://doi.org/10.1371/journal.pone.0127298>.



Xiao, C., & McCright, A. M. (2012). Explaining gender differences in concern about environmental problems in the United States. *Society and Natural Resources*, 25, 1067–1084.

## Affiliations

Susanne Bühner<sup>1</sup>  · Evanthia Kalpazidou Schmidt<sup>2</sup> · Rachel Palmén<sup>3</sup> · Sybille Reidl<sup>4</sup>

Evanthia Kalpazidou Schmidt  
eks@ps.au.dk

Rachel Palmén  
rpalmen@uoc.edu

Sybille Reidl  
sybille.reidl@joanneum.at

<sup>1</sup> Fraunhofer ISI, Breslauer Str. 48, 76139 Karlsruhe, Germany

<sup>2</sup> Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark

<sup>3</sup> Universitat Oberta de Catalunya, Av. Carl Friedrich Gauss 5, 08860 Castelldefels, Spain

<sup>4</sup> Joanneum Research, Sensengasse 1, 1090 Vienna, Austria