



INTEGRATING CONTEXT IN THE EVALUATION OF GENDER EQUALITY INTERVENTIONS AND BEYOND

EVANTHIA KALPAZIDOU SCHMIDT AND SUSANNE BÜHRER

DOI: 10.22163/fteval.2020.490

ABSTRACT

During the last few decades, the European Union and its member states have promoted gender equality policies in research organizations. However, evaluations of what policy interventions have accomplished have been limited. Based on a co-creation process, involving key European stakeholders, the H2020 funded EFFORTI project has aimed to fill this research gap by developing an evaluation framework to design and evaluate gender equality interventions in research and innovation. The analysis of interventions in a number of European countries has revealed the importance of considering the national and organisational contexts in gender equality policies and evaluation regimes. Context plays a crucial role when it comes to assessing outputs, outcomes and impacts of policy interventions, and needs to be factored into any explanation of change. Hence, evaluation approaches must ensure that design and evaluation of interventions are context-sensitive and theory-based to open the “black box” of programmes, thus avoiding that evaluation only looks at the effects without considering the underlying mechanisms producing them. Based on the experiences generated in the EFFORTI project, we herewith propose a context-sensitive logic model for interventions addressing gender based violence and harassment. The approach proposed may be an inspirational source for policy makers and evaluators working with gender equality interventions in complex contexts and within fields of study beyond gender equality.

BACKGROUND

Gender Equality (GE) is on the top of the EU agenda, as recently confirmed in the new GE strategy, “A Union of Equality: Gender Equality Strategy 2020-2025” (European Commission 2020). GE is also laid down in Article 8 of the Treaty on the Functioning of the European Union: “*In all its activities, the Union shall aim to eliminate inequalities, and to promote equality, between men and women*”. The scientific knowledge about the underrepresentation of women within research and innovation (R&I) systems and the most important hindrances are comprehensive and a

plethora of measures and toolkits exist that aim at the promotion of GE. In fact, progress has been made in almost all the EU countries during the past years, but at a rather slow pace. Thus, more evidence is needed that gender equality interventions have an impact and lead to improved R&I outcomes and processes, and innovations better suited to markets (European Commission 2017). Sound evaluation approaches can help to deliver such evidence.

The study presented herewith is based on an evaluation framework developed in the EFFORTI project¹ that adopts a theory of change approach which allows evaluators to systematically integrate various contextual factors in the design of evaluations (for more on the evaluation framework, see Kalpazidou Schmidt and Graversen 2020). One of the core assumptions of the developed evaluation framework is that context is important when evaluating the effects of gender equality measures in research and innovation systems (cf. Patton, 2008; Rog, 2012). Different contexts require different policies and measures to promote gender equality but also influence in various ways the effects of these interventions. The careful consideration of the context is also important for the transferability of results and the formulation of policy recommendations. By context, we mean the team, organisational and national contextual factors that facilitate or impede the intervention process. In this article, however, we focus on the latter level. By national context we mean the national or European framework conditions in force, which influence the effectiveness and efficiency of GE interventions. Typical contextual factors are the structure and competitiveness of the national research and innovation systems; the structure and main components of the welfare systems; labour market characteristics; existing gender equality policies including legislative measures; participation of women in tertiary education; horizontal and vertical gender segregation; gender pay gap or the share of women in decision-making positions (Reidl et. al 2019).

R&I INVESTMENTS, EVALUATION CULTURES AND GENDER REGIMES

Several countries were involved in this study and in the development and validation of the evaluation framework developed in EFFORTI, namely Austria, Denmark, Germany, Hungary, Spain and Sweden. These countries represent liberal, conservative, social democratic, Eastern and Western European model types of welfare regimes and research and innovation systems. This means, first, that they vary as to R&I expenditure, the share of female researchers, the size and variety of the R&I performing institutions and the R&I outputs and outcomes. A comparison between the studied countries reveals that Denmark and Sweden show a high constant R&I investment intensity level, followed by an increase in the proportion of women among researchers (2005-2015). Austria shows a similar pattern, but had a slightly lower increase in the share of female researchers during the same period. In Spain a slight decrease of R&I investment intensity has been noticed during the same period followed by a slight increase in the proportion of women among researchers while Hungary increased the per capita expenses on researchers but shows a decreasing share of women in R&I. Germany encounters a significant growth rate for female researchers between 2008 and 2015 (European Commission 2019) and a steady increase of R&I expenditures as well (OECD Research and Development Statistics RDS). Thus, a differentiated pattern emerges as to the development of the share of women researchers compared to the R&I investments (Striebing et al. 2020).

Second, these welfare state types represent different legal traditions, which have consequences for understanding the various policy approaches to promote gender equality. The countries represent hence different types of gender regimes in Europe, i.e. key policy logics of states in relation to gender (Striebing et al. 2020). Social democratic welfare countries, like Sweden and Denmark are considered as innovation leaders with high R&I investments, and public gender regimes that require gender mainstreaming and encourage a public debate about gender workplace equality and equal sharing of care responsibilities. Other countries, like Germany and Austria, which are strong innovators, are modelled on a conservative welfare state type characterized by a school system that requires one parent (usually the woman) to work part-time (Reidl et al. 2017) and modernizing domestic gender regimes. As regards Spain, which is a moderate innovator, gender equality policies in R&I appear to be comparable to Austria, showing a modernizing domestic gender regime, while Hungary with a low capacity to innovate shows an ambiguous gender regime (Striebing et al. 2020).

Third, the studied countries represent very different evaluation cultures, which have an influence on addressing and evaluating GE interventions. The development of an evaluation culture and capacity building vary among the studied countries in terms of establishing a vivid culture of evaluation and carrying out systematic evaluations of programmes and institutions. Thus, the intensity and quality of evalua-

tions of interventions is highly dependent on the degree of development of the national evaluation culture. Some interventions are not monitored - no data gathering mechanisms are built into the intervention, while others are monitored but no evaluation or impact assessment is carried out. In countries where strong evaluation cultures exist (such as Austria, Denmark, Germany and Sweden), the programmes are more comprehensively evaluated than for example in Spain or Hungary where the evaluation culture is weaker (Reidl et al. 2017).

METHODOLOGICAL APPROACH

Several methodological steps were undertaken to develop, test and validate the evaluation framework based on studies in seven national settings (Palmén et al., 2018). For the development of the framework, a thorough literature review was conducted, including the analysis of numerous smart practice examples. This step ended up with a comprehensive list of gender equality as well as research and innovation indicators that refer to the macro level (national level), the meso level (organisational) and the team level. In order to provide for relevant national context variables, a secondary data collection was carried out, where longitudinal data were identified, which describe the most relevant context variables for gender equality measures in European research and innovation systems. Concretely, besides the She Figures, numerous data from Eurostat, the Research and Innovation Observatory and the OECD were used to identify country differences. In addition, international comparative reports as well as national sources of information have been utilized that complemented the international datasets. Finally, expert interviews in the seven EU countries under investigation helped fill in still existing gaps that could not be covered by secondary data and sources.

This work was complemented with case studies to validate the evaluation framework. The case study work was based on the development of templates for the case description and common interview guidelines. Overall, 19 case studies of gender equality in R&I interventions were conducted, being implemented in Austria, Denmark, Germany, Hungary, Spain and Sweden. Each case study used documentary analysis and between four and twelve semi-structured interviews per case were carried out with policy makers, programme managers, practitioners and beneficiaries. A logic frame and a theory of change were developed for each case based on analysis of the design, implementation and impact assessment of each case. The feedback from the case study work supported the validation of the produced evaluation framework.

Based on the logic model and conceptual evaluation framework developed in the EFFORTI project, in the following section, we present an example of a stylised logic model for evaluation of interventions addressing gender based violence and harassment (GBVH) in research and innovation to illustrate the significance of the contextual factors for the effectiveness of interventions.

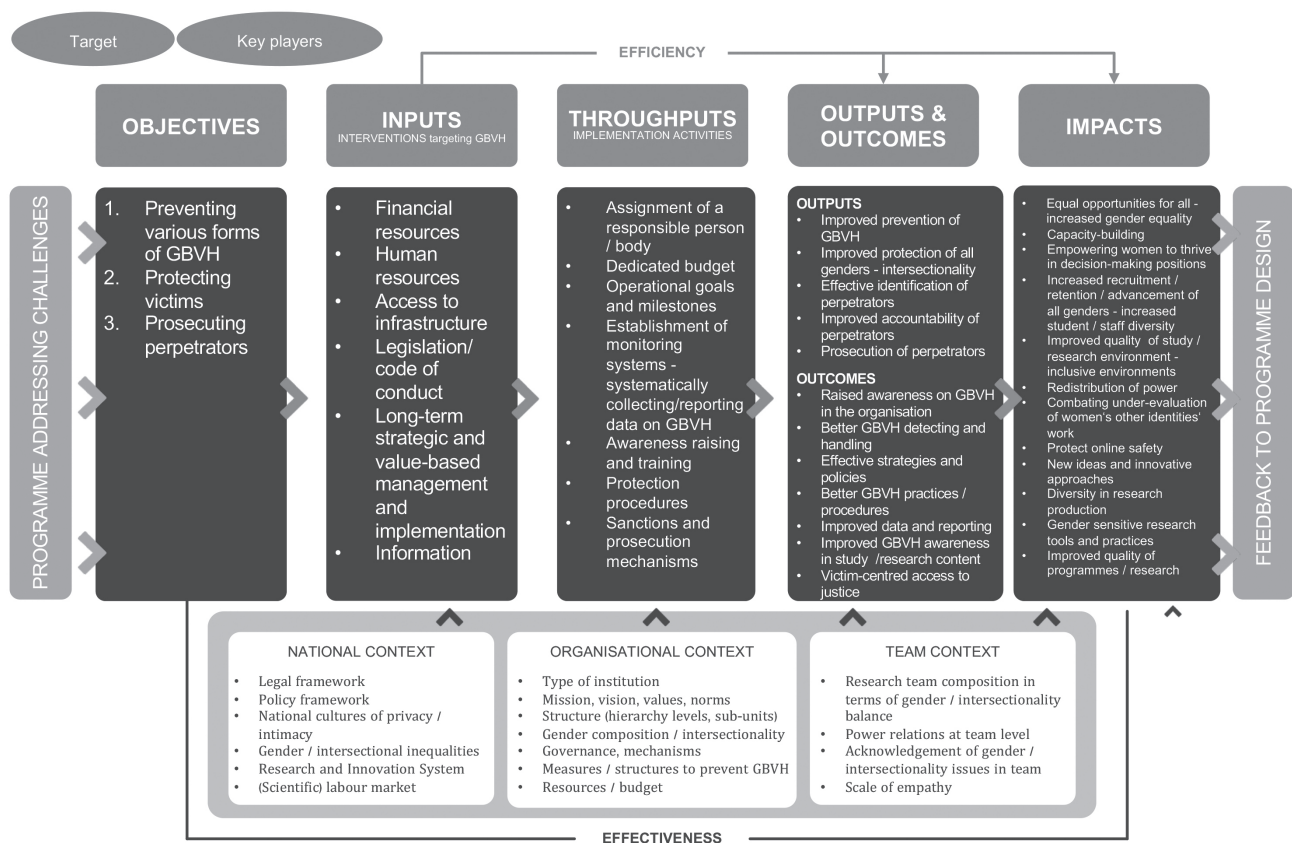
AN EXAMPLE OF A POTENTIAL CASE APPLICATION OF THE EFFORTI LOGIC MODEL

Context plays a crucial role when it comes to assessing measured outputs, outcomes and impacts of interventions. Adopting the theory of change approach enables and indeed requires factoring the context into any explanation of change (Kalpazidou Schmidt and Graversen 2020, Palmén et al. 2018)². Theory-based evaluations are used to open the “black box” of programmes, thus tackling complex settings. The “black box” issue refers to an evaluation approach that investigates programs primarily in terms of effects, without paying attention to how effects are produced (see Astbury and Leeuw 2010, Mathison 2005, Bush et al. 1995). Theory-based evaluations address in which way and under which conditions an intervention contributes to the intended and unintended impact (Döring and Bortz 2016). Theory-based approaches allow hence for assessing how particular means and instruments work to contribute to outcomes and impact in certain contexts. Theories are made explicit and evaluations build on elaborated assumptions while the engagement of all actors in the

evaluation process makes the different kinds of stakeholders’ assumptions transparent (Stame 2004). The theory of change approach uses evidence to identify, verify or challenge these assumptions and map the linkages between input, throughput, output, outcome, impact and context (Vogel, 2012a). For the evaluation practice, the theory of change is converted into an intervention logic model in order to reduce complexity and thus ease the empirical work and make feasible the concrete design and implementation of the evaluation of an intervention in a particular setting.

Below (figure 1) we apply the intervention logic model developed in the EFFORTI project to measures addressing gender-based violence and harassment (GBVH), offering a stylised model that incorporates the contextual factors in the intervention and evaluation process. Based on this intervention type, the presented model illustrates the various steps in the design and evaluation on the horizontal axis – inputs, implementation activities, and outputs, outcomes, and impacts – differentiated along three levels (team, organisational and country level), that allow for a comprehensive multi-level perspective, identifying the results of the policy measure and its preceding impact pathways. It is important to mention that the proposed logic model for evaluations of interventions addressing GBVH, although developed based on the experiences of the EFFORTI conceptual evaluation framework, has not yet been subject to concrete applications.

Figure 1: Stylised logic model for evaluation of interventions addressing GBVH



Source: own compilation

CONTEXTUAL CONDITIONS

This model is unique in its ability to account for potential contextual influence factors, at a team, organisational and national context level (see figure 1, three boxes at the bottom). As shown in the figure, the effectiveness of the intervention against GBVH is highly dependent on these frame factors. This consideration of contextual factors emphasises the frameworks' dynamic and responsive nature, and enables a more comprehensive and systemic depiction of the complex link between interventions aimed at improving conditions as regards GBVH and cope with its effects.

At the national level, countries have different types of research organisations (ROs) and higher education institutions (HEIs). The countries vary as well in terms of national legislative and policy contexts as regards GBVH and as to the mechanisms and instruments available to prevent and combat GBVH. This includes reporting, compiling data, and preventing GBVH, protecting victims and prosecuting perpetrators within HEIs and ROs, and in the awareness and willingness of their citizens to report GBVH (FRA 2014).

Countries vary in terms of legal references to potential biases including as to gender, minority status, age, parental status, academic position, sexual orientation, disability. They also vary in terms of their adoption of international frameworks that encourage, and in some cases, oblige states to take action to collect national data on sexual violence and harassment, to take measures to prevent violence and discrimination and to uphold the equality and dignity of their citizens. Such international frameworks include the UN Convention on the Elimination of all forms of discrimination against Women (CEDAW), the UN Convention on the Elimination of all forms of Racial discrimination (CERD) and the UN Convention on the Rights of Persons with Disabilities (CRPD) (United Nations 1965; 1979; 2006). Other international frameworks include the Council of Europe Convention on Preventing and combatting Violence against Women (FRA 2014), the EU Victims Directive and their orientation to the ILO global standards aimed at ending violence and harassment in the world of work (European Commission 2019).

The risks for GBVH are also maximised where structural inequalities exist and opportunities related to paid employment/academic success are ad hoc and dependent on individual powerful figures, who may enact toxic masculinities (national, organisational and team context). The difficulties of reporting GBVH and collecting data are also maximised in these circumstances (national and organisational context). Thus, particular attention needs to be paid to those in structurally dependent relationships, which are likely to facilitate GBVH: for example, relationships between PhD students and their supervisors; between exchange students or those in precarious positions and their line managers (team context). GBVH may also occur when differential cultural resources are drawn on by those in similar positions in the organisational hierarchy, for example, undergraduate students or minority staff (organisational and team context).

CONCLUDING REMARKS

Through the development and validation of a theory-based evaluation framework that aims to systematically integrate context factors at various levels, it became evident in the EFFORTI work that the context regarding research and innovation systems should always be taken into

account when designing gender equality policy interventions and evaluations. Research and innovation systems are complex involving a variety of stakeholders and vested interests, thus it is not always clear whether an intervention works in the anticipated way and which effects it has.

Ever complex interventions require increasingly complex evaluation approaches and acknowledging of the limitations of evaluations in dynamic contexts such as research and innovation systems. Complex systems involve multiple variables interacting in non-linear ways to produce outcomes and impacts. Evaluation is a complex concept in itself, implemented in complex systems. The design and instrumentation of evaluation must consider the complex systems in which the interventions operate. It is not adequate to present reality as a simple causal model - the models we use should address the complexity of systems (Kalpazidou Schmidt and Cacace 2019; Kalpazidou Schmidt et al. 2020).

The theory of change approach helps to mitigate the risks related to complexity in dynamic contexts and allows, at least to a certain extent, to open the "black box" – and answer the key questions, how and why a policy works, in which context and how to assess it. This particular way of designing and conducting evaluations based on a theory of change approach has proved to be a valuable tool to consider as to how different factors contribute to achieve impact.

Complex contextual conditions and choices as to evaluation approaches are characteristics also of other fields of study, different from GE in research and innovation. The approach proposed herewith may hence be an inspirational source for policy makers and evaluators working with interventions in complex contexts in other fields of study.

REFERENCES

- Astbury, B. & Leeuw, F.L.** (2010). Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation* 31(3), 363-381.
- Bush, Mullis, and Mullis** (1995). Evaluation: An Afterthought or an Integral Part of Program Development. *Journal of Extension*, 33 (2).
- Döring, N. and Bortz, J.** (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer.
- European Commission** (2017). Interim Evaluation: Gender equality as a crosscutting issue in Horizon 2020 Report of the Expert Group on "the Interim Evaluation of Gender Equality as a crosscutting issue in Horizon 2020" (retrieved from https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/interim_evaluation_gender_long_final.pdf, 02.07.2020).
- European Commission** (2019). She Figures 2018 (retrieved from <https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1/language-en>).
- European Commission** (2020). Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions: A Union of Equality: Gender Equality Strategy 2020-2025. Brussels 5.3.2020, COM (2020) 152 final (retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0152&from=EN>, 02.07.2020).

FRA: European Union Agency for Fundamental Rights (2014). Violence against women: An EU-wide survey: Results at a glance (retrieved from <https://fra.europa.eu/en/video/2014/violence-against-women-survey>).

Kalpazidou Schmidt, E. and Cacace, M. (2019). Setting up a dynamic framework to activate gender equality structural transformation in research organizations. *Science and Public Policy* 46: 321–38.

Kalpazidou Schmidt, E. and Graversen, E. K. (2020). Developing a conceptual evaluation framework for gender equality interventions in research and innovation. *Evaluation and Program Planning*, vol. 79.

Kalpazidou Schmidt, E.; Bühler, S.; Schraudner, M.; Reidl, S.; Müller, J.; Palmén, R.; Haase, S.; Graversen, E. K.; Holzinger, F.; Striebing, C.; Groó, D.; Klein, S.; Rigler, D.; Høg Utoft, E. (2017). A Conceptual Evaluation Framework for Promoting Gender Equality in Research and Innovation. A synthesis report. (retrieved from <https://www.efforti.eu/sites/default/files/2019-10/EFFORTI%20D3.3%20Conceptual%20Evaluation%20Framework%20Final.pdf>, 02.07.2020).

Kalpazidou Schmidt, E.; Ovseiko, P.V.; Henderson, L.R. and Kiparoglou, V. (2020). Understanding the Athena SWAN award scheme for gender equality as a complex social intervention in a complex system: analysis of Silver award action plans in a comparative European perspective. *Health Research Policy and Systems*, 18.

Mathison, S. (2005). *Encyclopedia of evaluation* (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc.

Mayne, J. and Johnson, N. (2015). Useful theories of change models. *Canadian Journal of Program Evaluation*, 119-142.

OECD Research and Development Statistics (RDS) 2020. (retrieved from: https://stats.oecd.org/Index.aspx?DataSetCode=GERD_TOE). <https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm>.)

Palmén, R.; Reidl, S.; Bühler, S., Kalpazidou Schmidt, E.; Holzinger, F., Groo, D.; Rigler, D., Striebing, C.; Arroyo, L.; Romano, M. J., Graversen, E. K.; Utoft, E. H.; Cecchini, M., Beranek, S.; Streicher, J.; Unger, M.; Koch, A. (2018). EFFORTI – Deliverable 4.2 Synthesis Report (retrieved from https://efforti.eu/sites/default/files/2019-05/D4.2_EFFORTI_Comparative_Report_28032018_FINAL.pdf, 02.07.2020).

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed). Thousand Oaks, CA: Sage 9.

Reidl, S.; Holzinger, F.; Streicher, J.; Beranek, S.; Unger, M.; Hafellner, S. (2017): EFFORTI – Deliverable 2.3. Comparative Background Report (retrieved from <https://www.efforti.eu/sites/default/files/2019-10/EFFORTI%20D2.3%20Comparative%20Background%20Report%20Context%20Analysis%20Final.pdf>, 02.07.2020).

Reidl, S.; Beranek, S., Palmén, R. (2019). What to consider regarding context? A Guideline (retrieved from https://www.efforti.eu/sites/default/files/2019-10/Guideline_context_final.pdf, 02.07.2020).

Rog, D. J. (2012). When background becomes foreground: Toward context-sensitive evaluation practice. *New Directions for Evaluation*, 135, 25–40.

Stame, N. (2004). Theory-based Evaluation and types of Complexity. *Evaluation*, vol. 10(1): 58–76.

Striebing, C.; Kalpazidou Schmidt, E.; Palmén, R.; Holzinger, F.; & Nagy, B. (2020): Women Underrepresentation in R&I. A Sector Program Assessment of the Contribution of Gender Equality Policies to Research and Innovation. *Evaluation and Program Planning*, vol. 79.

United Nations (2006). Convention on the Rights of Persons with Disability (retrieved from https://www.un.org/disabilities/documents/convention/convention_accessible_pdf.pdf).

United Nations (1979). Convention on the Elimination of All Forms of Discrimination against Women (retrieved from <https://www.ohchr.org/Documents/ProfessionalInterest/cedaw.pdf>).

United Nations (1965). International Convention on the Elimination of All Forms of Racial Discrimination. (retrieved from <https://www.ohchr.org/Documents/ProfessionalInterest/icerd.pdf>).

Vogel, I. (2012a). Review of the use of ‘Theory of Change’ in international development. UK: Department of International Development (retrieved from http://www.theoryofchange.org/pdf/DFID_ToC_Review_VogelV7.pdf, 02.07.2020).

Vogel, I. (2012b). ESPA guide to working with Theory of Change for research projects. (retrieved from <https://www.espa.ac.uk/files/espa/ESPA-Theory-of-Change-Manual-FINAL.pdf>, 02.07.2020).

AUTHORS

EVANTHIA KALPAZIDOU SCHMIDT

Aarhus University, Danish Centre for Studies in Research and Research Policy

Bartholins Allé 7, 8000 Aarhus C (Denmark)

E: eks@ps.au.dk

ORCID ID: 0000-0002-3204-0803

SUSANNE BÜHRER

Fraunhofer ISI

Breslauer Str. 48, 76139 Karlsruhe (Germany)

E: susanne.buehrer@isi.fraunhofer.de

ORCID ID: 0000-0002-1923-1437

Acknowledgements

The paper is based on the H2020 funded project EFFORTI (Grant agreement No 710470).

KEYWORDS

Evaluation, Gender Equality, Context, Logic Model, Gender-based Violence and Harassment, Research and Innovation