fteval JOURNAL
for Research and
Technology Policy Evaluation

May 2019, Vol. 47, pp. 140-145
DOI: 10.22163/fteval.2019.332

EVALUATION FRAMEWORK FOR PROMOTING GENDER EQUALITY IN RESEARCH AND INNOVATION:

HOW DOES GENDER EQUALITY INFLUENCE RESEARCH AND INNOVATION OUTCOMES AND WHAT IMPLICATIONS CAN BE DERIVED FOR SUITABLE EVALUATION APPROACHES?[1]

SUSANNE BÜHRER, SYBILLE REIDL, EVANTHIA K. SCHMIDT, RACHEL PALMEN, CLEMENS STRIEBING AND DORA GROO

DOI: 10.22163/fteval.2019.332

BACKGROUND AND PURPOSE OF THE STUDY

Despite all efforts undertaken in the past there is no comprehensive and rigorous analytical framework to consider all of the relevant variables in gender equality issues, although there have been a number of European Commission projects such as PRAGES, GENDERA, GenSET, STAGES and GENOVATE, which have explored the gender equality (GE) dimension with different foci. While all these previous studies have illustrated numerous evaluation approaches, concepts, indicators etc. to provide examples of measuring different kinds of impacts, a clear understanding of the mechanisms between different gender equality-related policy initiatives and interventions (inputs) and outputs/results is still not available. In order to address these challenges, EFFORTI (Evaluation Framework for Promoting Gender Equality in Research & Innovation), an EU funded project, aims to clarify the mechanisms between gender equality inputs and the expected results not only on gender equality itself, but also on research and innovation (R&I). The evaluation framework provides the theory and tools for analysing how gender equality-related interventions contribute to the achievement of the three European Research Area's main objectives on gender equality and how those achievements affect the desired outcomes of (responsible) research and innovation. The uniqueness of the evaluation framework is that it goes beyond conventional research and innovation indicators, taking into account also evaluation dimensions like providing answers to the Grand Challenges and the promotion of Responsible Research and Innovation.

With the rise of the idea 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 policy-making. At the same time, establishing causal relationships between policy interventions and observed changes poses a theoretical challenge as well as empirical and methodological problems. One approach to address these challenges is the theory-based impact evaluation approach (TBIE): In theory-based impact evaluation (TBIE), causality is often defined as a problem of contribution, not attribution. "Why and how" questions are typically being asked instead of "how things would have been without" like counterfactual approaches do. 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 the program and assessing its success by comparing theory with actual implementation. The "theories" to be investigated on how gender equality and R&I outcomes interrelate (intervention logics), which in turn link the allocation of resources to the achievement of intended results and finally impacts are still to be developed. These might be complemented by academic theories about public interventions and already existing empirical evidence from former evaluations and impact assessments. The actual results of GE policies will depend both on policy effectiveness and on other context variables. Context factors are organizational structures and cultures, as well as national and regional structures, capabilities and policies. The application of a theory based impact evaluation approach will allow us to take these different levels of influences on policy effectiveness - mechanisms and context - systematically into account. Furthermore, it allows us to develop context sensitive and policy specific theories of change.

1 Acknowledgement: The project EFFORTI that constitutes the basis for this contribution receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 710470. The authors would like to acknowledge the contribution of all project participants and all project activities to the ideas that underpin this paper.

METHODOLOGICAL APPROACH

Drawing on already developed and applied indicators in gender equality and R&I research (RIO Observatory, OECD STI Scoreboard etc.), but also on recent studies on RRI indicators (Ravn et al. 2015a, 2015b, European Commission 2015), we carried out a comprehensive desk research as a basis for the collection of a preliminary list of relevant indicators. Based on existing evidence, the project team first identified the most relevant indicators according to literature review; clustered these indicators into different categories, dimensions and sub-dimensions, which are based on GE-related literature and smart practice examples implemented in different organisations and contexts; and finally grouped these indicators according to an evaluation logic model. The indicators are differentiated between input, throughput, output, outcome and impact aspects. For each aspect, the indicators are illustrated at micro/ individual or team level, meso/organisational level and macro/policy or country level.

The indicators are based on the collection and review of "smart practices" implemented in Europe and beyond. The identification of smart practices was based on an assessment of the practices that are relevant, effective and efficient in the context that they operate in as to their quality of both evaluation and measurement (Kalpazidou Schmidt et al. 2017c). Smart practice examples evaluated measures of different nature 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 criteria of (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 (Kalpazidou Schmidt & Cacace 2017). Synthesising the typologies developed by Kalpazidou Schmidt and Cacace (2017) and the fields of action identified by the GENERA project and building on further theoretical and empirical experiences, we developed an intervention typology. Examples of impact stories were developed for a broad spectrum of these intervention types in order to provide examples of the mechanisms regarding intervention intentions and to provide a common framework for understanding the multi-faceted interventions of the cases that will serve as a testing ground for the further development of the tentative evaluation model.

CASE STUDY APPROACH FOR VALIDATION PURPOSES

The EFFORTI intervention logic model forms the conceptual basis for the case study work. The Intervention Logic Model considers inputs, throughputs, and outputs, as well as outcomes and impacts of the former two. The model also aims at showing how, once achieved, these objectives or effects can further affect desired R&I effects such as the number of patents and number of publications and citations, but also new R&I effects, such as providing answers to grand challenges and further promoting RRI. Additionally, the model includes three levels, i.e. team level (research quality, productivity, innovative outputs, and other RRI effects), organisational/ institutional level (workplace quality, recruitment capacity, efficiency, RRI orientation, competitiveness), and country/ system/ policy level (intensity, productivity, ERA orientation, etc.). However, some interventions will most likely overlap between different levels, which was taken into account in the development of the toolbox (EFFORTI Conceptual Evaluation Framework, D3.3, Kalpazidou et al. 2017.8). After having developed a first tentative evaluation framework, a series of case studies is foreseen to validate and further improve the model. 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). Case studies as a method have also been used extensively in evaluation research. We will use the case study method to inductively build on and validate the evaluation framework. The multiple case study work will shed light on those factors and mechanisms that shape and influence the effects of gender equality interventions in R&I on research and innovation outputs. It will attempt to explain what works (and what does not work) in what context and why. It will also explore whether the intervention is likely to work elsewhere and what is needed to make it work elsewhere. It will also attempt to explain how the national/ science system context influences the intervention in terms of the main contextual elements as well as the main agendas, strategies, and policies that shape the intervention. How the institutional context influences the intervention will also be taken into consideration – as will an assessment of whether the general conditions for effective gender equality policies are in place.

CASE STUDY EXAMPLE

In order to illustrate what insights can be gained from this approach for the Evaluation Toolbox, we subsequently outline exemplary results of a case study that addresses the ERA goal: *integration of gender dimension in research and education:*

"FEMtech Research Projects" is a funding scheme of the Austrian Research Promotion Agency (FFG), which supports projects in applied research, technology / product and process development that integrate the gender dimension in research content. It can be classified as an international good practice example in the context of fostering gender in research content because it is one of the few funding programmes that comprehensively targets the implementation of the gender dimension in scientific and technological research projects. By means of content analysis of project descriptions, interviews with policy designers, program managers and representatives of three funded research projects, the effects of this funding instrument especially its contribution to research and innovation were analysed.

In the beginning of the Case Study an intervention logic model was developed:

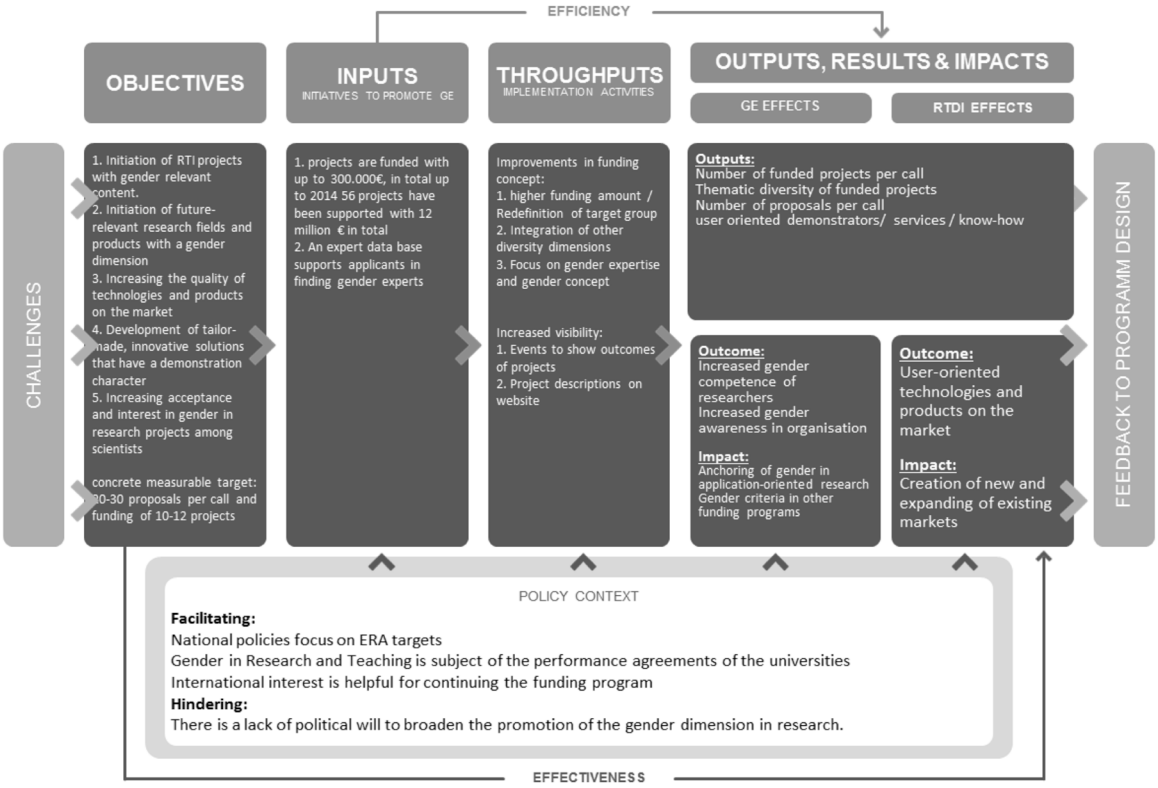

Figure 1. Intervention logic model "FEMtech Research Projects"

This log frame shows that the funding programme aims to initiate RTDI projects with gender relevant content in future-oriented research fields and in the long run wants to enhance the quality of technologies and products on the market. Funded projects shall develop tailor-made, innovative solutions that have a demonstration character. In addition, "FEMtech Research Projects" aims to increase acceptance and interest in the topic of integrating gender in research among scientists.

All in all 10-12 projects shall be funded per call with a maximum funding of 300.000€ per project. From 2008 until 2014, 7 calls have been launched. In total 56 projects were funded with a sum of 12 Mio. €.

Moreover, the logic model shows the expected outputs, outcomes and impacts of the funding scheme as they are stated in program documents[2] and formulated in interviews with policy makers and program managers (codes of interviewees: AU_CS2_02, AU_CS2_03, AU_CS2_04).

In the case study, it was then tried to measure possible outputs, outcomes and impacts of the funding program. The monitoring data for all seven calls between 2008 and 2014 shows a constantly rising number of submitted proposals with only one outlier in 2011.

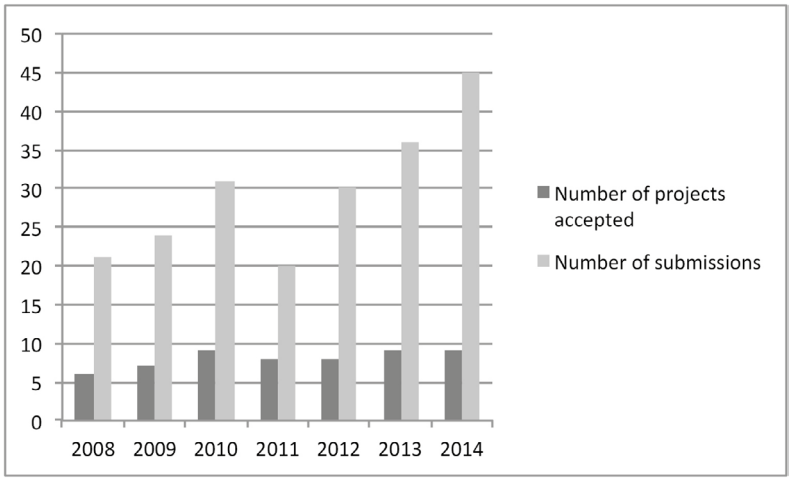

Figure 2: number of submitted and accepted projects per year
Source: FFG

1 See https://www.ffg.at/femtech-forschungsprojekte/5-ausschreibung

The number of funded projects stayed between six and nine per call because the amount of funding distributed was too low to fund more projects. The funded projects spread over six thematic categories from Energy/Ecology over Life Science to IT/communication, which indicates a thematic diversity of funded projects.

The measure's short-term output consists in the integration of gendered user involvement activities into technology development processes like gender divided test groups, gendered needs assessments, usability tests, participatory co-designing etc., ideally from the very beginning (see Nedopil/Schauber/Glende 2013; Rommes 2014). The result of this changed technology development process is information on gender-specific (and diversity-specific) user requirements for the product to be developed. But the analysis of the online-project descriptions shows, that not all of the projects focus on developing products, there were also studies funded to gain more gender-specific knowledge and projects that developed gender specific services. As the funded FEMtech research projects exhibit a broad range of research foci and project durations, the produced outputs vary widely as well. Also, most projects generated not one but several types of results. The most commonly produced result is the review of a product or service from a gender perspective. This can be explained with the considerably low funding of € 300,000 per project, which does not really make the development of a new product possible (AU_CS2_15). Less common are tutorials, didactic concepts / training concepts or manuals.

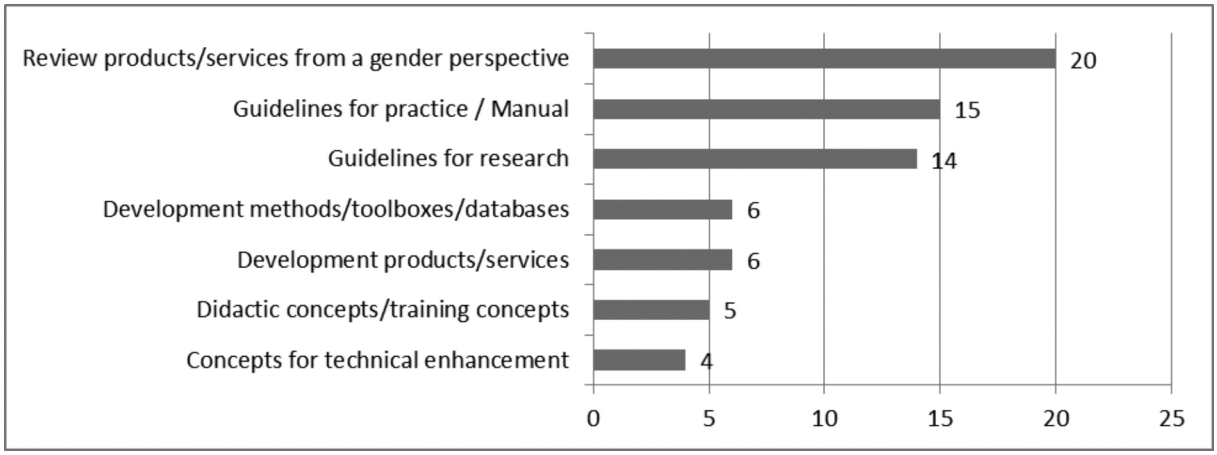

Figure 3: Type of results (number of projects)
Source: https://www.femtech.at/projekte, Analysis Joanneum Research

As many of the funded projects do not aim to develop new or improve existing products the outcome of the funding program cannot be measured only by counting user-oriented products and technologies on the market as it was expected in the intervention logic model. Another reason why this indicator is not useful is, that "FEMtech Research Projects" does not fund development processes until market entry. Therefore, information about the potential further development process after funding has ended is not available. Instead, in the case study the outcome of "FEMtech Research Projects" was measured by identifying different forms of further usage of project results. In 19 funded projects, starting points for further research were identified. 18 projects plan an application of project results in practice; another 12 are committed to apply the project results. Moreover, one interviewee reports of a market launch of a developed service in the upcoming months (AU_CS2_09).

Figure 4: Type of further use of results (number of projects)
Source: https://www.femtech.at/projekte, Analysis Joanneum Research

Follow-up projects are mentioned only in three project descriptions but are an important issue in the interviews with project representatives. Seven interviewees report of having already submitted another FEMtech project or a follow up project in another funding program.

In the interviews, also other outcomes of the projects were mentioned: the researchers gained gender competence in the course of the project and became self-confident regarding its practical use. For the research organisation the FEMtech Research Project also means a reference for further gender project applications. Some interviewees could also improve the scientific quality of the gender analysis in their research proposals for other funding schemes. The gained knowledge and competences regarding gender and/or new research methods were used in teaching, trainings and other research projects. Most interviewees also mentioned a sensitization of researchers regarding interdisciplinary and/or participative research through these projects.

This directly refers to the program target of increasing acceptance and interest in gender in research projects of scientists and the expected gender equality impact of anchoring gender in application-oriented research. The interviews showed an increasing awareness of the relevance of the gender dimension in research and also an increase of gender knowledge of researchers and representatives of companies who participated in the projects. This is also confirmed by the increasing quality of applications.

To investigate whether the funding program contributes to anchoring the gender dimension in application-oriented research, a social network analysis of funded organisations was conducted. It was examined whether the group of beneficiaries has grown from call to call. This analysis shows an expansion of research organisations and companies participating in "FEMtech Research Projects". This "spreading" is frequently happening via actors that submit regularly in the funding line, but with changing cooperation partners. It can be assumed that they have a multiplier function to involve other organisations in dealing with the gender dimension in research. But they can also take on a gatekeeper function in the future because they gain a substantial knowledge advantage, whereby other applicants with less experience are no longer competitive. All in all the community of organisations that already have conducted a FEMtech research project is still rather small compared to the number of research performing organizations in Austria. This can mainly be attributed to the fact that this funding scheme is rather small compared to other RTDI funding programmes. The lack of political will to broaden the promotion of the gender dimension in research, which was identified as a relevant context factor in the intervention logic model, reduces the expected impact of the "FEMtech Research Projects".

In the interviews, further possible effects of "FEMtech Research Projects" are addressed, which could not be investigated in the case study due to limited resources: FEMtech may have contributed to the implementation of gender criteria in other funding programmes. And interviewees report that organisations who conducted a FEMtech research project later on submit a proposal for a FEMtech Career project to start organizational change towards gender equality. This could not be investigated in this case study, as FFG monitoring data for other funding instruments could not be accessed.

We have presented some results of the "FEMtech Research Projects" case study and will now draw some conclusions from the case study for the development of the Evaluation Framework in EFFORTI: The case study shows that some expected effects cannot be detected due to lack of data. However, alternative ways of measuring outputs, outcomes and impact could also be demonstrated. The case study represents the first attempt to measure RTDI effects of FEMtech Research Projects as so far only concept and implementation evaluations have been carried out for this instrument. Indicators for impact assessment were outlined and will be included in the EFFORTI Evaluation Framework. With access to further funding data, impact measurement could be further developed.

DISCUSSION AND CONCLUSIONS

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, a first evaluation framework has been developed which was then used for the conduction of in total 19 case studies in seven EU countries (Austria, Denmark, France, Germany, Hungary, Spain, Sweden). The case studies cover a broad range of gender equality interventions, from mentoring instruments over structural change approaches up to incentives for integrating gender aspects into research and innovation projects. With this case study approach, we aim to validate and further develop the evaluation framework, a process of which the most recent results shall be shown and discussed at the Vienna Impact Conference.

Our approach of using a theory-based evaluation framework is appropriate even though it has hardly possible to measure concrete research and innovation outcomes and impacts of the GE programmes under consideration. One critique, however, can be that the theory of changes emphasizes differences between male and female researchers and might lead to the promotion of stereotypes. Furthermore, the work with log frames is rather linear and only partly suitable for complex environments, as we are fully aware.

The main and still unresolved problem is how to establish the link between the intervention and the research and innovation outcomes and impacts. Apart from subjective perceptions and anecdotal evidence, the interviewees could not contribute any confirmations.

The case studies underlined, however, the importance of the context yielding to the desired but also to some not desired effects. They also showed that the EFFORTI approach and the collection of indicators delivers a suitable background for programme evaluations.

REFERENCES

European Union (2016). European Innovation Scoreboard 2016. Brussels.

European Union (Ed.). Research Innovation Observatory (https.//rio.jrc.ec.europa.eu/en/stats/key-indicators).

European Commission (Ed.) (2015). Indicators for promoting and monitoring Responsible Research and Innovation. Brussels.

Kalpazidou Schmidt, E. and Cacace, M. (2017). Addressing gender inequality in science. the multifaceted challenge of assessing impact. Research Evaluation 2017, 1–13, doi. 10.1093/reseval/rvx003.

Kalpazidou Schmidt, E., Bührer, S., Schraudner, M., Reidl, S., Müller, J., Palmen, R., Haase, S., Graversen, E. K., Holzinger, F., Striebing,

C., Groó, D., Klein, S., Rigler, D. and Høg Utoft, E. (2017). A Conceptual Evaluation Framework for Promoting Gender Equality in Research and Innovation. Toolbox I - A synthesis report. EFFORTI - Deliverable 3.3.

Nedopil, C., Schauber, C. and Glende, S. (2013): The art and joy of user integration in AAL-projects. Brussels.

Nutley, S., Huw, D. and Walter, I. (2002). Evidence Based Policy and Practice: Cross Sector Lessons From the UK. Eds.. ESRC UK Centre for Evidence Based Policy and Practice. Working Paper 9. London.

OECD (2014). Science, Technology and Industry Outlook 2014, OECD Publishing. Paris.

OECD (2015). Science, Technology and Industry Scoreboard 2015. Innovation for growth and society, OECD Publishing. Paris.

Ravn, T., Nielsen, M. W. and Mejlgaard, 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) (http.//www.technopolis-group.com/morri/).

Ravn, T., Nielsen, M. W. and Mejlgaard, 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) (http.//www.technopolis-group.com/morri/).

Rommes, E. (2014): Feminist Interventions in the Design Process. In: Ernst, W. and Horwath, I. (Eds.): Gender in Science and Technology. Interdisciplinary Approaches. Transcript Verlag, Bielefeld. 41-55.

Sanderson, I. (2002). Evaluation, Policy Learning, and Evidence Based Policy Making, Public Administration 80(1), 1-22.

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.

Yin, R. (1994). Case Study Research. Design and Methods, Second Edition, Sage Publications, London.

AUTHORS

SUSANNE BÜHRER
Fraunhofer ISI
Breslauer Str. 48, 76139 Karlsruhe (Germany)
E: susanne.buehrer@isi.fraunhofer.de

SYBILLE REIDL
Joanneum Research
Sensengasse 1, 1090 Wien (Austria)
E: sybille.reidl@joanneum.at

EVANTHIA K. SCHMIDT
Aarhus University
Bartholins Allé 7, 8000 Aarhus C (Denmark)
E: eks@ps.au.dk

RACHEL PALMEN
Universitat Oberta de Catalunya
Av. Carl Friedrich Gauss 5, 08860 Castelldefels (Spain)
E: rpalmen@uoc.edu

CLEMENS STRIEBING
Fraunhofer CeRRI
Hardenbergstr. 20, 10623 Berlin (Germany)
E: clemens.striebing@iao.fraunhofer.de

DORA GROO
Association of Hungarian Women in Science
Napraforgo u 17, 1021 Budapest (Hungary)
E: dora.groo@nokatud.hu