

How to Implement Context-Sensitive Evaluation Approaches in Countries with still Emerging Evaluation Cultures

Susanne Buehrer

*Fraunhofer Institute for Systems and Innovation Research
Breslauer Str. 48, 76139 Karlsruhe, Germany*

Evanthia Kalpazidou Schmidt

*Aarhus University
Bartholins Allé 7, DK - 8000 Aarhus, Denmark*

Dorottya Rigler

*Association of Hungarian Woman in Science
1121 Budapest, Hungary*

Rachel Palmen

*Universitat Oberta de Catalunya,
Av. Carl Friedrich Gauss 5, ES - 08860 Castellde-fels, Spain*

crossref<http://dx.doi.org/10.5755/j01.ppaa.20.3.28371>

Abstract. *Evaluation cultures and evaluation capacity building vary greatly across the European Union. Western European countries, such as Austria, Germany, Denmark and Sweden, have been termed as leading countries in the evaluation as they have built up well-established evaluation cultures and carry out systematic evaluations of programmes and institutions. In contrast, in Central and Eastern European (CEE) countries, efforts continue to establish evaluation practices and further develop the current evaluation culture. In Hungary, for example, an established research and innovation evaluation practice does not exist, not one specifically considering gender equality in research and innovation evaluations with the exception of research and innovation programmes financed by the EU Structural Funds. Based on the results of a Horizon 2020 project, we apply a context-sensitive evaluation concept in Hungary that enables program owners and evaluators to develop a tailor-made design and impact model for their planned or ongoing gender equality interventions. The development of this evaluation was based on a thorough analysis of the literature and 19 case studies, building on documentary analysis and semi-structured interviews. The article shows that this evaluation approach is applicable also in countries with a certain catch-up demand of the existing overall evaluation culture. The special feature of the presented evaluation approach is, on the one hand, that the evaluation is context-sensitive. On the other hand, this approach makes it possible not only to depict effects on gender equality itself, but also to anticipate effects on research and innovation. Such effects can, for example, be a stronger orientation of research towards societal needs, which makes it particularly interesting for private companies.*

Keywords: *evaluation, impact assessment, emerging evaluation cultures, gender equality, private sector.*

Raktažodžiai: *vertinimas, poveikio analizė, naujosios vertinimo kultūros, lyčių lygybė, privatus sektorius*

Introduction

This article examines how evaluation of measures promoting gender equality can be carried out in countries with national research and innovation systems that still need to catch up, especially

with regard to the development of a comprehensive evaluation culture and practice. The focus of the Hungarian research and innovation system was chosen because of its relevance for the development of Europe as a whole and because gender equality is particularly in need of improvement here, especially within the business sector.

The article is structured as follows: In the subsequent section, we present recent findings on differences in evaluation cultures and practices across EU countries in general and describe the current situation in Hungary in particular. Here, we focus on issues related to the evaluation of gender equality as well. In the next section, we present our conceptual and methodological approach in developing a context-sensitive evaluation framework. We then present the results from the evaluation of a gender equality intervention in Hungary. The article wraps up with a discussion and conclusions.

State of the Art

The role of evaluations in European Research and Innovation Systems

The importance of research and innovation (R&I) for the prosperity and wealth of countries and their citizens is widely recognised. Numerous policy interventions have been designed and implemented in Europe to address systems and/or market failures (for an overview, see Research and Innovation Observatory, <https://rio.jrc.ec.europa.eu/> or OECD STI Compass, <https://stip.oecd.org/stip.html>). With the rise of the idea of evidence-based policy-making (Nutley et al. 2002; Solesbury 2001; Sanderson 2002), expectations have grown regarding the use of scientific evidence in policy-making, and evaluations are increasingly perceived as an instrument of evidence-based strategic intelligence (Kuhlmann et al. 1999). In this regard, evaluations may help to justify or re-direct funding, making public spending more efficient and effective but also enabling learning (Kuhlmann 2003). The growing importance of evaluations is reflected by the plethora of handbooks that were published since the turn of the century (OECD 1998; Fahrenkrog et al. 2002; Ruegg and Feller 2003; Miles et al. 2005).

Boras and Laatsit (2019) point out that different normative models about evaluation in a systemic perspective exist (Edler et al. 2008, Magro & Wilson, 2013, Arnold, 2004; Jordan et al., 2008; Hage et al., 2007) but that there is an increasing consensus about the need to take a systemic perspective (Smits and Kuhlmann 2004, Feller 2007). Furthermore, the authors emphasise the gap between innovation systems theory and the practice of evaluations, in particular the lack of theoretical and conceptual underpinning of evaluations (Molas-Gallart and Davies 2006). Boras and Laatsit (2019, 314) conclude: “We see system oriented innovation policy evaluation as a fundamental tool for creating strong, comprehensive and strategic policy advice”. However, whereas some European countries made good progress towards establishing an evaluation culture in their national innovation systems, others still lag behind (Edler et al. 2012, 2010).

Boras and Laatsit (2019) identify various groups depending on the four analytical criteria that they use to identify such systemic approaches. One of the criteria is temporality, which can be seen as a crucial aspect of country differences in the field of evaluation (Dahler-Larsen 2012). Boras and Laatsit (2019, 314) define temporality as “the extent to which there is a certain level of regularity in the evaluation of the three coverage elements (policy instruments, policy mix and socio-economic performance) and of the reports with systemic perspective”. The measures scores range from 2 (evaluations are conducted with a high level of regularity) over 1 (some evaluations are conducted regularly, but others more sporadically) to 0 (evaluations are done sporadically and ad hoc). Their analysis resulted in the following three groups: (1) Group 1 comprises Austria, Finland, Germany, Ireland, the Netherlands, the UK and Sweden. In these countries, evaluations are not only conducted regularly and frequently but also in a variety of ways. (2) The second group consists of Belgium, Denmark, Estonia, France, Hungary, Latvia, Lithuania, Poland, Slovenia, and Spain. Here, evaluations also take place frequently but less often than in the first group. (3) Finally, group 3 comprises Croatia, Czech Republic, Cyprus, Greece, Italy, Luxembourg, Malta, Portugal, Romania and Slovakia. In these countries, evaluations are conducted rather infrequently. Taking all four

criteria for system-oriented innovation policy evaluations, only six out of the investigated 28 EU countries meet the requirements (the Netherlands, Austria, Finland, Germany, Ireland and Sweden) whereas eight further countries (Denmark, France, Belgium, Poland, the UK, Estonia, Lithuania and Slovenia) are strong in instrument evaluation but lack policy-mix evaluations and assessment of the socio-economic performance. Furthermore, the regularity is somewhat smaller than in the first group. Five countries, among them Hungary (the others are Latvia, Spain, Czech Republic and Portugal), are catching up, attempting to make broad use of the available expertise and knowledge and thus laying the ground for future system approaches but currently showing an uneven regularity and reduced coverage of evaluation topics.

Based on a comparative analysis of seven selected countries (Austria, Denmark, France, Germany, Hungary, Spain and Sweden), Reidl et al. (2018) point out that evaluations are broadly routinised activities in R&I policy in countries like Germany, Austria, Denmark and Sweden and at least perceived as instruments of evidenced-based policy in France, Spain and Hungary. However, the objectives that are pursued with evaluations are slightly different. In Sweden and Denmark, evaluations continue to seek legitimation and accountability, reflecting the overall control-oriented purpose of evaluation. The Spanish approach focuses on achieving efficiency and accountability in all public administration actions. France seems to align its evaluation activities with the New Public Management paradigm and international standards in evaluation, paying more attention to input/output relationships for public funding. Hungary, described in more detail below, is working to build up a broader public policy evaluation culture.

The establishment of professional organisations and networks helped to promote evaluation as a topic in all seven countries. Evaluation societies were founded in Austria (1996), Germany (1997), France (1999), Denmark (2000), Spain (2001), Sweden (2003) and Hungary (2006) (Bachtler 2012). In some cases, the evaluation societies developed evaluation standards that are partly used in calls for tender as quality criteria (Germany, Austria). However, the mere existence of a professional, evaluation-fostering organisation does not necessarily lead to a comprehensive national discourse about evaluation as there are several barriers, such as an overall lack of communication and discussion, a missing collective vision and initiatives that could cut across the traditional boundaries of evaluation societies (Jacob et al. 2015). Reidl et al. (2018) also highlight recent developments towards further improvements of the respective evaluation cultures within European R&I systems, acknowledging that all countries aim to improve and promote evaluations as means of policy advice. The most important developments are:

- Increasing demand for and interest in impact assessments and evidence of effectiveness.
- An acknowledged need to further develop the existing evaluation methods towards more sophisticated approaches, including control group approaches, network analysis and econometric analysis, and improved methodologies to identify mid- and long-term impacts.
- The latter relates to an observed need of a better understanding of socio-economic effects in complex settings, being aware that the complexity of measures and programmes as well as the emergence of new topics in R&I challenge evaluation research and practice.

Despite the fact that conducting evaluations has evolved into becoming a vital profession in the countries mentioned above, there is a growing demand for capacity-building and skill development (Polverari 2015). Additionally, an increasing trend towards improved stakeholder engagement has been noticed in many countries (Pattyn 2017, Vedung 2010).

As regards the role of gender equality in R&I evaluations, Reidl et al. (2018) point out that the European R&I systems still lack gender-related initiatives and policies and thus do lack gender-sensitive evaluations in that field. This means that gender equality is usually neither considered in R&I policies nor in R&I-related evaluations, with some few exceptions where gender equality is part of the overall programme objectives or perceived as an overall strategic issue. Again, some critical country differences occur: Sweden and Denmark, for example, can rely on their long-lasting experience to promote gender equality and guidelines for reviewers, developed with a particular focus

on reducing gender bias. Nevertheless, even in these two countries, R&I programme evaluations do not (yet) take gender equality into account sufficiently. In France and Spain, strategic decision-makers at various councils and associations have developed road-maps with a long-term vision of improving gender equality, but these documents lack specific guidelines on how progress shall be achieved and measured. In Hungary, no prescriptions and no practice concerning the evaluation of gender equality in R&I exists. This is discussed in more detail in the following section.

The context for R&I evaluations in Hungary

In this article, we use Hungary as an illustrative case for a catch-up country as regards evaluation culture and practices. Thus, in this section, we discuss the context for R&I evaluations in Hungary providing a general frame for the following analysis. According to the RDI Act (Act LXXVI of 2014 on scientific research, development and innovation), NRDI Office (National Research, Development and Innovation Office) is in charge of developing the national strategy for scientific research, development and innovation. The tasks of the NRDI office include the strategic planning of calls for proposals for R&I projects and the coordination of the application process. It is also responsible for program evaluation and the design of the monitoring system.

The principles of the RDI Act foresee a regular monitoring and independent evaluation of support measures in the field of R&I. These monitoring and evaluation activities shall be in line with the principles of the Europe 2020 – Innovation Union document. The president of the NRDI Office is responsible for initiating evaluations and for monitoring the impact of R&I funding programmes. However, the availability of formal evaluation reports for R&I policies, programmes or support measures and the visibility of those evaluations in the R&I policy discourse remains limited. In practice, all the evaluations conducted so far have been ex-post evaluations.

A general review of the Hungarian innovation policies was carried out by the OECD (OECD, 2008). Its main recommendations included to improve the framework conditions for innovation, strengthen the human capital base for science, technology and innovation, improve the governance of the innovation system, promote business-sector-innovations, strengthen the connections within the innovation system, foster critical mass, excellence and relevance in the public research sector and maximise the benefits from international R&I collaboration.

As regards the international peer review of the Hungarian Scientific Research Fund's (OTKA, Országos Tudományos Kutatási Alapprogramok, integrated since 2015 into the operations of the NRDI Office) research proposal evaluations – whilst competence and the scientific communities' involvement were applauded, selection processes were flagged to be reviewed regarding 'fairness' and 'impartiality' (ESF, 2014:26).

The overall number of external evaluations of R&I programmes and institutions in Hungary remains limited, and the Hungarian government aimed to change this by pre-paring a dedicated regulation concerning the evaluation of R&I programmes. Moreover, the Monitoring Committee of EU-funded Operational Programmes (2014-2020) was expected to approve a comprehensive plan for monitoring and evaluation of these programmes, but there is no recent information about the process. According to the Research and Innovation Observatory (RIO country report 2015 Hungary p. 28), the impact assessment of the R&I funding programmes (funding period 2007-2013), initiated by the NRDI Office, shall serve to the further development of the evaluation culture, even if not published yet.

A further weakness of the Hungarian R&I system besides the issue of evaluation consists in a lack of strategic intelligence, i.e. the seldom use of tools like explicit foresight activities to design R&I policy concepts and strategies. However, the Smart Specialisation Strategy (S3) White Book foresaw the use of such innovation policy decision-making tools, i.e. evaluation, foresight, technology assessment and Delphi surveys (NIH 2014). The S3 strategy delivers reasons for an effective monitoring and evaluation mechanism as well (NIH 2014). The following three types of evaluation activities are foreseen: interim evaluation, ongoing evaluation and ex-post evaluation. According to

S3, evaluations shall be both normative and summative, following to pre-defined indicators (RIO country report Hungary 2015, 34).

The current lack of strategic intelligence is reflected by the fact that no comprehensive study was published in the field of R&I, with the following exception: the six RDI Sectoral White Books (NIH 2012) (i.e. agriculture and food processing; health industry; energy; ICT; environment protection; mobility, vehicle industry and logistics) related to the National Smart Specialisation Strategy (NIH 2014).

As for the Operational Programmes of the EU Structural Funds related to financing innovation, there are regular evaluations as prescribed in the Common Provision of Regulations P10F¹ P and other relevant regulations related to the EU Structural Funds, addressing ex-ante, ex-post, and interim evaluations.

For a long time, there has hence been no R&I policy monitoring or specific evaluation culture in Hungary. For example, a review of R&I policy concluded that Hungary has little experience in the design, implementation and evaluation of R&I strategies (EVAL-INNO, 2014, 28). This assessment is supported by the fact that the Hungarian R&I funding programmes are not systematically evaluated, apart from the mandatory evaluations of the Operational Programs co-financed by the Structural Funds. Another exception is the international evaluation of OTKA, described above, which was carried out by foreign experts based on an agreement with the European Science Foundation.

Since the establishment of the Science and Technology Observatory (called Kaleidoscope information service) within the National Innovation Office in 2012, the situation has changed slightly. (RIO country report 2015 – Hungary, p18). Meanwhile, the Kaleidoscope service provides interested stakeholders with overview data and reports. The Observatory's role has been taken over by the Department for Analysis and Information of the NRDI Office since 2015.

The Horizon 2020 Policy Support Facility offered the opportunity to evaluate the national R&I systems based on a peer-review basis, and Hungary was one of the first EU Member States that made use of this opportunity. The report of the expert groups was published in October 2015. The report identified the following four focus areas for the improvement of the Hungarian R&I system: i) R&I governance, funding and policy-making; ii) availability of human resources for R&I; iii) university-industry cooperation, technology transfer and entrepreneurship; iv) framework conditions for innovation in the business sector (European Commission, 2016).

The Peer Review 2016 (European Commission, 2016) highlighted that it was important for the R&I policy developments that programmes and support measures should be evaluated with the aim to use objective methods to assess the effectiveness and efficiency of specific measures so that policymakers and broader stakeholders can learn to improve policy, and programme managers are held to account. Despite a legal obligation to conduct programme evaluations, which was introduced in 2004, evaluation culture and the practice to ensure accountability, transparency and learning in R&I policy is poor, especially at the programme level.

The Peer Review emphasised the relevance of evaluation and formulated the following key messages for Hungary: Evaluations should be utilised by the highest level of policy-making and policy delivery. Evaluations should be based on sound evidence and delivered by external experts, be regularly planned and sufficiently budgeted. Furthermore, evaluations should be based on commonly agreed procedural standards to guarantee transparency. The review also stresses that a conducive evaluation culture and evaluation framework has to be developed. A common platform of practitioners, policymakers and policy implementers is suggested as a helpful element for community building and tool to support learning. Such a platform could also assist in linking Hungary to international communities, building peer pressure regarding good practice and thus establishing a favourable evaluation culture. The Review panel formulated its recommendations as follows (European Commission 2016, 39):

¹ COM(2013) 246 final, 2011/0276 (COD)

- “The panel supports a move towards increased evidence-based policy-making, including through the use of foresight and through the systematic evaluation of R&I policies, programmes and support measures. It calls for evaluations of the outputs and outcomes of programmes and projects to be managed in a clear and transparent way and to be delivered in a timely and efficient fashion, giving due publicity to them and eliminating undue bureaucracy”.

- “All priority R&I programmes should be rigorously evaluated at appropriate times using international reviews and standards. The outputs of those programmes should be evaluated against their objectives and funding. The systematic and meaningful international evaluation of the whole set of national R&I programmes should lead to incremental improvements of a core set of programmes that should remain stable over time to assure system predictability.”

Evaluation of gender equality in Hungarian R&I

There are no prescriptions or practices in Hungary concerning the evaluation of gender equality in R&I evaluations or the evaluation of gender equality interventions in R&I. As for the programmes financed under the EU Structural Funds, there are horizontal requirements concerning gender equality. These are evaluated in the framework of the overall evaluation exercises in the case of the Operational Programmes of the EU Structural Funds related to financing innovation (Füleki et al., 2018).

The conclusions of a workshop that was organised in the context of the EU funded Horizon 2020 project EFFORTI and that was dedicated to discuss, amongst other things, the role of gender equality in the Hungarian R&I system reiterated the main issues of evaluation culture and policy in R&I as follows. First of all, the participants concluded that there are no prescriptions and no practices in Hungary concerning the evaluation of gender equality in R&I evaluations or evaluation of gender equality initiatives in R&I. However, there is one exception: R&I programmes financed under the EU Structural Funds are evaluated with respect to gender equalities in the framework of the overall evaluation exercises.

From the point of view of the workshop participants, a strong recommendation was to translate this practice to all R&I programmes and projects in Hungary. Furthermore, the participants suggested that gender equality is incorporated in all public policy and development policy. This call for gender mainstreaming implies the assessment of the impact of measures on gender equality, also in the R&I fields. Additionally, appropriate indicators were sought out for input and output measurement (input and result indicators). Finally, when assessing the societal impacts of R&I programmes and projects, it was recommended to consider and assess also the impacts on gender equality (Füleki et al. 2018).

The conceptual and methodological approach

In the frame of the EFFORTI project and drawing on comprehensive desk research, a methodological approach was initiated to develop the conceptual evaluation framework that involved: (1) A literature review and mapping of state of the art with a wide-ranging collection of evaluation studies on gender equality intervention; (2) A collection of smart practices of interventions according to specifically defined criteria; (3) The development of a typology of gender equality interventions based on existing literature; (4) The identification and development of indicators (quantitative and qualitative for gender equality at different levels of intervention, i.e. team, organisation and country). The final conceptual evaluation framework² proposes a wide-ranging framework for capturing the complex context of interventions and their impacts. It opens the “black box” of the linkages between gender equality interventions and outputs, outcomes and impacts, pointing out which aspects of context are important for the design and evaluation of interventions.

It is an evaluation framework that goes beyond causality explanations, focusing on how the intervention contributes to achieving impact instead of attributing impact automatically to the

² For a more detailed description of the conceptual evaluation framework, see Kalpazidou Schmidt & Graversen 2020.

intervention. The framework addresses the questions (i) in which way and (ii) under which conditions an intervention causes the observed effects (Döring & Bortz 2016) and explores “not only whether the intervention works, but also how, for whom and in which context” (Van Belle, Marchal, Dubourg & Kegels 2010). This is articulated in an intervention logic model that explicitly states the assumptions of the intervention, which may help to understand the linkages between inputs, outcomes and impacts (Kalpazidou Schmidt & Graversen 2020). The framework needs to be tailored to local conditions and be designed with context-sensitivity in order to account for the local challenges in assessing the impact of gender equality interventions in R&I (Kalpazidou Schmidt & Graversen 2020).

The conceptual framework has been tested and validated through a comprehensive case study work. Nineteen case studies of gender equality interventions in R&I were carried out in Austria, Denmark, Germany, Hungary, Spain and Sweden. The case studies varied in scope (national, regional or institutional), targeted sector (Higher Education sector, Business Enterprise sector, Government sector), target beneficiary (university, company, individual researchers) and main objectives. The case study work aimed to create both a logic frame and a theory of change for each intervention that was able to factor in an assessment of the design and implementation of the interventions, and main outcomes (both in terms of gender equality and R&I) as well as taking into consideration contextual elements.

Methodological approach

The main objectives of the case study work in the Hungarian case were to: (1) carry out the concept, and implementation analysis as well as impact assessments for the selected case study; (2) develop a logic frame and a theory of change for the case; (3) validate the indicators developed in the conceptual framework useful for the case. The log-frame and the theory of change were developed based on analysis of the design, implementation and impact assessment of the case. During the case study work, explicit feedback was asked from policymakers, programme managers and evaluators in order to validate the theory-based evaluation approach.

The case study from Hungary focuses on a gender equality intervention that aims to promote networks among the female staff of a multinational company. This measure is implemented in large multinational companies operating in Hungary, which employ thousands of people, many of whom are women. Another reason for selecting the case is that the company runs not only factories and business service centres covering several branches of the industry but also research units across the country, making it a relevant actor in the Hungarian research, development and innovation sector.

We used two different approaches in relation with the case study methodology. First, we carried out a thorough analysis of the publicly available online information on the programme, as well as of internal company documents, namely organisational charts featuring the leaders of the initiative both at the global and the regional level. A list of all events organised by the network in the previous year, statistics, and status update documents were also used.

Secondly, we conducted five semi-structured interviews with Hungarian employees from different departments who played an active role in the network as programme managers, practitioners or beneficiaries. One of the interviewees was male. The employment levels of the stakeholders ranged from entry-level to top management positions. Our first interviewee was the initiator of the intervention in Hungary. This was followed by interviews with two managers of the company, one of whom was a leader and the other an active supporter of the initiative. Then we conducted an interview with a young participant of the project. Finally, we identified a person in a very high management position both at the company and within the network as the programme manager of the intervention, and we scheduled an interview with this person at a later stage.

We aimed for the most possible extensive coverage of questions regarding the history, the design, the concept analysis, the implementation and the impact assessment of the intervention. We focused on different aspects of the measure with each interviewee, according to their role in the company. We created the theory of change for the case study based on information obtained from the

first four interviews, and we presented it to the programme manager during the fifth one to have it validated.

All interviews lasted between 60 and 90 minutes and were carried out face-to-face at the company premises. In order to assure a sound empirical analysis, the interviews were recorded in audio format and transcribed. Research data from audio recordings, transcriptions and handwritten interview notes were handled, analysed and anonymised by the researchers that collected the data.

Figure 1 shows the intervention's logic frame, in which contextual factors are presented at the national, organisational and team level. While the team context shows the individual features that generally characterise women in work teams, the organisational context mirrors the norms and unspoken rules widely accepted in a given company. However, for our present paper, the national context is of paramount importance since it serves as the background of the selected intervention. The context at the macro (national) level includes policy and economic factors, e.g. the lack of gender equality policies at the national level; the relatively low participation of women in the labour market, especially in STEM (Science, Technology, Engineering, Mathematics) fields, as well as social and cultural factors, such as the prevailing traditional values in the Hungarian society regarding gender roles; the male-dominated organisational cultures with a high level of gender-based vertical segregation; the number of female STEM graduates, and the lack of work-life balance in Hungary.

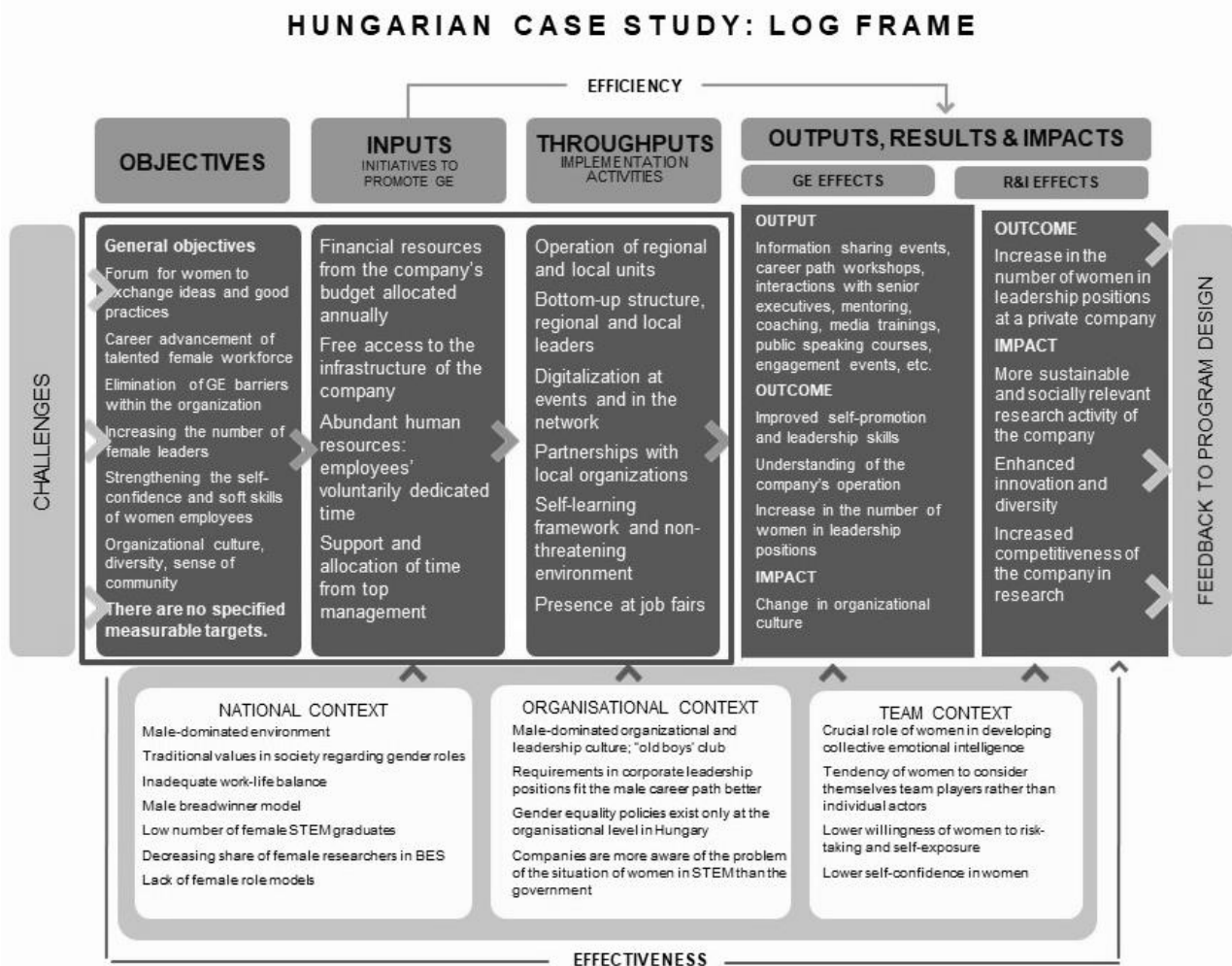


Figure 1. Logic frame for the Hungarian Case Study of a Network Programme in a larger company

Source: Authors.

Results

The multinational company has a global gender diversity initiative, which aims to aid the company in attracting, retaining, and developing the best talent in every country and every business segment it operates in. The programme began as a top-down initiative, which came from the recognition that the issue of the disproportionately low female representation in the management of the company needed to be addressed. To this end, a network of female staff was established to help the group members develop the attributes they need to advance their careers within the organisation.

A few years later, a local executive was asked to launch the initiative in the Hungarian branch of the company. The project was set in motion after a period of six months dedicated to studying existing networks and good practices in the private sector, which led to the development of a highly innovative, vibrant local group in the next several years. Since there is no membership fee or official membership application, it is difficult to estimate the number of members. At the time of the interviews, approximately 50 female employees had an active role in shaping the programme and organising the events of the local network in Hungary.

The Hungarian unit is somewhat different from the ones operating in other countries. The reason for this is the country's central geographical position. The company's presence in Hungary is influential in the region, which means that many activities are concentrated in a relatively limited area. Consequently, the cohort of employees is quite colourful in Hungary. This kind of diversity and the geographical closeness make the activities and the community of the local network extremely varied and vibrant.

The initiative's primary target group is ambitious female employees within the company who wish to advance their careers and become business leaders. A second target group includes all women working in the company and the future potential female workforce interested in joining the company. The initiative offers primarily mentoring, coaching and networking opportunities to stimulate the professional advancement of women employees. Regular trainings help to enhance the performance and the self-branding of the female workforce. The network also provides a self-learning framework and a women-friendly environment, enabling female employees to improve their professional and leadership skills.

The network initiative fits the overall priorities and strategies of the company and enjoys support from senior executives and top management. As the network has its own leadership and governance both at the global and regional/local levels, it is a highly autonomous body within the organisation, functioning voluntarily. High-potential female executives are often tapped to manage the units of the network, giving the women employees valuable, enterprise-wide executive experiences and developmental opportunities. The company allocates financial resources, time and infrastructure for the purposes of the network, but the active involvement of the members constitute the most valuable asset of the intervention.

The operation of the initiative is not overregulated, in spite of the fact that it stays within a well-defined frame. Organising the events means the bulk of the implementation process of the measure, which is usually carried out by the most active Hungarian members mentioned above. Each task is assigned to a volunteer member of the group, who bears the responsibility of organising the event in question.

The observed outputs of the intervention include an immense variety of activities and events that have been organised and completed since the creation of the Hungarian segment of the network, with hundreds of events reaching out to more than 30 thousand employees, both male and female. In general, the activities include both regular and non-recurrent events and programmes. Besides the networking and information sharing events, the specific outputs comprise career path workshops, interactions with senior executives, training programs, mentoring sessions, coaching activities, media training, public speaking courses, engagement events, specific services, such as babysitting arrangements at events, etc. The vast majority of the events organised by the network are open to anyone who works for the company.

The main observed outcome of the intervention is the professional development and career advancement of women employees who take part in the network. According to the interviewees, the motto of the initiative is: 'the more you put into it, the more you get out of it'. In a few cases, participation in the network even directly resulted in promotions, but in general, the intervention has a much more indirect positive effect on women's careers.

The initiative enhances the visibility of female workers within the company. Meetings and mentoring sessions with senior executives ensure that the skills and achievements of women are noticed by the management. In addition, contact with colleagues and leaders help female employees to gain an insight into the operation of the company and a better understanding of its strategies and procedures.

The network creates a self-organised, self-dependent and supportive community for female workers without excluding male employees. It increases the members' sense of belonging, thus boosting their engagement with the company. According to the programme manager, the network also stimulates unplanned and even overlooked learning processes within the organisation. What is learned in the women-friendly environment of the network initiative can be utilised routinely and almost automatically in the everyday work. This is an unintended positive effect of the network's operation.

One of the intended longer-term impacts of the measure is to improve the status of women in the company. In addition, women are being better trained in dealing with the challenges they face in the workplace. In the very end, such an effect is not only beneficial for the female program participants; it has positive impacts on the entire organisation because it brings about significant changes in the organisational culture and creates an environment in which all employees can thrive. These gradual but perceptible changes in the organisational culture and the operation of the company constitute the main indirect impact of the activities of the network. Additionally, the network may function as a best practice example for other companies in the private sector.

The major R&I-related outcome of the intervention is the enhanced number of women in decision making and leadership positions in a company that operates within the STEM fields, which is expected to boost the scientific performance of the given company. Impacts of the measure that are linked to R&I include a more sustainable and socially relevant research activity carried out by the company, an enhanced innovation driven by the diversity of the company and – in the long run - the increased scientific and economic competitiveness of the company.

By broadening knowledge and abilities, improving decision-making skills, increasing self-esteem and motivation, maximising career opportunities and improving employment marketability, the intervention encourages female employees to perform in an enhanced spirit of responsibility, which is expected to be reflected in the overall performance of the company. Moreover, increased awareness of gender issues contributes to better integration of women in the workplace, which may further enhance the company's overall performance and outputs.

Summing up, the long-term effects of the intervention might be improved visibility and leadership skills of the participants, their better understanding of the strategies and operations of the company, increased numbers of women in decision-making and leadership positions, changes in the organisational culture, and the enhanced innovativeness, competitiveness and research performance of the company. Gender-smart business innovation not only leads to the introduction of new features in the product and service design, but it also contributes to the emergence of new markets and the profitability of existing ones since a great number of products and services are far from gender-neutral. Therefore, boosting its innovation capacities by promoting gender equality helps the company thrive in the age of accelerated technological change.

Discussions and Conclusion

The herewith presented case study work and results showed that the evaluation framework and, in particular, the logic frame developed for the case has proved to be a valuable tool to consider

how different factors may contribute to the impact of interventions. The logic frame was valuable in identifying outcomes and impacts of gender equality interventions and possible R&I effects.

A general outcome arising from the case study work is that different activities contributed to the outcomes and impact of the intervention in combination with a complex array of contributory contextual factors, such as human and financial resources that can facilitate or hinder implementation and thus impact (Kalpazidou Schmidt & Graversen 2020, Palmén & Kalpazidou Schmidt 2019). The case studies provide an essential insight into how strongly national evaluation cultures shape the intensity and quality of programme evaluations: in countries with a robust evaluation culture, such as Austria, Germany, Denmark and Sweden, programmes will be evaluated more comprehensively and more frequently than, for example, in Hungary, where there is a weaker evaluation culture (Reidl et al. 2018).

Some limitations in connection with the study of the Hungarian case were identified involving the following: As the case study is a voluntary initiative of a large private enterprise, the evaluation of the initiative might only be carried out at the company's own discretion. Furthermore, the company's emphasis is on the continuous implementation of the intervention and not on its evaluation. This was reflected in the fact that there were no measurable targets and verifiable objectives stated in the internal policy documents. In addition, due to the ongoing character of the programme, an ex-post evaluation was not feasible either. However, the logic frame developed for the case has been a very useful tool and helped identify the main actors involved in the intervention, map the input and the contextual factors - and thus the facilitating and hindering factors - in the implementation process, but also the key outcomes and impacts, not only as regards gender equality, but also in terms of possible R&I effects.

In conclusion, the article discussed how evaluation cultures vary greatly across Europe with well-established evaluation cultures and systematic evaluations of programmes and institutions in western European countries, while in Central and Eastern European countries, in contrast, efforts continue to establish evaluation practices and to further develop existing evaluation cultures. We have focused on Hungary, where an evaluation culture within R&I is under development, also considering evaluations of gender equality in R&I. Based on the results of a Horizon 2020 project, we have applied a context-sensitive evaluation concept in a case study. There was an intervention in a branch of an international company operating in Hungary. The evaluation concept enables programme owners and evaluators to develop a tailor-made design and impact model for their planned or ongoing gender equality interventions. The case study work and the analysis showed hence that the used evaluation approach is applicable also in countries with a certain catch-up demand of the existing overall evaluation culture. The approach used did make it possible not only to depict effects on gender equality itself but also to anticipate possible effects on R&I, which makes it particularly interesting for private companies.

References

1. Arnold, Erik (2004): Evaluating research and innovation policy: a systems world needs systems evaluations. In: *Research Evaluation* (13), 3–17.
2. Bachtler John (2012): The evaluation of regional policy in Europe: culture, commitment and capacity. In Olejniczak Karol; Bienias Stanislaw & Kozak Marek (eds): *Evaluating the effects of regional interventions. A look beyond current Structural Funds' practice*. Ministry of Regional Development, Warsaw, p. 94-117.
3. Borrás, Susana & Laatsit, Mats (2019): Towards system oriented innovation policy evaluation? Evidence from EU28 member states. In: *Research Policy* 48 (2019) 312–321.
4. Dahler-Larsen, Peter (2012): *The Evaluation Society*. Stanford University Press, Stanford, CA.
5. Döring, Nicola, & Bortz, Jürgen (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer.
6. Edler, Jakob; Berger, Martin; Dinges, Michael; & Gök, Abdullah (2012): The practice of evaluation in innovation policy in Europe. In: *Research Evaluation* 21 (2012) pp. 167–182. doi:10.1093/reseval/rvs014.
7. Edler, Jakob; Joly, Pierre-Benoit; Kuhlmann, Stefan; Nedeva, Maria, Propp, Tilo; Rip, Arie; Ruhland, Sascha; & Duncan, Thomas (2006): *Understanding "Fora of Strategic Intelligence for Research and Innovation"*. The PRIME Forum Research Project. Final Report.

8. Edler, Jakob et al. (2010): INNO-Appraisal. Understanding Evaluation of Innovation Policy in Europe, Brussels/Manchester.
9. Edler, Jakob; Ebersberger, Bernd; & Lo, Vivien (2008): Improving policy understanding by means of secondary evaluation. *Research and Development Evaluation* (17), 175–186.
10. European Commission (2016): Peer Review of the Hungarian Research and Innovation system. Directorate-General for Research and Innovation. Brussels, 17Thttps://publications.europa.eu/en/publication-detail/-/publication/79eb4374-7ef4-11e6-b076-01aa75ed71a1/language-en17T.
11. European Science Foundation (ESF) (2014) Organisational evaluation of the Hungarian Scientific Research Fund (OTKA). Evaluation Report. European Science Foundation, Brussels – Budapest. 17Thttp://www.esf.org/uploads/media/otka_evaluation_01.pdf
12. EVAL-INNO (2014) Developing RTDI evaluation culture in South East Europe. Fostering evaluation competencies in research, technology and innovation – the EVAL-INNO experience. LIT Verlag, Vienna. 17Thttps://wbc-rti.info/object/document/14453/attach/EVAL-INNO_Final_Publication_WEB.pdf.
13. Fahrenkrog, Georg et al. (2002): RTD Evaluation Toolbox – Assessing the Socio-economic Impact of RTD Policies (EUR 20382 EN). Seville: IPTS.
14. Feller, Irwin (2007): Mapping the frontiers of evaluation of public-sector R&D programs. In: *Science and Public Policy* (34), 681–690.
15. Füleki, Katalin; Groó, Dóra; Kleinheincz, Ferenc; & Paksi, Veronika (2018): Country Note Hungary EFFORTI – Deliverable 2.2 (<https://efforti.eu/efforti/publication/country-note-hungary>).
16. Hage, Jerald; Jordan, Gretchen, B.; & Mote, Jonathan (2007): A theory-based innovation systems framework for evaluating diverse portfolios of research, part two: macro indicators and policy interventions. In: *Science and Public Policy* (34), 731–741.
17. Jacob, Steve; Speer, Sandra & Furubo, Jan-Eric (2015): The institutionalisation of evaluation matters: Updating the international atlas of evaluation 10 years later. *Evaluation* 21(1): 6–31.
18. Jordan, Gretchen B.; Hage, Jerald; & Mote, Jonathan (2008): A theories-based systemic framework for evaluating diverse portfolios of scientific work, part 1: micro and meso indicators. *New Direction of Evaluation*. 2008, 7–24.
19. Kalpazidou Schmidt, Evanthia & Graversen, Ebbe Krog (2020) Developing an evaluation framework for gender equality in R&I – a systematic non-linear approach for the integration of team, organisational and structural context factors in policy design and evaluation. In: *Evaluation and Program Planning* 79.
20. Kuhlmann, Stefan (2001): Evaluation as a Source of “Strategic Intelligence”. In: Philip Shapira and Stefan Kuhlmann (Editors): *Learning from Science and Technology Policy Evaluation*. Proceedings from the 2000 US-EU Workshop on Learning from Science and Technology Policy Evaluation, Bad Herrenalb, Germany. Pp. 11-21-11-59.
21. Kuhlmann, Stefan (2003): Evaluation of research and innovation policies: a discussion of trends with examples from Germany. In: *Int. J. Technology Management*, Vol. 26, Nos. 2/3/4, pp. 131-149.
22. Kuhlmann, Stefan; Boekholt, Patries; Georghiou, Luke; Guy, Ken; Héraud, Jean-Alain; Laredo, Philippe; Lemola, Tarmo; Loveridge, Denis; Luukkonen, Terttu; Polt, Wolfgang; Rip, Arie; Sanz-Menendez, Luis; Smits, Ruud (1999): Improving Distributed Intelligence in Complex Innovation Sys-tems. Final report of the Advanced Science & Technology Policy Planning Network (ASTPP), a Thematic Network of the European Targeted Socio-Economic Re-search Programme (TSER), Brussels/Luxembourg 1999. Munich Personal RePEc Archive.
23. Smits, Ruud; Kuhlmann, Stefan; & Shapira, Phil (Eds.) (2010): *The Theory and Practice of Innovation Policy*. An International Research Handbook. Edward Elgar, Cheltenham
24. Magro, Eburne & Wilson, James R. (2013): Complex innovation policy systems: towards an evaluation mix. In: *Research Policy* (42), 1647–1656.
25. Miles, Ian & Malik, Khaleel (2005): *SMART Innovation: A Practical Guide to Evaluating Innovation Programmes*, A Study for DG Enterprise and Industry, Brussels.
26. Molas-Gallart, Jordi & Davies, Andrew (2006): Toward theory-led evaluation: the experience of European science, technology, and innovation policies. In: *American Journal of Evaluation* (27), 64–82.
27. NIH (2012): Sectoral Strategic Whitebooks (ÁSFK), 17Thttp://nkfih.gov.hu/sectoral-strategic
28. NIH (2014): National Smart Specialization Strategy, 2014, 17Thttp://nkfih.gov.hu/szakpolitika-strategia/national-smart
29. Nutley, Sandra; Huw, Davies; Walter, Isabell (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.
30. OECD (1998): Best Practice Guidelines for Evaluation, PUMA Policy Brief No. 5, May 1998.
31. OECD (2008): OECD reviews of innovation policy: Hungary. OECD, Paris. 17Thttp://www.oecd-ilibrary.org/science-and-technology/oecd-reviews-of-innovation-policy-hungary-2008_9789264054059-en
32. OECD (2009): Enhancing Public Research Performance through Evaluation, Impact Assessment and Priority Setting, OECD. DSTI/STPTIP (2009)5, Paris.
33. Palmén, Rachel & Kalpazidou Schmidt, Evanthia (2019). Analysing facilitating and hindering factors for implementing gender equality interventions in R&I: Structures and processes. *Evaluation and Program Planning*, vol. 77.

34. Pattyn Valerie; Van Voorst Stijn; Mastenbroek Ellen & Dunlop Claire A. (2017): Policy evaluation in Europe. In Ongaro, Edoardo & Van Thiel Sandra (Eds): The Palgrave Handbook of Public Administration and Public Management. London, Palgrave Macmillan.
35. Plantenga, Janneke (2014): Searching for welfare, work and gender equality. Retrieved from http://www.foreurope.eu/fileadmin/documents/pdf/Workingpapers/WWWforEurope_WPS_no059_MS206.pdf
36. Polverari Laura (2015): The Monitoring and Evaluation of the 2014-20 Cohesion Policy Programmes. IQ-Net Thematic Paper 36(2), European Policies Research Centre, University of Strathclyde.
37. Reidl, Sybille; Holzinger, Florian; Streicher, Jürgen; Beranek, Sarah; Unger, Maximilian; Hafellner, Silvia (2018): EFFORTI Deliverable 2.3. Comparative Background Report (<https://efforti.eu/sites/default/files/2019-10/EFFORTI%20D2.3%20Comparative%20Background%20Report%20Context%20Analysis%20Final.pdf>).
38. Ruegg, Rosalie; Feller, Irwin (2003): A Toolkit for Evaluating Public R&D Investment, Models, Methods, and Findings, from ATP's First Decade. Gaithersburg: NST.
39. Sanderson, Ian (2002): Evaluation, Policy Learning, and Evidence Based Policy Making, Public Administration 80(1), 1-22.
40. Solesbury, William (2001): Evidence Based Policy: Whence it Came and Where it's Going. ESRC UK Centre for Evidence Based Policy and Practice. Working Paper 1. London.
41. Van Belle, Sara B; Marchal, Bruno; Dubourg Dominique & Kegels, Guy. (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.
42. Vedung Evert (2010): Four Waves of Evaluation Diffusion. Evaluation, pp. 263-277.

Susanne Buehrer, Evanthia Kalpazidou Schmidt, Dorottya Rigler, Rachel Palmen

Kaip įgyvendinti kontekstui jautrius vertinimo metodus šalyse, kuriose dar tik formuojasi vertinimo kultūra

Anotacija

Vertinimo kultūros ir vertinimo gebėjimų stiprinimas Europos Sąjungoje labai skiriasi. Vakarų Europos šalys, tokios kaip Austrija, Vokietija, Danija ir Švedija, buvo vadinamos pirmaujančiomis šalimis, nes jos sukūrė nusistovėjusią vertinimo kultūrą ir sistemingai vertina programas ir institucijas. Priešingai, Vidurio ir Rytų Europos šalyse toliau stengiamasi nustatyti vertinimo praktiką ir toliau plėtoti dabartinę vertinimo kultūrą. Pavyzdžiui, Vengrijoje nėra nusistovėjusios mokslinių tyrimų ir inovacijų vertinimo praktikos. Niekur nėra atsižvelgiama į lyčių lygybės klausimus atliekant mokslinių tyrimus ar kuriant ir diegiant inovacijas. Išimtį sudaro tik projektai finansuojami iš ES struktūrinių fondų. Remdamiesi projekto „Horizontas 2020“ rezultatais, Vengrijoje buvo atliktas tyrimas pritaikant jautrius nuo konteksto priklausančius tyrimų metodus. Šis metodas leido programas įgyvendinti ir vertinti asmenims sukurti individualų jų planuojamų ar vykdomų lyčių lygybės intervencijų dizaino ir poveikio modelį. Šio vertinimo rengimas buvo pagrįstas nuodugnia literatūros analize ir atlikta 19 atvejų analize, pastaroji buvo pagrįsta dokumentine analize ir pusiau struktūrizuotais interviu. Straipsnyje argumentuotai pagrindžiama, kad kontekstui jautrus vertinimo mechanizmas gali būti taikomas šalims, kuriose kritinio vertinimo kultūra dar tik plėtojasi. Ypatingas pateikto vertinimo metodo bruožas yra tas, kad vertinimas yra jautrus kontekstui. Kita vertus, šis požiūris leidžia ne tik pavaizduoti poveikį pačiai lyčių lygybei, bet ir numatyti poveikį moksliniams tyrimams ir inovacijoms. Toks poveikis gali būti, pavyzdžiui, stipresnis mokslinių tyrimų orientavimas į visuomenės poreikius, todėl šio tyrimo išvados ypač vertingos privačioms įmonėms.

Susanne Buehrer, Project Manager, Fraunhofer Institute for Systems and Innovation Research, Germany

E-mail: susanne.buehrer@isi.fraunhofer.de

Evanthia Kalpazidou Schmidt, Associate Professor and Research Director, Aarhus University, Denmark

E-mail: eks@ps.au.dk

Dorottya Rigler, Researcher, Association of Hungarian Woman in Science, Hungary

E-mail: dorottya_rigler@yahoo.com

Rachel Palmen, Senior researcher, Universitat Oberta de Catalunya, Spain

E-mail: rpalmen@uoc.edu

Susanne Buehrer, Projectų vadovė, Fraunhofer sistemų ir inovacijų tyrimų institutas, Vokietija

El. paštas: susanne.buehrer@isi.fraunhofer.de

Evanthia Kalpazidou Schmidt, Docentė ir tyrimų vadovė, Aarhus universitetas, Danija

El. paštas: eks@ps.au.dk

Dorottya Rigler, Tyrėja, Vengrijos moterų moksle asociacija, Vengrija

El. paštas: dorottya_rigler@yahoo.com

Rachel Palmen, Vyriausioji tyrėja, Oberta de Catalunya universitetas, Ispanija

El. paštas: rpalmen@uoc.edu



This article is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 (CC BY 4.0) License (<http://creativecommons.org/licenses/by/4.0/>).