GENDER ACTION+

Project acronym:

GENDERACTIONplus

Project title:

Gender Equality Network to Develop ERA Communities To coordinate Inclusive and sustainable policy implementation

Grant agreement No: 101058093

Start date of project: 1 June 2022

Duration: 36 months

Deliverable 6.3
GEP impact evaluation system

Due date of the deliverable	30.11.2024
Submission date	27.11.2024
File name	D6.3 GEP impact evaluation system
Organisation responsible for the deliverable	Higher Education Authority (HEA)
Author(s)	Jennie Rothwell, Angela Wroblewski, Julie Anne Young
Status	Final
Dissemination level	PU

DOCUMENT REVISION HISTORY

Revision History			
Version	Date	Modified by	Comments
1.0	19/04/2024	Jennie Rothwell, Julie Anne Young	First draft
2.0	27/05/2024	Marcela Linková	Comments on the first draft
2.1	28/05/2024	Jennie Rothwell, Julie Anne Young,	Implementation of the comments, including additional information
2.2	05/06/2024	Jörg Müller, Chloe Mour, Policy CoP	Comments by expert reviewers and Policy CoP
3.0	02/07/2024	Julie Anne Young, Jennie Rothwell, Angela Wroblewski	Implementation of comments. addition of Angela Wroblewski as co-author
4.0	12/11/2024	Ella Ghosh	Comments by Quality Editor
5.0	13/11/2024	Julie Anne Young, Jennie Rothwell, Angela Wroblewski	Implementation of the comments, including additional information
6.0	14/11/2024	Julie Anne Young, Jennie Rothwell	Version 6 shared with the Consortium and expert reviewers
6.0	25/11/2024	Julie Anne Young, Jennie Rothwell	Implementation of the comments
7.0	26/11/2024	Julie Anne Young, Jennie Rothwell	Document sent to Coordinator
7.0	27/11/2024	Martina Fucimanová	Document submission

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AC	Associated Countries
CoP	Community of Practice
EC	European Commission
EGET	European Gender Equality Taskforce
EIGE	European Institute for Gender Equality
ERA	European Research Area
EU	European Union
GE	Gender equality
GEP	Gender Equality Plan
GBV	Gender-based violence
GDPR	General Data Protection Regulation
HEA	Higher Education Authority (Ireland)
HEI	Higher Education Institution
KIF	Committee for Gender Balance and Diversity in Research (Norway)
MS	EU Member States
NA	National Authorities
NCP	National Contact Point
NIP	National Impact Plan
RFO	Research Funding Organisation
R&I	Research & Innovation
RPO	Research Performing Organisation
SWG GRI	Standing Working Group on Gender in Research and Innovation
WP	Work Package

EXECUTIVE SUMMARY

This document outlines guidance on establishing an evaluation framework for gender equality plan (GEP) implementation. This guidance is aimed at national authorities (NA) but also considers the role and responsibility of the European Commission (EC) and other important actors including research funding organisations (RFOs) and research performing organisations (RPOs) in ensuring that GEP implementation is a successful and sustainable endeavour.[1]

Evaluation at national level is key to understanding the impact of GEPs as an instrument for institutional change. Effective evaluation ensures that the relevancy, effectiveness, sustainability and impact of GEP implementation are measured and analysed.

As defined by the EC, a GEP is a "set of commitments and actions that aim to promote gender equality in an organisation through institutional and cultural change" (European Commission, 2021: 11). Developing and maintaining a GEP is a necessary part of the eligibility criteria for Horizon Europe funding. Evaluation of GEPs should, therefore, form an important part of member states (MS) and associated countries (AC) national policy framework relating to gender equality in research and innovation (R&I) (European Commission, 2020). The benchmarking analysis of monitoring and evaluation of GEPs carried out by GENDERACTIONplus (Knapińska and Chrobak-Tatara, 2023) provides the main input and background for this guidance document, in addition to definitions and requirements from the EC relating to GEPs. Only a few countries have evaluation policies in place and not necessarily specific to GEP implementation. This highlights a need for EC level guidance to promote evaluation.

In D6.2 (Guidance on GEP monitoring), it was noted that GEP implementation varies across MS. While developing an evaluation framework may only be immediately relevant for the countries that have strong monitoring in place, ideally national authorities would commit to monitoring and evaluation concurrently. The guidance provided in this document offers adaptable advice on what is important and necessary when developing a framework for evaluating the implementation of GEPs. This adaptability ensures NA, in cooperation with the EC, can take into account their own context-specific requirements and provide impactful measures.

---

[1] We use the terms national and national authorities throughout, but this guidance also addresses regional authorities and other bodies that may have responsibility for advancing gender equality and monitoring and evaluating GEPs. It is also recognised that there are differing national structures for R&I and gender equality across MS, and that this guidance is intended to be adapted to different national contexts.

1. INTRODUCTION

1.1 About the project

Building on the Horizon 2020 project GENDERACTION, the overall goal of GENDERACTIONplus is to contribute to the coordination of the gender equality and inclusiveness objectives of the new European Research Area (ERA) through the development of two communities of practice (CoPs), one consisting of representatives of NA and the second consisting of representatives of RFOs. The network is made up of a total of 22 member states (MS) and 3 associated countries (AC), as well as 26 project partners and 14 associated partners.

Adding the plus sign to the title of the previous GENDERACTION project not only indicates that it is a follow-up project but also makes it explicit that this project also addresses diversity and intersectionality (the gender+ approach).

Specifically, the GENDERACTIONplus project aims to:

- Develop strategic policy advice on existing and emerging policy solutions.
- Enhance the policy-making process by engaging with stakeholders, civil society organisations, and citizens.
- Build capacities, competence, and expertise for gender equality and mainstreaming in R&I among the policy and RFO community members, with special attention to countries with a less comprehensive policy.
- Create an impact through communication, dissemination, and exploitation.

Thematically, the project focuses on:

- Intersectionality and inclusiveness.
- Gender-based violence (GBV).
- Gender dimension in research and innovation.
- Monitoring and evaluating gender equality actions in ERA.
- Promoting institutional change through GEPs.

GENDERACTIONplus aims to achieve the following impacts:

- Advance policy coordination among MS and AC and through stakeholder and citizen engagement.
- Improve research careers and working conditions in European R&I, by developing policy dialogue and solutions on inclusion and intersectionality, combating GBV, and promoting institutional changes through GEPs.
- Improve research quality and the social responsibility of knowledge by integrating the gender dimension into R&I.
- Reduce geographic inequality by targeting less experienced/engaged countries and regions.

1.2 Structure of the guidance document

This document outlines guidance for national authorities and the EC on the development of an evaluation framework for GEPs at national level. It is divided into the following main parts:

- Firstly, the policy context is summarised with reference to the importance of the Pact on R&I in Europe, the four interlinked outcome deliverables of Action 5 of the ERA Policy Agenda, the Ljubljana Declaration on Gender Equality in Research and Innovation 2021 and the requirements for Horizon Europe funding. The ERA Policy Agenda 2025-2027 was under negotiation at the time of developing this guidance.
- Secondly, the benchmarking analysis of monitoring/evaluation of GEPs carried out by GENDERACTIONplus provides the context of the evidence and analysis (Knapińska and Chrobak-Tatara, 2023).
- Thirdly, the evaluation framework and the definition of key concepts are outlined. Monitoring and evaluation, their role in a policy cycle and challenges to their implementation in the context of gender equality is discussed.
- Fourthly, the principles for evaluation outline the attributes for successful evaluation.
- Next, the role of key actors in the evaluation framework is considered. These actors include the EC, NA, RFOs and RPOs.
- Lastly, evaluation guidance is proposed, aimed at national authorities and the EC. The final conclusions outline the importance of having strong evaluation mechanisms in place in order to effectively assess the effectiveness of GEP implementation at national level.

2. POLICY CONTEXT

Endorsing the Ljubljana Declaration on Gender Equality in Research and Innovation 2021 affirms the EU's commitment to advancing gender equality in the ERA (Council of the European Union, 2021b).

The high-level priority areas put forward by the Declaration include:

- Ensure fair, open, inclusive and gender equal career paths in research, and consider intersectional perspectives on gender inequalities.

- Facilitate mutual learning opportunities through form-follows-function robust governance.

- Employ existing and newly developed tools, such as GEPs, to facilitate systemic institutional change and remove institutional barriers.

- Address and counteract GBV; including root causes and tackling obsolete power dynamics within research and academia.

- Support active and participative monitoring and evaluation to ensure continuous improvement.

- Leverage synergies to enhance gender equality achievements within the ERA, but also within complementary fields such as the European Higher Education Area, structural funds, innovation, as well as in international cooperation.

These priorities position GEPs as a key instrument of institutional change. Additionally, the declaration acknowledges the need to:

> *develop tools for the monitoring and evaluation of national gender equality policies in R&I and gender equality actions, including Gender Equality Plans. Such tools are essential to evaluate actions aimed at achieving institutional change, collect information on Gender Equality Plans uptake rates at national level, and consider intersectional data* (Council of the European Union, 2021a).

This specific reference to monitoring and evaluation underlines its importance within the policy implementation framework, rather than as an add-on.

Similarly, the Pact on R&I in Europe underlines priorities for joint action in support of the ERA, including monitoring and evaluating national gender equality policies and plans in R&I (Council of the European Union, 2021a). In alignment with the Ljubljana Declaration, the Pact on R&I in Europe underlines priorities for joint action in support of the ERA and acknowledges and highlights the commitment to monitoring and evaluation in relation to gender equality.

Under the heading of gender equality, equal opportunities for all and inclusiveness, the Pact outlines commitments to:

- Ensure fair, open, inclusive and gender-equal career paths in research to facilitate systemic institutional and structural change in R&I funding and performing organisations.

- Counteract GBV and sexual harassment.

- Remove inequities regardless of gender, racial or ethnic origin, religion or belief, disability, age or sexual orientation.

- Monitor and evaluate national gender equality policies and plans in R&I.

Following on from this, the four interlinked outcome deliverables of Action 5 of the ERA Policy Agenda 2022-2024 reflect the key priorities of the Pact:

- Develop a policy coordination mechanism to support all aspects of gender equality through inclusive GEPs and policies, and a dedicated EU network on their implementation.

- Strategy to counteract GBV including sexual harassment in the European R&I system and to assure gender equal and inclusive working environments through institutional change in any research funding or performing organisation.

- A policy approach to inclusive gender equality that addresses gender mainstreaming and opening to intersectionality with other diversity dimensions to advance the new ERA.

- Develop principles for the integration and evaluation of the gender dimension in R&I content in cooperation with national RFOs (European Commission, 2021: 8).

- These actions and outcomes outline the commitments made by the EU to support gender equality initiatives in R&I, in particular supporting monitoring and evaluation. In turn, this report aims to provide an adaptable framework for use by national authorities to evaluate the impact of implementation of GEP actions.

The ERA Policy Agenda 2025-2027 was under negotiation at the time of developing this guidance document. While this document is aimed specifically at national authorities, it should be noted that for public bodies to be eligible to receive Horizon Europe funding, a GEP or an equivalent strategy is required. The EC has established the mandatory requirements and recommended thematic areas for GEPs (European Commission, 2021).

Table 1: Horizon Europe GEP requirements and thematic areas

Mandatory requirements	Thematic areas
Be a public document	Work-life balance and organisational culture
Have dedicated resources	Gender balance in leadership and decision making
Include arrangements for data collection and monitoring	Gender equality in recruitment and career progression
Be supported by training and capacity building	Integration of the gender dimension into research and teaching content
	Measures against GBV including sexual harassment

Source : European Commission, 2022a

As such, it is expected that most institutional GEPs will address the areas above and be collecting related data and information to support GEP implementation.

National authorities should consider the incorporation of these requirements and recommendations in the development of an evaluation framework. Additionally, if Horizon Europe requirements are taken as the minimum standard for GEPs, there is increased opportunity to benchmark across the ERA.

3. EVIDENCE AND ANALYSIS

GENDERACTIONplus undertook a benchmarking survey amongst members of the consortium. The benchmarking analysis conducted within the framework of WP 6 depicts the current landscape for GEP requirements in R&I at national/regional level (Knapińska and Chrobak-Tatara, 2023). It provides an overview of the overall policy framework on GEP requirements concerning ERA stakeholders and presents relevant policy developments, mainly targeting national authorities. Responses to the questions relating to GEP monitoring and evaluation were received from 16 national/regional authorities from 15 countries. The benchmark report relating to GEPs highlighted a number of findings:

- Sweden was the only country to indicate having a specific national evaluation system, however this encompasses national priorities relating to gender equality more broadly, rather than solely focusing on GEP implementation. This highlights the need for guidance to be developed, especially for the MS/AC where GEPs are not an established practice. It is also noted that countries outside the GENDERACTIONplus consortium or that did not respond to the survey may be advancing an evaluation system at national level.

- Seven other countries that responded to the survey (Austria, Croatia, Czechia, Ireland, Israel, Poland, and Spain) indicated that there are imminent plans to develop national level evaluation systems.

- The GEP eligibility criterion for Horizon Europe funding has led to an increase in the number of GEPs and related activities such as capacity building. This demonstrates the appetite for continued learning and development in relation to GEP implementation, even in MS where GEPs are not a national requirement.

- Incorporating Horizon Europe guidelines as mandatory and necessary baselines may form an effective part of evaluation guidance.

- Comparison of GEP monitoring and evaluation across MS/AC is challenging as there is not sufficient data or examples of good practice.

- Currently, it is difficult to have a unified model or guidance across MS/AC. A model that is flexible or has multiple levels may be required.

- Some MS/AC activities are grounded in laws/legislation, while others are based on policy. In some MS, RFOs played a role in the implementation of national GEP requirements and encouraged RPOs to adopt GEPs. There is a need for continued reflection and consideration of the role of RFOs in GEP implementation, monitoring and evaluation.

- There is a need to move beyond the minimum of the Horizon Europe GEP compliance check in order to highlight the links between different areas e.g. inclusive gender analysis in research content and to move towards intersectional approaches to inclusive gender equality work. Moving beyond the minimum levels of compliance underlines the importance of driving change through more ambitious monitoring and evaluation frameworks.

- There is a gap between requirements for public and private institutions in relation to their obligations to develop a GEP, and differences in requirements for different sectors (higher education, government, non-governmental organisations, etc).

- In some countries, institutions are encouraged to implement gender equality activities, but this is often not a requirement. With the introduction of the GEP requirement in Horizon Europe, there is a need to move beyond encouragement and voluntary activities, which are often undertaken on top of existing work without adjustment of pay/workload.

Other areas of the GENDERACTIONplus benchmark analysis also highlighted that:

- There is a lack of unified understanding around intersectionality and related terminology and concepts which may impact the design of inclusive and intersectional GEPs (Holt Zachariassen, Ghosh and Woods, 2023). As such, this may also impact the evaluation of intersectional actions included in GEPs. The EC notes that "inclusive gender equality plans and policies will only be effective if they are developed, implemented and evaluated with the trust, support and partnership of marginalised groups" (European Commission, 2022a: 44).
- Similarly, there is a lack of knowledge and understanding on how to integrate the inclusive sex/gender analysis in research and innovation content (Korsvik, Gonzalez and Dvorackova, 2023: 54). Taking this into account, this may also impact the evaluation of actions to incorporate the inclusive gender analysis in R&I content.
- While it has long been an important mission, the lack of change in relation to gender-based violence (GBV) in ERA demonstrates that actions to address GBV should also be considered in GEP monitoring and evaluation (Bondestam, Lundqvist and Young Håkansson, 2023: 61).
- As there is no requirement to have a national action plan (as in previous ERA period), there is only limited information available regarding MS/AC priorities, objectives and planned or implemented measures (Wroblewski, 2023: 35).

As noted by the Standing Working Group on Gender in Research and Innovation (SWG GRI) in 2021 a lack of requirements relating to GEPs, however, is not indicative of a lack of promising practices, noting that "progress has been achieved through softer measures or more bottom-up approaches, which may be related to differences between countries and the socio-cultural factors that affect gender equality policy design" (ERAC, 2021: 4). As such, some national authorities may wish to evaluate other institutional strategies such as equality statements or policies, or pilot initiatives.

Referencing the work of the EFFORTI project, the EC notes that "evaluating the impact of interventions in the context of multifaceted structural drivers of inequality has been a persistent challenge for gender equality over the much longer period in which action has been taken" (European Commission, 2022a: 11). With this challenge in mind, this document aims to provide practical guidance to assist in establishing a framework for a national level evaluation of GEP implementation.

3.1 Country examples

While no respondents indicated having an evaluation system that solely focuses on GEPs, there are examples of monitoring and evaluation practices that consider gender equality policy implementation more broadly.

While there are examples of good practices in relation to monitoring work being done, there is a lack of consistency at national level, especially in relation to approaches to evaluation.

Table 2: Country evaluation examples

Sweden	All public sector bodies in Sweden report annually to the Gender Equality Agency (Jämställdhetsmyndigheten) on gender mainstreaming in their organisations. These reports make reference to national gender equality goals and ask organisations to reflect on their own progress based on the previous three years.

Ireland	In 2016, the Higher Education Authority (HEA) conducted the first National Gender Equality Review of Irish Higher Education Institutions (HEI). A Second Review of Gender Equality in Irish HEI was conducted over the course of 2022. An expert group evaluated the progress made since the 2016 Review and the perception of gender equality among higher education staff. The Second Review made recommendations to ensure the continued advancement of gender equality in the higher education sector. Initial monitoring based on these recommendations began in 2023.
Norway	The Committee for Gender Balance and Diversity in Research (KIF) undertakes an annual review of higher education and R&I institutions with GEPs. In this context, GEP monitoring uses reports of gender-disaggregated data on research careers, including on women in grade A positions, and students. The Ministry of Education and Research commissions these reports annually. The equality and discrimination law requires institutions to report on equal pay every other year.
Spain	Spain does not consider itself to have a national GEP monitoring system as such. There is a requirement to develop measures against sexual harassment and gender-based harassment and these are monitored on an annual basis. The Women and Science Unit of the Spanish Ministry of Science and Innovation is responsible for the *Científicas en cifras* biennial report which is used as a monitoring tool.

Source: Knapińska and Chrobak-Tatara, 2023

4. EVALUATION FRAMEWORK FOR GENDER EQUALITY POLICIES IN ERA

4.1 Definitions of key concepts

4.1.1 Evaluation and monitoring

Evaluation determines the merit, worth, or value of things. The evaluation process identifies relevant values or standards that apply to what is being evaluated, performs empirical investigation using techniques from the social sciences, and then integrates conclusions with the standards into an overall evaluation or set of evaluations (Scriven, 1991).

For the purpose of this report and in line with OECD, policy evaluation is defined as a "structured and objective assessment of an ongoing or completed policy or reform initiative, its design, implementation and results. Its aim is to determine the relevance and fulfilment of objectives, efficiency, effectiveness, impact and sustainability as well as the worth or significance of a policy" (OECD, 2020: 15). Similarly, evaluation is "the systematic assessment of the design, implementation and outcomes of an intervention (policy or programme). This involves understanding how an intervention is being, or has been, implemented and what effects it has, for whom and why. It identifies what can be improved and estimates its overall impacts and cost-effectiveness" (HM Treasury, 2020: 9). The main purpose of an evaluation may differ due to the specific context of a policy/programme. Evaluation standards stipulate that good evaluation should focus on the utilisation of the evaluation results. Typical purposes include:

- To comply with accountability requirements.
- To assess the effects as well as the effectiveness and efficiency of the policy/programme.
- To provide empirical evidence for informed decision making.
- To learn from experiences with policy/programme implementation in order to further develop the policy/programme.

Evaluation questions may be posed at different stages of the policy process. Ex-ante, prior to the introduction of a policy (during the policy design phase), during the implementation of the policy, or after a designated period of implementation (ex-post).

Figure 1: Evaluation stages
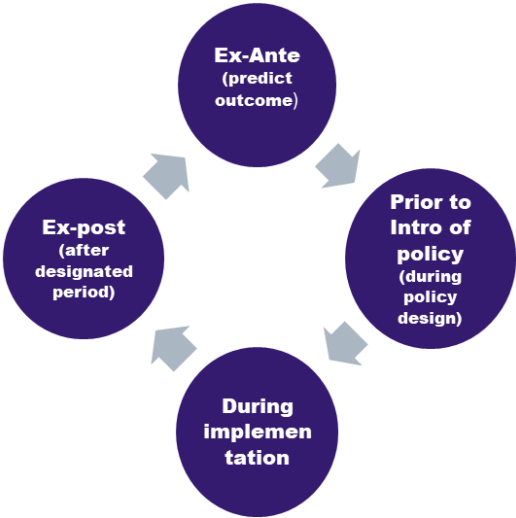
Source: Own elaboration

An ex-ante evaluation is concerned with determining whether the design of a policy is such that the desired objectives can be achieved. Assumptions about the expected mode of action are subjected to close scrutiny and, ideally, supported by theoretical considerations or experience with other policies.

The aim of an accompanying evaluation is to analyse the implementation of a policy in order to determine whether it is being carried out in accordance with the design, thus ensuring the realistic achievement of the objectives. An ex-post evaluation is a retrospective assessment of a measure's implementation. It involves examining whether the objectives were achieved, whether the desired changes were realised, and whether the cost-benefit ratio was favourable.

It is important to distinguish evaluation from social science research in general, as well as from monitoring. Evaluation asks specific questions which are answered by social science approaches. Furthermore, evaluation has defined specific standards (Yarbrough et al., 2010) and aims at being transparent regarding its values (Schwandt and Gates, 2021). The dominant social science doctrine for many decades prided itself on being value free. As such, the distinction between evaluation and research is evident in not just the particular questions posed but also the contextual framework that emerges from a specific client-contractor relationship, the political interest in the outcomes, and the resistance encountered in the field.

In contrast to evaluation, monitoring is defined as a continuing function that uses the systematic collection of information on specified indicators to provide information to key stakeholders on an ongoing intervention. This information should include indications of the level of progress and achievement of the objectives as well as the use of any allocated funds (see also D5.1 and D5.2). As already described, evaluation aims at answering specific questions going beyond monitoring questions but builds ideally on monitoring data (e.g. Wroblewski and Leitner, 2022). Hence, in an ideal case, monitoring and evaluation go hand in hand; neither is more important than the other. **Monitoring ensures that the right thing is done, while evaluation ensures that the right outcomes are achieved.** Table 3 and Table 4 summarise the main differences between evaluation and monitoring.

Table 3: Definition of evaluation and monitoring

Evaluation	Monitoring
The systematic and objective assessment of a planned, ongoing or completed intervention, its design, implementation and results. The aim is to determine relevance, coherence, effectiveness, efficiency, impact and sustainability. Evaluation also refers to the process of determining the worth or significance of an intervention.	A continuing process that involves the systematic collection or collation of data (on specified indicators or other types of information). Provides the management and other stakeholders of an intervention with indications of the extent of implementation progress, achievement of intended results, occurrence of unintended results, use of allocated funds and other important intervention and context-related information.

Source: OECD, 2023.

Table 4: Typical questions

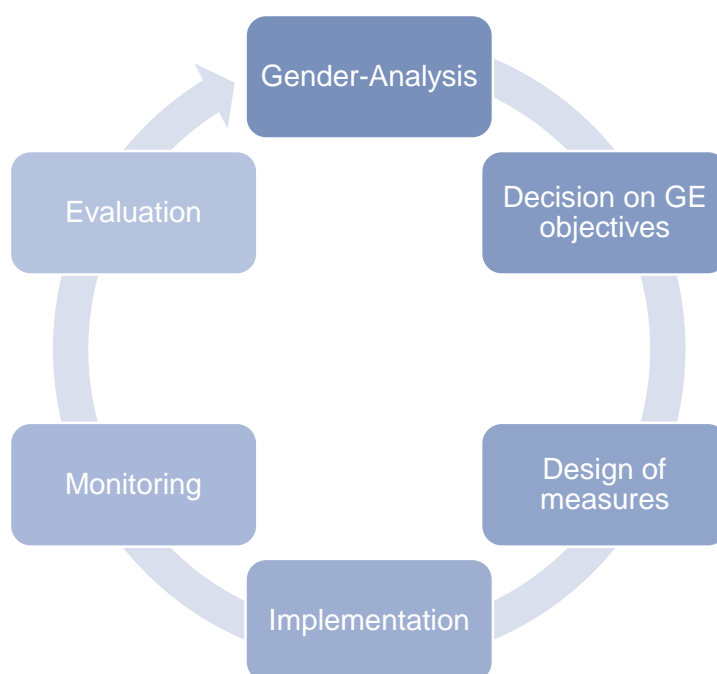| Evaluation	Monitoring
- Does the policy/programme work? - Why does the policy/programme work? - For whom does the policy/programme work? - Does the policy/programme produce unintended effects?	- Is the context changing? If so, how? - Is the policy/programme being implemented? - Are there deviations from the original concept? - Are envisioned targets met?

• Is the policy/programme effective and efficient? • What is the relationship between costs and benefits?	• Is there a need for adaptation of the policy/programme due to these deviations? • Does the implementation contribute to the achievement of objectives?

Source: *Own elaboration*

4.1.2 Monitoring and evaluation as elements of a complete policy cycle

The GENDERACTIONplus approach is based on the concept of a complete policy cycle, which encompasses the development and implementation of policies. Evaluation should either build on the results of monitoring and assess implementation or be implemented as an ex-post evaluation. In the first case, evaluation focuses on the implementation of the policy and possible deviations from the concept which may lead to an ineffective implementation, or which makes goal achievement unrealistic. In the second case, the evaluation takes place after a certain period of policy/programme implementation and focuses on goal achievement and impacts.

Figure 2: Complete cycle for the development and implementation of gender equality policies

Source: *Wroblewski, Angela (2024)*

4.1.3 Challenges to the monitoring and evaluation of gender equality policies

Against this background, the development of meaningful indicators that can be used for gender analysis, monitoring and evaluation represents a significant challenge. For that reason, the monitoring of a GEP ideally contains indicators which allow the assessment of GEP implementation as well as its outcomes.[2] As gender equality is a complex construct, a gender indicator can only be an approximation. As Tony Beck (1999: 7) puts it: "An indicator is an item of data that summarises a large amount of information in a single figure, in such a way as to give an indication of change over time, and in comparison, to a

---

[2] The following section is based on Wroblewski and Leitner (2022).

norm". Hence, a gender indicator is defined as an indicator which captures gender related change over time.

The deviation between the indicator and the construct to be measured must be reflected on and considered in the interpretation. It is also important to note the difference between indicators and statistics. Statistics present facts while indicators involve comparison to a norm and their interpretation. In this context the conceptualisation of gender and its equivalent in empirical evidence is of specific relevance. While gender from a theoretical point of view is seen as socially constructed (Butler, 1990; West and Fenstermaker, 1995; West and Zimmermann, 1987), in administrative data gender is usually coded in a dichotomous way (female/male). The variable sex or gender available in empirical data does not provide information about gender (Hedman, Perucci, and Sundström, 1996; United Nations Economic Commission for Europe and World Bank Institute, 2010; Döring, 2013). In addition, sex and gender interact with each other e.g. medical clinical trials being conducted primarily by men with the main reference being the male body. Another example is when gender research from the 1960s mainly focused on women and has been mainly conducted by female researchers (Stefanick and Schiebinger, 2020). Gender refers to norms, behaviours and roles associated with being a woman, man, girl or boy, as well as the relationships with each other. As a social construct, gender can change over time. Furthermore, both sex and gender produce inequalities that intersect with other social and economic inequalities.

When discussing gender-based discrimination, gender intersects with other factors of discrimination such as age, socioeconomic status, disability, ethnicity, gender identity and sexual orientation (van der Haar and Verloo, 2013; Walby, Armstrong, and Strid, 2012; Verloo, 2006). To approach this complex construct in empirical analysis, other variables are considered. The availability of information regarding other relevant characteristics like disability, care responsibilities or gender identity is the exception rather than the norm. The assumption that specific characteristics like care responsibilities mainly affect women may lead to an unintended emphasising of gender stereotypes and supports the identification of discrepancies as gender-based even though they are based on other characteristics (Stadler and Wroblewski, 2021; Degele, 2008). This problematic aspect gains additional relevance because available data might be gender biased, especially in the case of administrative data. The production of administrative data tends to overrepresent realities which are male dominated. This becomes a problem if such administrative data is used for analysing gender imbalances e.g. when data from labour market statistics is used to analyse gendered patterns of employment because official statistics only consider paid employment (D'Ignazio and Klein, 2020; Criado-Perez, 2019; Hedman et al., 1996).

Gender indicators are not merely statistics on men and women. They highlight the contributions of all genders to society, to science and research and their different needs and challenges. To depict this complex picture adequately, a set of indicators that covers all relevant aspects is necessary. The interpretation of one isolated indicator may be misleading. In the context of gender equality policies, monitoring must contain indicators addressing all objectives of the GEP. Monitoring should contain indicators about women's representation in all fields and at all hierarchical levels, indicators which represent structural barriers for women such as women's participation in decision making, and indicators which demonstrate the integration of the gender dimension in research content and teaching.

Data availability differs for these dimensions which consequently effects the validity of indicators. For example, it is easier to depict women's representation compared to the gender dimension in research content and teaching (European Commission, 2019a-d; European Commission, 2018). Regarding the objective of gender balanced representation in all fields and at all hierarchical levels, there is good availability of data in most cases. Institutions know about the gender composition of students and staff in different disciplines as well as in decision making bodies/structures. Furthermore, information about the share of women at different hierarchical levels is usually available. There is less data available regarding structural barriers for women's careers e.g. information about the representation of women in different stages of appointment procedures is not available by default. Limited data is available regarding the integration of the gender dimension in research and teaching content. Even less information is available in relation to gender-based violence including sexual harassment.

To compensate for this lack of data, different data sources are combined, including administrative data which is electronically available (e.g. student or staff records) or project or publication repositories (to identify projects and publications with gender content). However, it is not always possible to extract gender relevant information from electronic data management systems (e.g. in the context of recruitment). Hence, the development of indicators for monitoring and evaluation often requires an adaptation of existing data sources or the establishment of new data collection mechanisms and specific data collection tools e.g. survey. Indicators can be either quantitative (e.g. number, percentage, ratio) or qualitative (e.g. assessment in qualitative terms). Especially in evaluation, qualitative and quantitative approaches are combined to compensate for the shortcoming of both approaches (e.g. Flick, 2018; Mertens, 2017).

The challenges described above refer to the development of indicators for monitoring and evaluation, but it should be noted that developing indicators for evaluation faces two more challenges. First, it is difficult to measure change. This requires not only meaningful indicators to measure the expected outcome but also information for at least two timepoints. Second, this change has to be attributed to the policy/programme. Hence, it is necessary to establish a causal relationship between the change and the intervention. This requires a research design that provides a counterfactual. Evaluation designs based on experimental or quasi-experimental designs allow for causal conclusions (Shadish et al., 2001). For instance, it is difficult to measure the sectoral impact of interventions that have a wider effect beyond the implementing organisation; for example, the benefits of mentorship programmes or scholarships for early career researchers that move to other organisations, sectors or countries.

It is also important to highlight that monitoring and evaluation will remain a challenge for NA if there is insufficient human and financial resources to commission and support the work.

4.2 Key actors in the evaluation framework

In the context of the ERA, two principal objectives of evaluation can be identified: (1) to provide decision-makers with relevant information to enable evidence-based decision making, and (2) to provide input for the further development of a policy/programme, in this context GEPs, in order to achieve the objectives in a more efficient and effective manner. Consequently, evaluation is strongly focused on the utilisation of evaluation results with two different audiences.

It is important to distinguish between these two distinct audiences: decision-makers and implementers. Decision-makers can be national authorities or the management of an institution implementing a GEP. They are typically responsible for initiating a policy or a GEP. However, they are not typically involved in operational implementation. Implementers, on the other hand, are experts responsible for the implementation of the GEP or its actions. Consequently, evaluation reports must address these two groups specifically and provide recommendations for them.

Michael Quinn Patton (2008) developed the Utilisation Focused Evaluation, an approach based on the principle that an evaluation should be judged on its usefulness to its intended users. Evaluations should, therefore, be planned and conducted in ways that enhance the likely utilisation of both the findings and of the process itself in order to inform decisions and improve performance. Patton (2008) argues that research on evaluation demonstrates that the intended users of any evaluation are more likely to use the findings if they feel ownership of the evaluation process, and if they have been actively involved from the start. As such, evaluation aiming to support learning processes should be participatory and time flexible. Importantly, it needs evaluators who are competent in conducting such complex evaluations.

Utilisation-focused evaluation is linked to or can play a central role in organisational learning. This presupposes the willingness of the institution implementing a policy/programme to participate in a process of reflection and to allow a critical examination of the implementation of the policy/programme. There must be an explicit commitment from the management to the further development of the policy/programme, a basis of trust among the people involved in the implementation and an explicit willingness to learn from experience, i.e. to adapt the concept of the policy/programme or the implementation if necessary.

It is necessary to determine who should be responsible for commissioning and carrying out an evaluation. Typically, a utilisation-focused evaluation is conducted by external experts who possess expertise in evaluation, gender equality, and organisational development in the relevant field (in our context R&I). Focusing on the organisational learning aspect, it is assumed that an institution implementing a GEP has an intrinsic interest in the evaluation and its results. This ensures that the evaluation can be conducted without significant resistance and that the recommendations can be implemented effectively.

If experts involved in policy/programme implementation have concerns that the evaluation could be misused (e.g. to cease or downsize a successful policy/programme), this will lead to resistance and make it more challenging to engage in a constructive reflection process.

What is the role of the EU, national authorities and the RPOs themselves in the context of the GEPs? The European Commission and national authorities should aim at fostering and sustaining this intrinsic interest. This could be supported by targeted incentives such as:

- Creating awareness of evaluation as part of the professional implementation of GEPs.
- Linking the continuation of GEP-related funding to the implementation of an evaluation and proof that the recommendations have been considered.
- Contribution to the costs of the evaluation.
- Providing tools and training to build up basic evaluation expertise.
- Anchoring these topics in a European and national discourse on the development and implementation of GEPs.
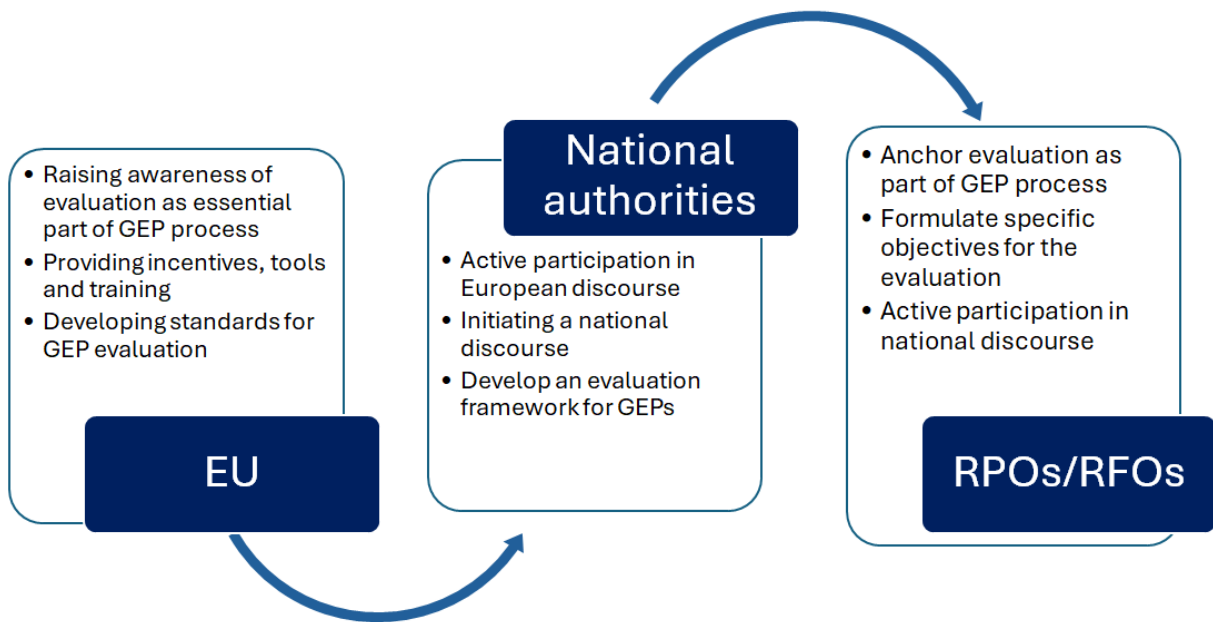
The primary role of the EC is to initiate and advance a discourse with a view to fostering an alignment and coordination among member states and RPOs in the ERA ecosystem. This should also encompass the establishment of common standards for the evaluation of GEP implementation.

The role of national authorities would be threefold:

- National authorities should take an active role in the European discourse described above.
- They should initiate a national discourse involving the RPOs and RFOs under their remit with the aim to arrive at a common understanding of the role of evaluation, standards for evaluation and the utilisation of evaluation.
- National authorities should develop a national framework for evaluation of GEP implementation which includes the development of specific standards related to evaluation and mechanisms of collecting and analysing GEP evaluations carried out by institutions. The latter could be used to identify good practices which could be an input to the national discourse.

The main challenge of the evaluation framework described is raising awareness among RPOs and RFOs about the relevance of evaluation for the effective and efficient implementation of GEPs. It is essential to create a climate in which evaluation is considered a valuable tool for the continued improvement of GEPs, rather than a bureaucratic requirement or a mere cost factor. Productive and constructive dialogue at European and national levels will assist in developing this climate.

Figure 3: European and national discourse to establish an evaluation framework

EU
- Raising awareness of evaluation as essential part of GEP process
- Providing incentives, tools and training
- Developing standards for GEP evaluation

National authorities
- Active participation in European discourse
- Initiating a national discourse
- Develop an evaluation framework for GEPs

RPOs/RFOs
- Anchor evaluation as part of GEP process
- Formulate specific objectives for the evaluation
- Active participation in national discourse

Source: Own elaboration

5. PRINCIPLES FOR EVALUATION

Outlining evaluation principles is key to forming a cohesive evaluation framework.

- Evaluation should be based on a **feminist approach** (Brisolara et al., 2014; Podems, 2011)**.** It is essential to examine the power structures that exist within the organisation and the wider environment, both in the analysis of the context and in the implementation and impact of measures.
- Evaluation should follow a **participatory approach** (Patton, 2008). This requires the identification of all relevant stakeholder groups and including them in the implementation of the evaluation. The objective is to gain insight into the perspectives of underrepresented groups with regard to the measures in question, their implementation and the resulting impact.
- Evaluation should follow a **utilisation-focused approach** (Patton, 2008). This highlights the importance of evaluation in the further development of measures or GEPs, ensuring that evaluation is not merely a pro forma exercise.
- **Evaluation design should allow for flexibility** to be able to respond to changes in the context or to address relevant issues emerging during evaluation implementation.
- Evaluation should refer to a **comprehensive and intersectional understanding of gender equality**. On the one hand, this requires that the multidimensional equality objective of the EC is the guiding principle and therefore all topics proposed by the Commission should also be considered in the evaluation. If all topics are not currently addressed in the GEP, it is imperative to identify any blind spots and incorporate them into the future development of the equality plan.
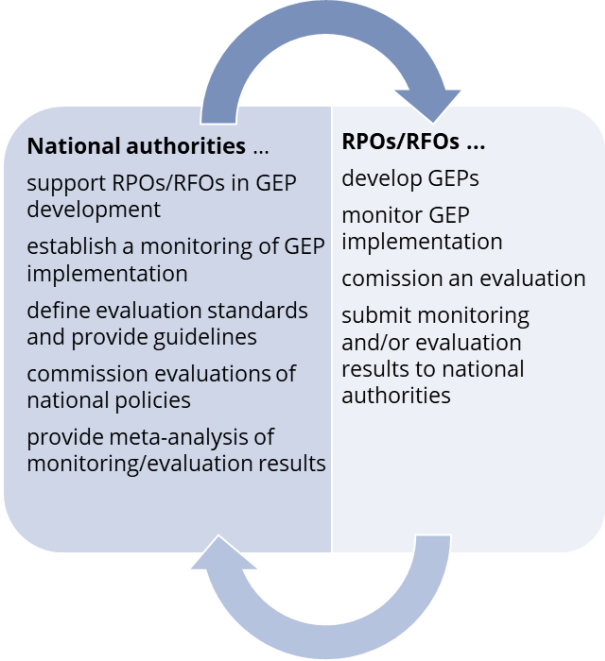
The **profile of an evaluator or evaluation team** is complex and comprises theoretical knowledge and social skills. Gutknecht-Gmeiner et al. (2017) include the following competencies in the profile of an evaluator with gender expertise:

- In-depth knowledge of gender mainstreaming as a political strategy.
- A comprehensive understanding of relevant gender theories (including feminist, critical male/masculinity, queer theory) and intersectional theory.
- Competence to conceptualise and facilitate reflection processes with those affected by gender equality policy/GEP and those involved in its implementation.
- Capacity to generate and facilitate discussions on gender issues.
- Capacity to deal with attacks and resistance.
- Capacity to deal with gender-specific power and domination structures and relationships.
- Ability to recognise gender-specific disparities in all dimensions of the evaluation object and its context as well as to take these factors into account throughout all steps of the evaluation. This entails the capacity to apply or develop gender-specific criteria and indicators for the evaluation objects/questions.
- Ability to adopt a gender-sensitive approach when conducting the evaluation, particularly with regard to data collection methods and communication with target groups.

National authorities should consider the competencies and knowledge that are relevant to their national context. It is possible for national authorities to apply these standards to evaluations that they commission themselves. Furthermore, they can also provide incentives for HEIs or RPOs/RFOs that have their GEP evaluated to apply these standards. For instance, if an evaluation of the GEP is agreed as part of a performance agreement with a HEI, quality standards for the evaluation could also be agreed upon. Such activities facilitate the development of a culture that values and utilises monitoring and evaluation.

A crucial aspect of fostering such a culture is the manner in which national authorities approach monitoring and evaluation. If monitoring and evaluation are requested from RPOs/RFOs, these should be employed within the context of a policy discourse. A meta-analysis of monitoring outcomes or evaluations may serve as a foundation for a discourse on the advancement of GEP implementation.

National authorities ...

support RPOs/RFOs in GEP development

establish a monitoring of GEP implementation

define evaluation standards and provide guidelines

commission evaluations of national policies

provide meta-analysis of monitoring/evaluation results

RPOs/RFOs ...

develop GEPs

monitor GEP implementation

comission an evaluation

submit monitoring and/or evaluation results to national authorities

Source: Own elaboration

6. ROLE OF KEY ACTORS IN THE EVALUATION FRAMEWORK

6.1 Role of the European Commission

While methodologies for monitoring and evaluation of individual institutional GEPs exist, there is an immediate need for agreed monitoring and evaluation indicators for evaluation at European level to support inclusive gender quality in the ERA. Learnings from the GENOVATE, EFFORTI and CALIPER projects (GENOVATE, 2017, EFFORTI, 2019, CALIPER, 2023), and the Impact Driver Model (Mergaert et al., 2022) provide guiding principles and suggested methodologies for evaluation. The EC should emphasise the need to fully incorporate an intersectional and inclusive approach to gender equality in GEPs (and related monitoring and evaluation) to address groups that are at risk of discrimination. While it is acknowledged that intersectional analysis of GEP implementation remains a challenge, the EC needs to emphasise the critical importance of embedding this perspective in policy implementation, monitoring and evaluation. As outlined by GENDERACTIONplus (2024b), "by considering the workings of multiple systems of disadvantage across dimensions such as gender, ethnicity, and class, it is not possible to simply focus on academic careers and progression without considering how equality dimensions affect academic advancement differently for people of different backgrounds. Intersectionality for policymakers is therefore an analytical tool and lens that can be used to ensure that policies are inclusive and recognise the diversity of groups in academia".

The EC could consider how the Horizon Europe GEP requirements and recommended thematic areas could be used as a starting point for European level monitoring and evaluation. The EC should also place greater emphasis on the need for national level monitoring and evaluation of GEPs, as this is critical to measuring and understanding the impact of GEPs as an instrument for institutional change and ensuring the sustainability of gender equality work. Member states should take an active role in advancing the goals of the ERA Policy Agenda 2025-2027 at national level.

Guidance from the EC would assist national authorities and responsible bodies in developing strong approaches to national level evaluation that are reflective of country-specific contexts while also providing a consistent baseline for benchmarking across the ERA. This guidance should be accompanied by training/instructions for national authorities on what is required by the EC.

6.2 Role of national authorities and member states

Ideally, national authorities would be able to commit to a monitoring and evaluation cycle (rather than solely monitoring), however, the observed differences in GEP implementation and related monitoring at national level means this is not currently happening in practice. As such, there is a need to articulate a sequence of steps/actions for national authorities to follow in relation to implementing an evaluation framework at national level.

The limits of evaluation should be understood as national authorities are not evaluating individual institutional GEPs; rather, they are providing a national level framework for the evaluation of GEP implementation. This ensures any evaluation measures are based on identifiable and specific markers, thus providing clarity for all stakeholders. As such, there is a need for national authorities to understand their own objective and related intended use of the evaluation in that context. As part of the national framework, national authorities could conduct a meta-analysis of evaluation conducted at RPO and RFO level.

With these considerations in mind, national authorities should provide common evaluation standards for their country and provide guidance and/or training in relation to this. National authorities also have a responsibility to take action based on the results of a meta-evaluation.

The evaluation should follow a participatory approach (Patton, 2008). This requires the identification of all relevant stakeholder groups and including them in the implementation of the evaluation. The objective is to gain insight into the perspectives of underrepresented groups with regard to the measures in question, their implementation and the resulting impact.

Sufficient human and financial resources are needed in order to develop an evaluation framework, to set evaluation standards and to commission an evaluation.

6.3 Role of institutions (RPOs and RFOs)

As part of professional GEP implementation institutions should engage in regular monitoring and evaluation of their GEP and provide relevant self-assessment information to the relevant ministry/national authorities/evaluation bodies. Institutions should commit to critical self-assessment and evaluation to maximise the impact of their GEP. The tendency to only report on what is going well should be avoided; it is beneficial for national authorities to understand where common challenges lie so that appropriate national level interventions/incentives/actions can be put in place to mitigate/overcome these challenges.

In order to carry out high quality evaluations, a number of conditions need to be met at the institutional level: The willingness to engage with the evaluation results, i.e. to allow critical reflection on the status quo of gender equality and the implementation of the GEP, and the willingness to learn from the evaluation results to further develop gender equality policies/the GEP.

It is important to develop gender competent evaluation expertise within the institution to commission and support high quality evaluations. RPOs/RFOs should allocate sufficient resources and a realistic time frame to carry out institutional level evaluation. When awarding an evaluation contract, it is important to ensure that the evaluator(s) has/have gender expertise in the team.

In the context of GEPs, RFOs may have a dual role. In their capacity as an employer/public body/Horizon Europe participant, they develop and implement an internal GEP, which aims to address gender inequalities within their own organisation. In their capacity as funding bodies, they may also establish equality measures or a GEP as a criterion for the allocation of funding. Given the focus of this guidance on national authorities, the role of RFOs as initiators of GEPs at applicant institutions is not addressed further here.

Overall, it is assumed that RFOs act in accordance with national regulations and thus support national efforts and measures or may themselves be affected by national regulations. Where a participatory approach is provided by their national authority, RFOs should engage fully with this and comply with the minimum standards for evaluation.

6.4 Role of national authorities in monitoring and evaluation of GEPs

As described in chapter 4, evaluation builds on strong monitoring which provides information on the implementation of policies. However, the expectations for monitoring differ from the perspective of a national authority and an RPO/RFO implementing an institutional GEP.

The aim of a national authority is to know the extent to which the RPOs and RFOs under their remit have (1) developed a GEP, (2) implemented a GEP and (3) if these GEPs support structural change at institutional level. National monitoring should address the first two questions. To answer the third question, a national evaluation framework should be developed which provides incentives for RPOs/RFOs to monitor their institutional GEPs and which also defines the role of the national authority in that context.

National authorities also play a central role in developing an evaluation culture at national level. First, authorities should evaluate state policies and programmes under their remit. Second, national authorities are in a position to create the framework conditions and incentives for higher education

institutions to evaluate their gender equality policies and GEPs in terms of effectiveness and efficiency. Additionally, national authorities should consider intersectionality as crosscutting in evaluating the impact of GEP implementation and consider how taking an intersectional lens can assist in creating a robust and relevant evaluation framework.

As such, it is crucial for national authorities to define and demand standards of good evaluation.

It is therefore recommended that they establish fundamental criteria for the evaluation of gender equality policies in general and GEPs in particular. This also includes the definition of competency profiles for evaluators. Mutual learning between member states that are more advanced and less advanced may be required in order to develop the capacity of national authorities in relation to this.

What is the aim of monitoring?

- To be informed about extent of GEP implementation. How many institutions are eligible for EU funding? How many institutions comply with national regulations?
- To be able to assess the quality of GEP implementation. Do the GEPs fulfil the national requirements?
- To be able to act if needed.

Why should national authorities support GEP evaluation at institutional level?

- To support the effective and efficient implementation of GEPs.
- To ensure that GEP implementation meets specific standards.
- To be able to take supportive action if needed.

7. CONCLUSIONS AND RECOMMENDATIONS

7.1 Background to recommendations

The purpose of this guidance is to define the key actors in the evaluation of GEP implementation and set out suggested ways to advance a framework of monitoring and evaluating GEP implementation at scale in the ERA. Ideally, national authorities would be able to commit to monitoring and evaluation at the same time but the varying levels in GEP implementation and related stages of monitoring at national level means this is not currently possible in all member states.

It is not intended to provide a systemic framework of additive stages in tandem with the monitoring stages proposed by GENDERACTIONplus D6.2 (Rothwell and Young, 2024). As specific indicators are not defined, the guidance provided by national authorities, ideally in cooperation with the EC, should suggest dimensions for the development of indicators, considering each national context. As outlined by other EU projects, localised contexts must be understood when considering evaluation of GEP implementation (Kalpazidou Schmidt et. al., 2017: 42, 66). A country that has only just introduced a GEP requirement will have different cultural, political, and organisational contextual policies than a country that has had successive years of GEP implementation.

7.2 Recommendations for national authorities

While specific recommendations for national authorities relating to intersectional policymaking have been outlined by GENDERACTIONplus (GENDERACTIONplus, 2024a), the points below focus specifically on evaluation of GEP implementation. For clarity, the guidance is grouped under four headings: Setting the national context for evaluation, engaging with institutions, carry out analysis, and creating an evidence base for change.

Setting the national context for evaluation

- National authorities should clearly define the objectives of their inclusive gender equality work or policy in order to foster a supportive environment for the evaluation of GEP implementation. Providing institutions with a greater understanding of these objectives will assist in linking institutional work to national goals.

- National authorities should seek to align objectives with existing policy areas (e.g. Horizon Europe GEP requirements and recommended thematic areas) to create an evaluation framework. This will ensure that national authorities can establish a framework that is responsive to the European policy context while also being adaptable to their national context.

- A benchmark should be created (if not already) to understand current standing of GEP implementation and to evaluate compliance with national requirements and recommendations. This benchmark can be used in successive years to analyse how recommendations and interventions have had an impact on implementation.

- National authorities should advocate for strong guidance from the EC related to monitoring and evaluation of GEP implementation, the development of gender indicators, along with the necessary resources and training to ensure evaluation competence.

- National authorities should engage with the gender equality competence facility to be established by the EC to ensure access to support for the evaluation of GEP implementation.

- While there remains a lack of unified understanding around intersectionality in the ERA, national authorities should ensure that long-term evaluation goals incorporate intersectional and multidimensional approaches, in order to build capacity towards inclusive GEPs.

- National authorities should define what it means to integrate an intersectional approach in inclusive gender equality work and consequently involve underrepresented groups when identifying goals and measures.

- National authorities should establish a strong monitoring of GEP implementation as a starting point for GEP evaluation.

- National authorities should develop their capacity to define minimum standards for evaluation.

Engaging with institutions

- National authorities should clearly communicate the purpose of the national evaluation to stakeholders. This should include what information will be collected, what it will be used for, and how the relevant data protection requirements will be met.

- National authorities should communicate clearly on the quality standards for the evaluation of GEPs.

- National authorities and institutions should consider the learnings and recommendations for institutional level GEP evaluation already provided by projects such as GENOVATE, CALIPER, and EFFORTI.

- Institutions should be incentivised to participate in meta-analysis of GEP evaluations by providing the necessary information. For example, institutions can provide national authorities with their GEP evaluation which can feed into a meta-analysis to understand the current national state of play.

- National authorities should facilitate cooperation and mutual learning among institutions to avoid 'reinventing the wheel'. This mutual learning could also be supported by institutions sharing their GEPs on relevant sites e.g. genderportal.eu.

- A clear timeframe for reporting should be part of the guidance provided to institutions. This should include reporting deadlines and the expected timeline for national level reports in order to manage expectations.

- If there is no current GEP requirement at national level, considerations for other gender equality work undertaken by institutions should be considered for inclusion in a monitoring and evaluation framework.

Carry out analyses

- While each national authority should create indicators that are context specific there are common questions that should be asked. For example, Are GEPs meeting the criteria as outlined in the evaluation framework? Does GEP implementation contribute to changes in institutional culture? What would have been the situation without intervention?

- National authorities and institutions should build capacity for supporting evaluation within their organisation and/or identify appropriate and relevant evaluators/evaluation bodies that can support this work. National authorities and institutions should consider the relevant evaluator competencies and knowledge required for their national context when commissioning or conducting an evaluation.

- The findings from national monitoring should be leveraged to identify additional areas for evaluation.

Create an evidence base for change

- To promote mutual learning, national authorities should share information on good practices and areas of common challenge.

- Using the outcomes of national level evaluation, national authorities should make recommendations for action in areas of common challenge or consider how these challenges could be addressed through a national approach.
- National level evaluation is not intended to assess the merit of individual institutional GEPs or the performance of individual institutions in relation to gender equality. However, where a lack of progress or considerable challenges are identified by an institution through the evaluation process, national authorities can seek to put supportive measures in place for the institution.

7.3 Conclusions

With variance in GEP adoption and implementation across MS/AC, evaluation at national level is central to understanding the impact of GEPs as an instrument for institutional change. Effective evaluation at national level can assess system level change, identify common challenges and provide opportunities for benchmarking and mutual learning.

The guidance proposed by GENDERACTIONplus in D6.2 and D6.3 aims to underline the important role of monitoring and evaluation in long-term structural change management, both for individual institutions and for a set of stakeholders within a national context. Rather than be limited to aspects of compliance, monitoring and evaluation can be harnessed in order to further develop institutional sustainability for inclusive gender equality work and enhance the evidence base for institutional and national change.

8. REFERENCES

Bondestam, Fredrik; Lundqvist, Maja; and Young Håkansson, Susanna (2023): *D3.1 Benchmarking report on GBV and SH targeting national authorities and RFOs*, GENDERACTIONplus, retrieved 7 November 2024: https://genderaction.eu/wpcontent/uploads/2023/05/101058093_GENDERACTIONplus_D3.1_Benchmarking-report-on-GBV-and-SH-targeting-national-authorities-and-RFOs.pdf

Beck, Tony (1999): *Using Gender-sensitive Indicators. A Reference Manual for Governments and Other Stakeholders*, London: Commonwealth Secretariat.

Brisolara, Sharon; Seigart, Denise; and SenGupta, Saumitra (2014). *Feminist Evaluation and Research: Theory and Practice*, New York: Guilford Publications.

Butler, Judith (1990): *Gender Trouble: Feminism and the Subversion of Identity*. New York/London: Routledge.

CALIPER Project (2023): *Linking research and innovation for gender equality*, DOI: 10.3030/873134, retrieved 7 November 2024 https://cordis.europa.eu/project/id/873134

Council of the European Union (2021a): *Council recommendation on a Pact for Research and Innovation in Europe*, 13701/21, retrieved 7 November 2024: https://era.gv.at/public/documents/4620/Pact_for_RI_26112021_EN.pdf

Council of European Union (Slovenian Presidency) (2021b): *Ljubljana Declaration on Gender Equality in Research and Innovation,* retrieved 7 November 2024: https://www.gov.si/assets/ministrstva/MIZS/Dokumenti/PSEU/Ljubljana-Declaration-on-Gender-Equality-in-Research-and-Innovation-_endorsed_final.pdf

Criado-Perez, Caroline (2019): *Invisible Women: Data Bias in a World Designed for Men*. New York: Abrams Press.

Degele, Nina (2008): *Gender/Queer Studies: Eine Einführung* [Gender/Queer Studies: An Introduction]. Stuttgart: UTB.

D'Ignazio, Catherine and Klein, Lauren F. (2020): *Data Feminism*. Cambridge: The MIT Press.

Döring, Nicola (2013): Zur Operationalisierung von Geschlecht im Fragebogen: Probleme und Lösungsansätze aus Sicht von Mess-, Umfrage-, Gender- und Queer-Theorie [On the Operationalisation of Gender in Questionnaires: Problems and Approaches from the Perspective of Measurement Theory, Survey Theory, Gender Theory, and Queer Theory]. *Gender. Zeitschrift für Geschlecht, Kultur und Gesellschaft,* 5(2), 94-113. Retrieved 7 November 2024: http://www.nicola-doering.de/wp-content/uploads/2014/08/D%C3%B6ring-2013-Zur-Operationalisierung-von-Geschlecht-im-Fragebogen.pdf

EFFORTI Project (2019): *Evaluation Framework for Promoting Gender Equality in R&I*, DOI:10.3030/710470 retrieved 7 November: https://cordis.europa.eu/project/id/710470

Kalpazidou Schmidt, Evanthia; Bührer, Susanne; Schraudner, Martina; Reidl, Sybille; Müller, Jörg; Palmen, Rachel; Haase, Sanne; Graversen, Ebbe Krog; Holzinger, Florian; Striebing, Clemens; Groó, Dora; Klein, Saskia; and Høg Utoft, Ea (2017): D3.3 *A Conceptual Evaluation Framework for Promoting Gender Equality in Research and Innovation,* EFFORTI, retrieved 7 November 2024: https://efforti.eu/publications/project-reports

European Commission (2018): *Monitoring the Evolution of Benefits of Responsible Research and Innovation. The Evolution of Responsible Research and Innovation: the Indicators Report*. Brussels. Retrieved 7 November 2024 https://data.europa.eu/doi/10.2777/728232

European Commission (2019a). *European Research Area Progress Report 2018*. Brussels. Retrieved 7 November 2024 https://ec.europa.eu/info/publications/era-progress-report-2018_en

European Commission (2019b). *ERA Monitoring Handbook 2018*. Brussels. Retrieved 7 November 2024: https://ec.europa.eu/info/publications/era-progress-report-2018_en

European Commission (2019c). *ERA Progress Report 2018. Data gathering and information for the 2018 ERA monitoring – Technical Report*. Brussels. Retrieved 7 November 2024: https://ec.europa.eu/info/publications/era-progress-report-2018_en .

European Commission (2019d). *She Figures 2018*. Brussels. Retrieved 7 November 2024: https://www.etag.ee/wp-content/uploads/2019/03/She-Figures-2018-1.pdf

European Commission (2020): *Gender equality in research and innovation,* retrieved 7 November 2024: https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/democracy-and-rights/gender-equality-research-and-innovation_en

European Commission (2021): *European Research Area Policy Agenda. Overview of actions for the period 2022-2024*, retrieved 7 November 2024: https://commission.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

European Commission (2022a): *Approaches to inclusive gender equality in research and innovation (R&I)*, retrieved 7 November 2024: https://data.europa.eu/doi/10.2777/004694

European Commission (2022b): *Horizon Europe Work Programme 2021 – 2022 General Annexes 2022*, retrieved 7 November 2024: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/wp-call/2021-2022/wp-13-general-annexes_horizon-2021-2022_en.pdf

ERAC SWG on Gender in Research and Innovation (2021): *Report by the ERAC SWG on Gender in Research and Innovation on Gender Equality Plans as a catalyst for change*, ERAC 1202/21, retrieved 7 November 2024: https://data.consilium.europa.eu/doc/document/ST-1202-2021-INIT/en/pdf

Flick, Uwe (2018): *Doing triangulation and mixed methods*. The Sage Qualitative Research Kit 8. London: SAGE Publications Ltd.

GENDERACTIONplus. (2024a): *A new ERA of inclusion of intersectionality,* retrieved 7 November 2024: https://genderaction.eu/wp-content/uploads/2024/02/GENDERACTIONplus_Position-paper_A-new-ERA-of-inclusion-and-intersectionality.pdf

GENDERACTIONplus. (2024b): Research careers: the ground for gender+ inclusion*, retrieved 14 November 2024: https://genderaction.eu/wp-content/uploads/2024/11/GENDERACTIONplus_Position-Paper-7_Inclusive-Research-Careers.pdf

GENOVATE Project (2017): *Transforming organisational culture for gender equality in research and innovation*, retrieved 7 November 2024: https://cordis.europa.eu/project/id/321378

Gutknecht-Gmeiner, Maria; Eckstein, Kerstin; and Wroblewski Angela (2017). *Genderkompetenz und Genderexpertise – Anforderungen an Kompetenzprofile von Evaluator(inn)en*. Zeitschrift für Evaluation, 16(1), 218-222.

Hedman, Brigitta; Perucci, Francesca; and Sundström, Pehr (1996): *Engendering Statistics: A Tool for Change*. Stockholm: Statistics Sweden. Retrieved 7 November 2024: https://ams-forschungsnetzwerk.at/downloadpub/gender_1996_statistics_601.pdf .

HM Treasury (2020): *Magenta Book. Central Government guidance on evaluation.* London. Retrieved 7 November 2024:
https://assets.publishing.service.gov.uk/media/5e96cab9d3bf7f412b2264b1/HMT_Magenta_Book.pdf

Holt Zachariassen, Heidi; Ghosh, Ella; and Woods, Ross (2023), *D2.1 Benchmarking report on terminology and policy on intersectionality*, GENDERACTIONplus, retrieved 7 November 2024: https://genderaction.eu/wp-content/uploads/2023/05/GENDERACTIONplus_D2.1_Benchmarking-report-on-terminology-and-policy-on-intersectionality.pdf

Knapińska, Anna and Chrobak-Tatara, Magdalena (2023), *D6.1 Benchmarking analysis of monitoring/evaluation of GEPs*, GENDERACTIONplus, retrieved 7 November 2024: https://genderaction.eu/wp-content/uploads/2023/07/GENDERACTIONplus_D6.1_Benchmarking-analysis-of-monitoringevaluation-of-GEPs.pdf

Korsvik, Trine Rogg; González, Lydia; and Dvorackova, Jana (2023), *D4.1 Benchmarking and assessment report on guidelines for sex/gender analysis*, GENDERACTIONplus, retrieved 7 November 2024 https://genderaction.eu/wpcontent/uploads/2023/12/GENDERACTIONplus_D4.1_Benchmarking-and-assessment-report-on-guidelines-for-sex_gender-analysis-.pdf

May, Judith V. and Wildavsky, Aaron B. (eds.) (1978): *The Policy Cycle*. Beverly Hills/London: Sage.

Mergaert L, Cacace M, Linková M. (2022): Gender Equality Impact Drivers Revisited: Assessing Institutional Capacity in Research and Higher Education Institutions. *Social Sciences*. 2022; 11(9):379. https://doi.org/10.3390/socsci11090379

Mertens, Donna M. (2017): *Mixed Methods Design in Evaluation*. Thousand Oaks, CA: SAGE.

OECD (2020): *Improving Governance with Policy Evaluation. Lessons From Country Experiences.* Retrieved 7 November 2024: https://www.oecd.org/en/publications/improving-governance-with-policy-evaluation_89b1577d-en.html

OECD (2023): *Glossary of Key Terms in Evaluation and Results-Based Management for Sustainable Development* (Second Edition), Paris: OECD Publishing, Retrieved 7 November 2024: https://doi.org/10.1787/632da462-en-fr-es

Patton, Michael Q. (2008), *Utilization Focused Evaluation. The new century text*, 4th ed., Thousand Oaks et al.: SAGE.

Podems, Donna R. (2011). Feminist Evaluation and Gender Approaches: There's a Difference? Journal of Multidisciplinary Evaluation 6(14): 1-17.

Rothwell, Jennie; Young, Julie Anne (2024), *D6.2 Guidance on Establishment of Gender Equality Plan Monitoring Systems at National Authority Level*, GENDERACTIONplus, retrieved 7 November 2024: https://genderaction.eu/wp-content/uploads/2024/06/GENDERACTIONplus_D6.2_Guidance-on-establishment-of-GEP-monitoring-systems.pdf

Schwandt, Thomas A.; Gates, Emily F. (2021): *Evaluating and valuing in social research*, New York: The Guilford Press.

Scriven, Michael (1991): *Evaluation thesaurus*. Newbury Park, CA: Sage.

Shadish, William R.; Cook, Thomas D.; Campbell, Donald T. (2001): *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston/New York: Houghton Mifflin Company.

Stefanick, Marcia L., & Schiebinger, Londa (2020): Analysing how sex and gender interact. *The Lancet*, 396(10262), 1552-1554. doi:10.1016/S0140-6736(20)32346-1

Stadler, Bettina, & Wroblewski, Angela (2021): Wissen in Zahlen: Potenziale von Gender-Monitoring im gleichstellungspolitischen Prozess am Beispiel österreichischer Universitäten [Knowledge in numbers: Potentials of Gender Monitoring in the Equality Policy Process Using the Example of

Austrian Universities]. *Gender – Zeitschrift für Geschlecht, Kultur und Gesellschaft*, 13(2), 142-158. doi: 10.3224/gender.v13i2.10

United Nations Economic Commission for Europe (UNECE), & World Bank Institute (2010): *Developing Gender Statistics: A Practical Tool*. Geneva: United Nations. Retrieved 7 November: https://unece.org/statistics/publications/developing-gender-statistics-practical-tool

Van der Haar, Marleen, & Verloo, Mieke (2013): Unpacking the Russian doll: gendered and intersectionalized categories in European gender equality policies. *Politics, Groups, and Identities*, 1(3), 417-432. doi: 10.1080/21565503.2013.816246

Verloo, Mieke (2006): Multiple inequalities, intersectionality and the European Union. *European Journal of Women's Studies*, 13(3), 211–228. doi: 10.1177/1350506806065753

Walby Sylvia, Armstrong, Jo, & Strid, Sofia (2012): Intersectionality: Multiple inequalities in Social Theory. *Sociology*, 46(2), 224–240. doi: 10.1177/0038038511416164

West, Candace, & Fenstermaker, Sarah (1995): Doing Difference. *Gender & Society*, 9(1), 8-37. Retrieved 7 November 2024: http://www.csun.edu/~snk1966/Doing%20Difference.pdf

West, Candace, & Zimmermann, Don H. (1987). Doing Gender. *Gender & Society*, 1(2), 125-151. doi: 10.1177/0891243287001002002

Wroblewski, Angela (2023), *D5.1 First report on monitoring ERA action implementation at national level*, GENDERACTIONplus, retrieved 7 November 2024, https://genderaction.eu/wp-content/uploads/2023/09/GENDERACTIONplus_D5.1_First-report-on-monitoring-ERA-action-implementation-at-national-level.pdf

Wroblewski, Angela (2024), *D5.2 Second report on monitoring ERA action implementation at national level,* GENDERACTIONplus retrieved 25 November 2024, https://genderaction.eu/wp-content/uploads/2024/11/GENDERACTIONplus_D5.2_Second-report-on-monitoring-ERA-Action-5-implementation-at-national-level.pdf

Wroblewski, Angela & Leitner, Andrea (2022): Relevance of Monitoring for a Reflexive Gender Equality Policy. In: Wroblewski, Angela & Palmén, Rachel, (eds.) *Overcoming the Challenge of Structural Change in Research Organisations: A Reflexive Approach to Gender Equality*. Emerald Publishing Limited, pp. 33-52. https://doi.org/10.1108/978-1-80262-119-820221003

Yarbrough, Donald B.; Shula, Lyn M.; Hopson, Rodney K.; Caruthers, Flora A. (2010): *The Program Evaluation Standards: A guide for evaluators and evaluation users* (3rd. ed). Thousand Oaks, CA: Corwin Press.