



# Redesigning Equality and Scientific Excellence Together



## Project Information

Topic:	SwafS-09-2018-2019-2020 Supporting research organisations to implement gender equality plans
Funding Scheme:	EU H2020 - Coordination and Support Action
GA Number:	101006560
Start date:	01/01/2021
Duration in months:	48
Project Coordinator:	UNIVERSITE DE BORDEAUX

RESET aims to address the challenge of Gender Equality in Research Institutions in a diversity perspective, with the objective to design and implement a user-centered, impact-driven and inclusive vision of scientific excellence.

## Consortium partners





Redesigning  
Equality and  
Scientific  
Excellence  
Together

# Monitoring and Evaluation Plan

## Document Information

<b>Title</b>	Monitoring and Evaluation Plan
<b>Deliverable No.</b>	D2.1
<b>Version</b>	1.1
<b>Type</b>	X Report <input type="checkbox"/> Demonstrator <input type="checkbox"/> ORDP <input type="checkbox"/> Ethics <input type="checkbox"/> Other
<b>Work Package</b>	2
<b>Work Package Leader</b>	Sciences Po
<b>Issued by</b>	Sciences Po
<b>Issued date</b>	30.06.2021
<b>Due date</b>	30.06.2021
<b>Dissemination Level</b>	<input type="checkbox"/> Public <input checked="" type="checkbox"/> Confidential <i>only for members of the consortium (including the EC)</i>

### LEGAL NOTICE

The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

### Copyright

© Copyright 2021 The RESET Consortium

Consisting of:

UNIVERSITE DE BORDEAUX  
ARISTOTELIO PANEPHISTIMIO THESSALONIKIS  
UNIwersytet LODZKI  
UNIVERSIDADE DO PORTO  
RUHR-UNIVERSITÄT BOCHUM  
OULUN YLIOPISTO  
FONDATION NATIONALE DES SCIENCES POLITIQUES

*This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the RESET Consortium. In addition, an acknowledgement of the authors of the document and all applicable portions of the copyright notice must be clearly referenced.*

*All rights reserved.*

*This document may change without notice.*



Main Authors	
Name	Organization
Maxime FOREST	Sciences Po

Quality Reviewers	
Name	Organization
Marion PAOLETTI, Maryna RADCHUK	Bordeaux University

## Abbreviations

<b>AUTh</b>	Aristotle University of Thessaloniki (GR)
<b>EC</b>	European Commission
<b>EU</b>	European Union
<b>GA</b>	Grant Agreement
<b>GDPR</b>	GDPR General Data Protection Regulation
<b>GEP</b>	Gender Equality Plan
<b>M&amp;E</b>	Monitoring and Evaluation
<b>RESET</b>	Redesigning Equality and Scientific Excellence Together
<b>RUB</b>	Ruhr University Bochum (DE)
<b>SC</b>	Steering Committee
<b>ScPo</b>	Fondation Nationale des Sciences Politiques (Sciences Po, FR)
<b>UBx</b>	Bordeaux University (FR)
<b>UOULU</b>	Oulu University (FIN)
<b>UL</b>	University of Łódź (PL)
<b>WP</b>	Work Package

## Executive Summary

This *Monitoring and Evaluation Plan* (D.2.1) is submitted by Sciences Po as part of Work Package 2 – Monitoring and Evaluation (M&E). The purpose of WP2 is to design and implement a comprehensive M&E approach and its instruments, enabling GEP implementing partners to self-assess their performance, providing impartial assessment and supporting GEPs' sustainability beyond the lifetime of RESET. The primary objective of this WP is to impartially assess whether the project as a whole and the respective GEPs are delivering the promised activities and changes. This first requires providing RESET partners with a sound framework and relevant capacities for self-evaluation, while impartially monitoring the organization and delivery of the project's activities both at consortium and partners' level. More specifically, this deliverable is the output of Task 2.1 – Design of the M&E plan implemented by Sciences Po over the first six months of the project. It draws upon the notion of "Theory of Change", defining the generic impact pathway of the project, e.g. how the different project's inputs and activities are expected to deliver the intended outputs and outcomes, and under which (pre)conditions. This impact pathway, first introduced to the partners at the first M&E Capacity building session delivered on June 15<sup>th</sup>, 2021, will be further tailored to each partner and updated throughout the project. It is complemented by a set of M&E instruments, including documentary analysis (of deliverables, GEPs or other statutory documents), on-site visits (including focus groups and M&E workshops), bilateral virtual M&E meetings and M&E sessions during project meetings, as well as experimental evaluation. The methodological value, conditions of implementation and potential relevance of each M&E instrument, are thoroughly reviewed. A referential for evaluation capabilities, highlighting the knowledge and skills to be demonstrated for the M&E of the project is annexed to the deliverable. A checklist for GEP design, implementation and sustainability is also provided in annex, to be further tailored to the specificities of RESET – such as the intensive use of co-design methodologies or the focus on intersection inequalities and submitted in an updated form by M12 of the project as D.2.2. The content of the first M&E capacity-building session delivered to the RESET community, is also provided as an addendum to the deliverable.

## Acknowledgement

This deliverable elaborates upon the M&E instruments developed by Sciences Po for the evaluation of the SUPERA project (Horizon 2020, 2018-2022). For the sake of cumulativeness, and with the kind authorization of the SUPERA community, it is based on the state of the art regarding the notion of Theory of Change and the review of M&E instruments developed by Sciences Po for SUPERA. It also includes in annex the checklist for M&E of GEP design and implementation elaborated as part of this project, which will serve as a basis for co-designing the one to be submitted on M12 by RESET. Yet, the structure of the deliverable, its contents, the list of references and the building blocks and timeline of our M&E approach have been substantially updated, based on the experience of their first implementation and to reflect the specifics of RESET. The resources planned by Sciences Po on WP2 are consistent with this cumulative approach, with only a limited number of p/m devoted to the kick-off phase of the project, and resources targeted on knowledge transfer to the coordinator and RESET partners.

## Table of contents

<b>Abbreviations</b>	<b>3</b>
<b>1. Introduction</b>	<b>7</b>
1.1 Grounds for the monitoring and evaluation of RESET	7
1.2 Purpose of the monitoring and evaluation	8
<b>2. Setting the stage for monitoring and evaluating change</b>	<b>10</b>
2.1 Capability framework	10
2.2 Guiding methodological principles	12
<b>3. Developing a theory of structural change for gender equality</b>	<b>17</b>
3.1 Defining the theory of change	17
3.2 Process for elaborating a theory of change under RESET	22
<b>4. Proposed Monitoring &amp; Evaluation plan</b>	<b>24</b>
4.1 Evaluation questions	24
4.2 Methods and tools	28
4.3 Core components and timelines of execution	33
<b>Annex I. EES Evaluation Capabilities Framework</b>	<b>38</b>
<b>The EES framework refers to three key categories of capabilities:</b>	<b>38</b>
<b>Annex II. Checklist for GEP design and implementation</b>	<b>40</b>
Why a checklist?	40
Purpose (and limitations) of checklists for action	40
Purpose of a checklist for GEP implementation as part of RESET	41
Review of available checklists	42
GEAR steps and impact drivers	43
Checklists developed as part of sister projects	46
Proposed checklist for GEP evaluation	48
About resources	49
About stakeholders' inclusion	51
About top leadership (enduring) support	53
About GEP visibility	55
About GEP sustainability	56
<b>REFERENCES</b>	<b>58</b>

## List of figures

Fig. 1 New gender-related criteria under Horizon Europe	15
Fig. 2 Typical presentation of a logic model as a result chain	18
Fig. 3 Basic generic theory of change	19
Fig. 4 Mandatory GEP process requirements under Horizon Europe (2021)	21
Fig.5 Updated impact pathway, UCM, SUPERA project (April 2020)	22
Fig.6 Impact pathway including external variables, (06.2021)	23
Fig.7 Dimensions to be evaluated	25
Fig.8 Gender Mainstreaming cycle	44

## List of tables

Tab. 1 Generic evaluation questions on GEP	24
Tab. 2 SPICED indicators	29
Tab. 3 Typology of indicators	29

# 1. Introduction

## 1.1 Grounds for the monitoring and evaluation of RESET

The monitoring and evaluation approach to be implemented for RESET draws upon program and policy evaluation theory and practice that has developed since the 1970s with view to enhance the quality of projects, programs and policies in terms of design, implementation and outputs. This domain, which has developed at the intersection of different types of knowledges (practice-based vs. theoretical) and disciplines (political science, management science, sociology of organizations, economics, psychology), has raised crucial issues, such as the relevant timing(s) for evaluation (*ex-ante*, continuous and/or *ex-post* ?), the competencies to be mobilized by the evaluators, the participatory dimension of evaluation, but also its very purpose: from better informing the design of programs or policies, through offering a learning potential for on-going actions or supporting better implementation, to measuring their actual impact.

Key methodological issues posed to the evaluation of a project aiming at bringing structural changes for gendering higher education and research will thus be considered, especially with regards to aspects such as a) the use of mixed methodologies combining quantitative and qualitative instruments; b) the timing and overarching purpose of the evaluation design, thought so as to be formative, supportive, responsive and summative and c) the endorsement of capabilities frameworks or standards (as the one developed by the European Evaluation Society). We will also pay specific attention to evaluation approaches and tools developed and implemented in science and innovation to monitor and evaluate gender equality programs and policies, and to the use of “theories of change” which identify the pre-conditions, causal effects and intervening variables for intended changes to be delivered by planned actions.

Our approach is also informed by our own experience of structural change for gender equality. From 2014 to 2017, Sciences Po coordinated the EGERA project, standing for Effective Gender Equality in Research and the Academia, which delivered Gender Equality Plans (GEP) at seven RFOs based in the EU and Turkey. Provided the diversity of institutional and policy contexts, organizational settings, involved disciplines, stakeholders, and windows of opportunity for change, this project offered a rich material for devising our evaluation approach. This approach has been first tested as part of the SUPERA project, standing for Supporting Equality in Research and the Academia, launched in June, 2018 under Horizon 2020. Now in its fourth year of successful implementation, SUPERA has demonstrated the validity of our monitoring and evaluation framework, also pointing out some room for improvement in terms of building capacities for self-monitoring and assessment of delivered activities at implementing partners. It is based on this experience and acknowledging the specificities of RESET, that we opted to annex a first version of the Checklist for GEP monitoring and evaluation, initially due for Month 12, to the present deliverable, so that it can support GEP design.

This Monitoring and Evaluation (M&E) Plan thus relies upon two bodies of knowledge: the one accumulated in the field of program and policy evaluation and the one gained through conducting and evaluating structural change for integrating gender equality and the gender dimension in research organizations. Cumulatively elaborating on the evaluation framework of SUPERA, we

have directed our resources towards further improving this framework and delivering Monitoring and Evaluation support activities to partners from the early stage of the project <sup>1</sup>.

## 1.2 Purpose of the monitoring and evaluation

Project and program evaluation has become commonplace in science and innovation, where not only it constitutes a key step for the validity of many scientific protocols, but also now largely commands the allocation of resources to particular projects. And yet, notably due to this direct relation between evaluation and the distribution of funding, the intrinsic value and purpose of monitoring and evaluation is often neglected or seen as disconnected from project's implementation. It is thus worth reminding that proper M&E are not only a mandatory feature for EU-funded projects or initiative, but also a condition and resource for successful implementation.

The extensive literature on the evaluation of projects, programs and policies, insists on its multi-dimensional character, often distinguishing between *formative* and *summative* evaluation. Formative evaluation refers to an evaluation that starts from the stage of project development, and will continue throughout its life cycle, with the dual objective of a) assessing ongoing project activities and b) providing (through self-assessment and/external monitoring) information which are relevant for improving the project or program (NSF, 2002: 8-10). Hence, by definition, a *formative* evaluation is two-dimensional: it focuses both on implementation (how the project is being conducted, in relation to planned activities, milestones and outputs), and progress towards set objectives, the logics being that the former conditions the latter. Progress evaluation requires to collect data: a) to benchmark participants progress towards set objectives, b) to evidence the factors of success or hindrance for achieving these objectives, and c) to assess the impact of carried out activities. This information will later inform a *summative* evaluation, which purpose is to assess ex-post the performance of a project or program in terms of outcomes, impact and outreach. Unanticipated outcomes (positive or negative) are also important to summative evaluation. Additionally, it shall also offer a learning potential for other (future) projects and programs, possibly outside the realm of implementation of the one evaluated.

While formative evaluation can hardly be distinguished from project implementation, and can involve both insiders (people involved in the daily implementation of project's activities) and outsiders (impartial evaluators), summative evaluation is often entrusted to fully external evaluators, who do not count among projects' implementers and are thus expected to be more objective. Another commonly used typology of evaluation refers to the stage on which it focuses: design, implementation, outcome or impact.

In all cases, the purpose of evaluation is not limited to determining the value and performance of a project or program. It is also intended to contribute to:

- a) Increasing the quality of its **design**, by better informing *ex-ante* the conception and planning of project's activities and the definition of its objectives

---

<sup>1</sup>The first M&E capacity-building session devoted to implementing structural change, was delivered on June 15<sup>th</sup>, 2021 to equip project teams with common background knowledge. This cumulateness is also reflected in the way Sciences Po distributed its resources over the project, intensifying human resources devoted to RESET from M6 of the project.

- b) Increasing the quality of its **implementation**, by providing real-time information about the (un)effective implementation and management of the project, identifying gaps and strategies for adaptation, factors of success or hindrances/resistances`
- c) Increasing the **sustainability** of a project by highlighting the most efficient use of resources, as well as focusing on factors that conditions sustainability, such as a valid ex-ante diagnosis, the collection of relevant data, stakeholders' mobilization or the institutionalization of relevant activities beyond the project's timeframe.

Provided its characteristics and scope, the evaluation of a project as RESET is meant to be both formative (implementation and progress-oriented) *and* summative. Hence, it should provide the consortium, the European Commission and the broader audience interested in bringing forward gender equality strategies in research and the academia, with a robust body of information and data on the project's implementation, impartial recommendations for the further enhancement of the project during its entire lifecycle, and a fair assessment of its overall performance with regards to a) its set objectives and b) the specific conditions for its implementation (in terms for instance of institutional or policy framework, stakeholders' participation or windows of opportunity for change) as those having emerged during the course of the project.

It is now worth highlighting the specificities of *policy* evaluation. First, policies usually pursue higher stakes than projects and programs, such as reducing poverty, increasing equality or improving health *in general*, and provide the framework for projects and programs to be developed and implemented at a domain or issue-specific level. Second, policies are elaborated through a different cycle, different processes and procedures and with the participation of different stakeholders and actors than for projects or programs. As they resort to the domain of public bodies and institutions, policies imply *higher standards* in terms of transparency, accountability, and stakeholders' participation. The purpose of their evaluation is not only to provide information which are relevant for their design, improvement and assessment, but also to contribute to the overall legitimacy of policymaking, by demonstrating that policies are evidence-based, properly monitored and aimed at producing measurable effects (Boussaguet, Jacquot, Ravinet, 2006). One of the added values of policy evaluation, is to pay a great deal of attention to the legal and institutional contexts in which a particular policy is being implemented.

Designing and implementing a GEP within a university can arguably be compared to designing a gender equality policy. Firstly, because GEPs or strategies have long been coined as one of the most popular soft policy instruments in the field of gender equality and non-discrimination on the grounds of sex or gender. Eventually, in some relevant contexts such as Sweden or Spain, two contemporary champions with regards to gender equality policies, plans aimed at fostering gender equality, preventing gender-based violence and implementing gender mainstreaming have paved the way for the adoption of hard law, thus contributing to shape the "policy style" in this area (Alonso and Forest, 2012, Sainsbury and Berqvist, 2009). Secondly, because unlike a research project, GEPs are not meant to bring about specific research outputs nor the adoption of a particular product, but to trigger cultural and organizational changes that involve both structures and individuals, processes and "way of doing things". Thirdly, because the comprehensive and holistic set of actions and measures a GEP is meant to be – as emphasized in the recommended building blocks and areas of action promoted by the European Commission

in the realm of Horizon Europe,<sup>2</sup> is necessarily embedded within the broader framework of existing legislations and policies at the subnational, national and supranational levels. Hence, the specific attention of policy evaluation literature for “path -dependent” features, logics of actors or the cognitive dimension of policy change (See Lombardo and Forest, 2012; Schmidt, 2010; Lombardo, Meier and Verloo, 2009), are also valuable to the monitoring and evaluation of a project aimed at advancing gender equality at universities such as RESET

In this multi-dimensional context, monitoring will be an instrument for project’s implementation and progress evaluation, and for the self-assessment by respective partners of their own activities and performance. As noted by the Equality Challenge Unit (2014), distinguishing monitoring and evaluation can prove challenging, as those occasionally overlap. As suggested by Dahmen and Peterson (GenderTime, 2017a: 4), quoting Funnell and Rogers (2011), “one way to distinguish between monitoring and evaluation is to focus monitoring on what is happening in a project without necessarily explaining why (...), which is instead left to the evaluation process”.

**Monitoring** can be defined as “a continuing function using systematic data collection on specific indicators to provide (...) indications of the extent of progress and achievements of objectives”.

**Evaluation** will be understood as a more comprehensive and in-depth exercise (Espinosa, Bustelo and Velasco, 2016: 5) also bringing a more generic value for other similar projects, programs or policies through analysing factors of success or failures, identifying good practices and addressing their cognitive and interactional dimensions.

Having defined our starting point and the scope and purpose of monitoring and evaluation, we shall now focus on two more practical dimensions: first, the skills and competencies we believe are required for conducting the monitoring and evaluation of a project as RESET and second, the sound principles that should guide this process.

## 2. Setting the stage for monitoring and evaluating change

### 2.1 Capability framework

Due to the variety of definitions of evaluation, and to the even greater variety of actors claiming to perform evaluation work, it has been a growing preoccupation of the broad evaluation community, to set up minimum quality criteria. The emergence of national evaluation associations and more recently, the one of transnational networks aiming at sharing practices and higher standards – such as the European Evaluation Society, have offered venues for collectively building capabilities or competencies framework. Those are not only meant to determine the competencies that an evaluator should ideally master, but also the values,

---

<sup>2</sup>[https://ec.europa.eu/info/sites/default/files/research\\_and\\_innovation/strategy\\_on\\_research\\_and\\_innovation/documents/ec\\_rtd\\_gp-faqs.pdf](https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_gp-faqs.pdf)



interpersonal skills and ethical principles he or she should demonstrate to ensure quality evaluation. Another spur for such an endeavour, is the very context in which an evaluator operates: since evaluation largely conditions the resources, sustainability, credibility and publicity of a project, program or policy, notwithstanding the legitimacy of its promoters and potential well-being of its beneficiaries, it requires especially high standards with regards to the transparency, fairness and impartiality of the evaluation process. From that perspective, *external* evaluation has long been considered a guarantee of impartial evaluation, while quantitative indicators are often thought to bring *actual* evidence of success or failure in achieving the ultimate goals of a project or program. Yet, evidence suggests that quality and ethical criteria need to be defined also for external evaluation and that quantitative indicators are not immune to the same (and other) biases as qualitative ones.

Calling upon evaluators who are not involved in the design and life of the project or program being evaluated, does not mean that they cannot be heavily biased towards its very objective, either for ideological reasons, because they have been involved in the design of an alternative product/policy, or due to methodological issues. Failing to account for gender, age and other grounds of diversity, for instance, is known to introduce potential bias in evaluation, as it can do for any experimental protocol. Hence, drawing upon a fully externalized evaluation does not prevent from questioning the evaluation framework, interests and competencies of the evaluators, and even requires specific criteria ensuring a fair evaluation and an appropriate level of knowledge of the context in which the project, program or policy unfolds.

Similarly, quantitative-based evaluation is *not* a proxy for objective evaluation. As reminded by the US National Science Foundation quoting Sudman (1976), “the most common misconception about sampling is that large samples are the best way of obtaining accurate findings. While it is true that larger samples will reduce sampling error (...), sampling error is the smallest of the three components of error that affect the soundness of sample designs. The two other errors – sample bias (primarily due to the loss of sample units) and response bias (responses or observations that do not reflect “true” behaviour, characteristics or attitudes) – are much more likely to jeopardize the validity of findings” (NSF, 2002: 25). Hence, evaluators are invited to give priority to procedures aimed at controlling those bias, thus pointing out certain competencies and skills.

In sum, there is no guarantee of an objective, methodologically sounded and impartial evaluation without self-reflexivity, peer-learning and a set of carefully identified competencies. So as to frame our evaluation work, we will refer here to the Capabilities Framework elaborated by the European Evaluation Society (See: Annex I). Acting as platforms for cross-disciplinary dialogue and “communities of practice” (Wenger, 1998), a number of national evaluation societies have developed competencies frameworks. Yet, due to its transnational character, we consider that the one developed in a collaborative and cumulative way through two large surveys (2009 and 2011) by the European Evaluation Society, informs about a broader set of criteria. Those do not merely focus on the sole competency and skills of individual evaluators, as those are not the only intervening variable in the quality of evaluation: “evaluation outcomes are affected by the behaviors of other actors (commissioners; other stakeholders, etc.) let alone by the characteristics of the enabling and evaluation governance environment. Judging evaluators capabilities solely through examination of the quality of evaluation products and their results is therefore risky”.

It is thus necessary to “interrogate competencies or capabilities in terms of disciplinary content as well as delivery, social interaction and/or management skills”<sup>3</sup>. As the evaluation team of RESET is based at the French National Foundation of Political Science, it also logically endorses the knowledge and principles of the specific research and action field constituted by *policy* evaluation. Although this field has demonstrated less institutionalization in France than in other contexts, Sciences Po hosts two of the most experienced and recognized French evaluation research teams: the LIEPP and the OFCE, which both justify of embedded gender expertise. Provided the French tradition of policy evaluation, this entails a specific attention for interactional dynamics and the cognitive dimension of policy and institutional change, complemented by competencies in quantitative evaluation drawn from econometrics. The RESET evaluation team adheres to the general principles stated in the Deontological Charter adopted by Sciences Po specifically regarding Policy and project evaluation, to prevent conflicts of interest, ensuring transparency and the integrity and sincerity of the methods used<sup>4</sup>.

## 2.2 Guiding methodological principles

Elaborating upon the EES capabilities framework and the practice of policy evaluation grounded in political science and the gender and policy scholarship, we have adopted the following guiding methodological principles:

### *Formative and summative evaluations*

It is expected that the M&E plan of RESET should include both formative *and* summative evaluations. Priority will nonetheless be given to formative evaluation, aiming at continuously informing the conduct of the project, from its early design stage to its finalization, so as to provide partners with relevant background knowledge to effectively implement RESET. Summative evaluation will be based on the same elements, with view to draw lessons for the sustainability of GEPs and strategies adopted by partner organizations, and for similar projects and initiatives to be carried out elsewhere in the future.

### *Quantitative and qualitative evaluation*

There is a strong case for projects as RESET to adopt a mixed approach combining quantitative and qualitative data collection and analysis. Evaluation practice, notably in the field of science and innovation related projects, advocates for such a combination, which allows for selecting the least imperfect or biased method for each evaluation task and tends to control the imperfections and bias inherent to each of these methods. In the specific case of the evaluation of strategies and plans aimed at increasing gender equality and gender awareness in research and the academia, there are also specific reasons to do so: first, structural change takes time. In other terms, changing universities and other Research Performing Organizations (RPOs) towards the intended objectives of the project, will take far more time than the one allowed by the project’s funding. Hence, ultimate (quantitative) evidence of certain changes (for instance in the cohorts of students opting for some disciplines or of academics achieving senior or decision-making positions) might be hard to collect within the lifetime of the project. This limitation requires to consider other ways to assess the change process, making possible to forecast whether longer-term quantitative changes are likely to take place. Another argument stems from the fact that

<sup>3</sup><https://www.europeanevaluation.org/sites/default/files/surveys/EES%20EVALUATION%20CAPABILITIES%20FRAMEWORK.pdf>

<sup>4</sup> [https://www.sciencespo.fr/liepp/sites/sciencespo.fr.liepp/files/Charte\\_deontologie\\_Liepp-lpp-ldep\\_v1\\_0.pdf](https://www.sciencespo.fr/liepp/sites/sciencespo.fr.liepp/files/Charte_deontologie_Liepp-lpp-ldep_v1_0.pdf)

gender equality is not a standard, neutral area for initiating institutional change. It affects personal beliefs, widely shared cultural references and stereotypes, as well as the distribution of power, knowledge and assets within a particular organization and society (Mergaert and Forest 2015), thus triggering specific resistances (Lombardo and Mergaert, 2016).

Therefore, information and data touching upon or related to these aspects, are not necessarily easy to collect. This task is tightly framed by GDPR provisions at each partner institution, as documented in RESET Data management plan (D9.5), submitted in April, 2021, and can be also complicated by specific legal provisions regulating the collection of data about personal circumstances such as gender, sex, sexual orientation, age or ethnicity. Universities and research organizations being power institutions, collecting data is also about power and making such data public may trigger significant levels of resistance. For these different reasons, it will be crucial: a) to identify a restricted number of valid quantitative indicators, apt to capture key areas of change and to document the achievement of RESET's intermediate and longer terms goals and b) to adopt a broader set of qualitative indicators and qualitative data collection methods, so as to grasp institutional and cognitive dynamics at play, to understand the fundamental features of each organizational context or to analyze how pursued changes "hit" structures and individuals into the intended – or another – direction.

RESET monitoring and evaluation will draw upon a set of data collection methods reviewed in section 4 of this deliverable, including (but not exclusively): focus groups with key categories of stakeholders, interviews with key informants, documentary analysis, smaller and larger surveys. With respect to the use of co-design promoted by RESET, complementary interactive methods inspired from social design will also be used for M&E data collection.

### ***Participatory evaluation***

Monitoring and evaluating RESET will also rely upon stakeholders' participation. This participatory dimension will be twofold: first, the theory of change and the monitoring and implementation plan outlined in this deliverable (see: section 3), will be fine-tuned, and adapted at the level of each implementing partner, to reflect its own GEP priorities, institutional context and implementation timeline. With regards to the theory of change to be deployed following a similar blueprint, it will entail to collectively identify which are the pathways of change to follow, and the main change factors to trigger. Second, for the fine-tuning of quantitative and qualitative indicators and the collection of key comparable data, Sciences Po team will rely upon participatory methods during face-to-face focus groups and M&E capacity-building activities, as showcased during the capacity-building session delivered on June 15<sup>th</sup>, 2021 (see Addendum).

### ***Bias-free evaluation***

Controlling for potential methodological biases with regards to sampling and data collection methods will receive a great deal of attention. This will be especially the case with regards to including all categories of stakeholders (including administrative staff and students) when considering project's activities and preventing gender bias that may be projected on men's and women's attitudes or expectations. Provided the outstanding relevance given to intersecting inequalities in RESET project design, this will also include the attention paid to intersectionality,

as specific groups may be specifically targeted and/or differently affected by implemented actions. Giving a voice to these groups during the monitoring and evaluation work, as well as considering relevant indicators will thus be ensured whenever possible.

### Cumulative evaluation

Evaluating structural change towards gender equality and the integration of the gender dimension in research, cannot only rely upon the generic evaluation literature. It is indeed paramount, to draw upon the rich body of knowledge accumulated on the conduct and evaluation of the dozens of so-called “sister” projects implemented with EU funding<sup>5</sup>. This is especially important due to those projects being implemented in a wide variety of countries, institutions and academic disciplines, embodying the huge diversity of contexts and challenges associated with promoting gender equality and the gender dimension in research content in the European Research Area. Since FP7, funded projects have adopted similarly diverse approaches, addressing an increasingly holistic set of issues, articulating different theories of change and leveraging different methods. In terms of monitoring and evaluation, most include(d) a dedicated work package, usually entrusted to a partner acting as an impartial evaluator: a university, an NGO or a consultancy company, endowed with different M&E capabilities. A few projects opted for internalizing M&E, while calling upon external evaluators to carry out punctual evaluation work.

This cumulative experience has been considered and several evaluation frameworks adopted under EU-funded projects, have informed our approach. This is the case of the M&E of the GENOVATE project. Conducted by the Complutense University of Madrid, it fully considers the inputs from the evaluation literatures, defining a theory of change adapted to the project’s purpose. It also draws upon the insights from the critical analysis of gender mainstreaming implementation in several domains, thus highlighting the specific hindrances and resistances faced by social and institutional change aiming at gender equality, and it focuses on three fundamental dimensions of change: ideas, structures, and people (See: GENOVATE’s *Guidelines for evaluating Gender Equality Action Plans*, 2016). Similarly, we paid attention to the *Monitoring handbook* developed by GenderTime project (2017), which provides insights and recommendations on a usually less theorized element. It emphasizes that “implementation processes in change projects are often complex and multilayered and can at times also be confusing for those involved. Monitoring is useful as it can produce detailed and structured information about what is happening in the change project and how the interventions are going (...). This information, if fed back to the practitioners involved in the change process, can contribute to their increased motivation” (Dahmen and Peterson 2017:5). Along with this baseline argument, the handbook provides clear indications with respect to what should be monitored and when in such a project. A key recommendation is that monitoring should be an integrated and collaborative part of the project, tailor-made to account for specific local contexts.

In terms of cumulativeness, crucial insights have been brought by the implementation of the M&E approach designed by Sciences Po for the SUPERA project. RESET M&E plan thus largely draws upon the concept elaborated for SUPERA, while bringing some innovations to reflect: a) RESET’s focus on co-design methodologies and intersectionality, and b) the lessons learnt from SUPERA Monitoring and Evaluation, notably with regards to supporting partners in building their own M&E capacities and defining their respective project’s impact pathways.

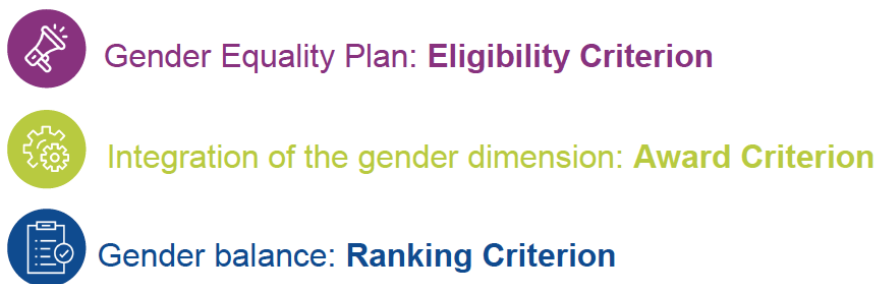
---

<sup>5</sup> According to the Gender Equality Unit of DG Research and Innovation, by May 2021, over 30 “sister projects” had already been funded, involving over 200 universities, research performing and research funding organizations.

### Supportive evaluation

The experience gathered in monitoring and evaluating such projects, also advocates for an evaluation which is not only formative, but also clearly aims at supporting implementing partners in achieving their goals. EU-funded GEP projects are setting up a new, higher standard in dealing with gender imbalances and bias in research and the academia. It is therefore crucial that they result successful and rich in experiences and information for other similar initiatives across the EU, especially as Horizon Europe introduces new eligibility, selection and ranking criteria in relation to submitting institutions to adopt a GEPs, to integrating the gender dimension in the project design or to gender balance in research teams<sup>6</sup>.

**Fig. 1 New gender-related criteria under Horizon Europe**



Under Horizon Europe the **integration of the gender dimension into R&I content is mandatory** unless the topic description explicitly includes a sentence such as the following:

"In this topic the integration of the gender dimension (sex and gender analysis) in research and innovation content is not a mandatory requirement."

Source: DG Research and Innovation, 2021

Besides, while evaluation can occasionally be carried out from a very different set of values and goals than those endorsed by the evaluated project or program, or at least from a perspective primarily dictated by the wise use of public resources, this is not the case here. First, having coordinated and evaluated similar projects, the evaluator fully adheres to the objectives of the project, that it considers also its own and thus intends to use monitoring and evaluation as a tool for success, rather than a way to simply assess the value-for-money of a project. Second, while one of the purposes of this monitoring and evaluation, is to ensure that the project is properly carried out along its planned tasks and milestones, and that its resources are used wisely, our primary commitment will go to achieving its overarching goals and ensuring that lessons learnt are taken on board for this projects and other projects in the future. Supportive evaluation shall also dictate a way to communicate and interact with implementing partners that is conform to this approach.

Supportive evaluation should be distinguished from the role of a mentor, though. This difference is twofold: first, the context of RESET, UOULU and RUB, the two mentor universities of the consortium have been entrusted with a specific contribution to supporting partners implementing their first GEP. This contribution, and more specifically their roles in WP1, relies upon their long-

<sup>6</sup>[https://ec.europa.eu/info/sites/default/files/research\\_and\\_innovation/strategy\\_on\\_research\\_and\\_innovation/documents/ec\\_rtd\\_g\\_ep-fags.pdf](https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_g_ep-fags.pdf) See also the latest EU Council conclusions on the new European Research Area of December, 2020: <https://data.consilium.europa.eu/doc/document/ST-13567-2020-INIT/en/pdf>

standing record in addressing the gender dimension in research. The perspective of Sciences Po will be M&E driven, operationalizing the comparative knowledge built in conducting and evaluating gender equality strategies in research and the academia, and relying more extensively on policy evaluation. Second, for the external evaluator, being supportive does not preclude from adopting the independent, non-implementing perspective to assess partners' performance in delivering intended changes. This assessment will also cover the role of mentors in sharing their experience, helping to set the standard for GEP beginners and enhancing their own policies.

### *Responsive evaluation in convulsed times*

Finally, based on the multi-layered experience detailed above, it is considered key that the evaluation carried out for RESET will be responsive to the project's realities and needs. Responsiveness means here to adapt to potential changes brought to the project's structure, timelines or methods based on practice and lessons learnt, so that monitoring and evaluation remain fully useful to project's implementation and management. Responsive evaluation has proved especially important in the aftermath of the COVID-19 pandemics. The experience gained by Sciences Po in the monitoring of the SUPERA project, will be fully mobilized, as the SUPERA community proved to be especially resilient, swiftly adapting to the new reality and this not only in terms of resources management or project implementation, but also in terms of covered topics, GEP actions implementation or M&E. As the social, financial, and sanitary consequences of the pandemic may also impact RESET, lessons learnt in terms of responsiveness will be of specific value and guide the responses to be provided from WP2. Responsiveness may also entail to address the potential impact of the backlash on gender equality observed in various EU member states on the implementation of strategies for which top leadership support and stakeholders' participation are required. Such circumstances might eventually require the adaptation of the framework theory of change, monitoring & evaluation plan and of the set of tools for self-assessment outlined in section 3 and 4 of this deliverable.



### 3. Developing a theory of structural change for gender equality

As we intend to be cumulative and draw upon the extensive knowledge gained by the scholarships on program and policy evaluation, on the one hand, and the scholarship interested in gender mainstreaming, on the other, we consider of utmost importance to elaborate on a theory of (structural) change for gender equality. It has been indeed one of the key contributions of the literature on program evaluation, to highlight the relevance of establishing evaluation work on the solid grounds of a theory of change adapted to the structure and objectives of the evaluated intervention (program, project or policy alike). In this section, elaborating further on our approach designed for the SUPERA project, and drawing upon the insights from the first M&E capacity-building session delivered to RESET partners during which partners were invited to elaborate their own theory of change, we will first provide a definition of what is intended by a theory of change and build the case for its use in a project such as RESET, emphasizing the specificities of an intervention aiming at producing *structural* changes in participating organizations. Second, we will introduce the process through which implementing partners, supported by WP2 leader, will tailor their own theories of change/impact pathways.

#### 3.1 Defining the theory of change

##### 3.1.1 *This is not about theory!*

An immediate clarification is required: the “Theory of change” is *not* what is usually understood by a theory, in the sense of a “supposition or a system of ideas intended to explain something, especially one based on general principles independent of the thing to be explained” (Oxford Dictionary). Instead, the theory of change we will refer to here, is meant to be a building block for implementation and impact evaluation, which helps defining ex-ante (as in the present case) or ex-post, *how* the changes, effects and impacts intended to be achieved by an intervention are to be (or have been) attained. As such, the theory of change derives from program theory, a long-established area of evaluation literature that can be tracked back to the 1960s. As put by Funnel and Rogers (2011), program theory invites to clarify the theoretical underpinnings of an intervention such as a project, program and policy, by evidencing the logical connection and articulation between inputs and (expected) outcomes and explaining how the intervention is actually meant to work and to produce an impact. Over the past decade, the theory of change has developed as a particular, more applied sub-field of program theory. As different terms have blossomed to describe the “models depicting how interventions are meant to work” (Mayne, 2015: 119-120) such as logic model, results chain, outcome pathway, action or implementation theory, different reviews of program and change theory literatures have been proposed (see for instance Sharpe, 2011; Vogel, 2012; Stein and Valters, 2012 and Mayne, 2015) which will help us to provide valid definitions and build the case for its use in the design of the M&E of RESET.

According to Sharpe (2011: 72), a **theory of change** “consists of a set of statements that describe a particular program, explain why, how, and under what conditions the program effects occur, predict the outcomes of the program, and specify the requirements necessary to bring about the desired program effects”.

Patricia Rogers (2014) indicates that it can be developed for any level of intervention (from a project to a policy) and independently from the fact that “the objectives and activities can be identified and tightly planned beforehand” or change and adapt “in response to emerging issues and to decisions made by partners and other stakeholders” (Rogers, 2014: 1). As pointed out by Mayne and Johnson (2015), theories of change have plenty of possible uses in designing, managing and evaluating interventions. While it has been discussed especially in monitoring and evaluation (See Funnell and Rogers, 2011, part. V), it thus applies to a broader set of stages in the development and implementation of a project, program or policy.

With regards to *when* a theory of change should be developed, it is acknowledged that elaborating it from the design phase, helps better structuring and planning the intervention. However, a theory of change can be usefully elaborated, although under different circumstances and with different inputs and purposes, at every stage of an intervention. For instance, prior to its evaluation or even *ex-post*, to better grasp how the objectives of a particular project have been attained and if not, what went wrong and how? Generally speaking, “the development of a program theory is necessary when hoping to determine why a program is succeeding or failing and if and where program improvement should be focused” (Sharpe, 2011: 72).

In terms of components, a **theory of change** usually refers to:

- a) project/program activities or inputs
- b) the intended outputs or outcomes
- c) the mechanisms through which the intended outcomes are achieved.

### 3.1.2 Why is it crucial for quality evaluation?

Developing a theory of change at some point of an intervention, and even more meaningfully prior to its implementation, appears to be crucial in several ways for its proper monitoring and evaluation. Already at the stage of designing and planning the activities of a project, it can support some sort of *ex-ante* evaluation of the proposed activities, to ensure they will effectively reach their targets and objectives, and help designing a monitoring system – that is, deciding upon which information or data should be collected, when, by whom and for which purpose. It also appears to be crucial for designing meaningful evaluation questions, methods and tools and explaining causal relations between inputs/activities and outputs/outcomes and ultimately, impacts. With regards to the later justification, Mayne (2015) insists on the fact that “to understand how and if an intervention is working, we need to understand how the activities of the intervention are expected to lead to the desired results– both (a) the causal pathway from activities to outputs to a sequence of outcomes to impacts and (b) the causal assumptions showing why and under what conditions the various links in the causal pathway are expected to work” (Mayne, 2015: 121). To put it clearly, a logic model as the generic one below does not by itself constitutes a theory of change. But it specifies “the causal mechanisms (that) transforms a logic model into a theory of change.” Chen (2015).

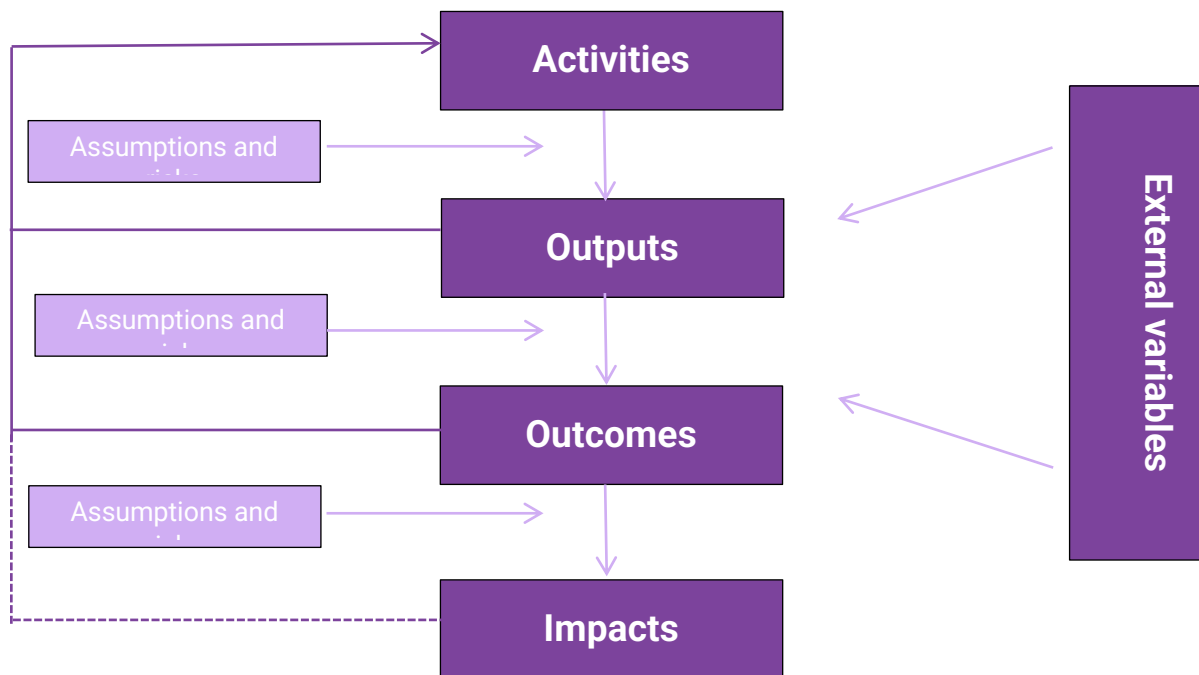
**Fig. 2 Typical presentation of a logic model as a result chain**





Another important term thus consists in the “assumptions” made about how the change process happens. These assumptions aim at “making explicit the interpretations of how change might happen” and are meant to be relevant to the context of implementation. They might consist in considerations about the beliefs or rationales of the stakeholders, hypothetical cause-effect relations or contextual organizational feature that might trigger intended changes. As pointed out by Mayne (2014: 128), project’s or program’s activities are rarely the sole cause of a result. Hence, by incorporating assumptions about intended or observed changes, as well as key external variables, a theory of change offers a broader understanding of the process at play.

**Fig. 3 Basic generic theory of change**



**Source:** Based on United Nations Children’s Fund, Supplementary Program Note on the Theory of Change, Peer Review Group meeting, 11 March 2014, UNICEF, New York, 2014, p. 4. [www.unicef.org/about/execboard/files/PRG-overview\\_10Mar2014.pdf](http://www.unicef.org/about/execboard/files/PRG-overview_10Mar2014.pdf).

Along with these aspects, which are the basic justification for elaborating a theory of change and grounding monitoring and evaluation into that theory, there are other reasons for taking such a step. In particular, especially if its design is collaborative, involving the main stakeholders interested to the implementation and success of the intervention, a theory of change helps building a shared understanding of the main components and objectives of a project, and of the way it should ideally work to produce its full impact. This is especially relevant for a project of such a complex nature and ambition as RESET. Indeed, partners are likely to demonstrate different degrees of understanding of the project at a first stage, depending on their previous experience of structural change and degree of gender knowledge. Additionally, due to the specificities of the call, a project such as RESET deploys two levels of implementation that only partially overlap: a project level, structured by work packages, deliverables and project’s milestones and a GEP level, structured by GEP activities and their respective timelines and monitoring mechanisms. Hence, collectively building a shared understanding of what is intended in terms of changes, of the mechanisms through which it should happen and of the impact it

should eventually achieve beyond the mere, static description of project's activities and goals, can reveal key for the success of RESET. This shared understanding, translated into a common presentation of how the project is intended to work, deliver its activities and produce impact, can also reveal an important support for project's communication, both internally (to stakeholders and beneficiaries) and externally (towards funders, peer organizations, policy makers or the general public), thus contributing by itself to a successful implementation.

### 3.1.3 What is specific about structural change for gender equality?

Now that we have illustrated why it is valuable to elaborate a theory of change to ground our monitoring and evaluation approach, it is important to highlight the specificities of implementing a project aimed at triggering structural changes to foster gender equality and integrate the gender dimension in research. While a project consisting in testing a new service or product will surely also entail behavioral and possibly organizational changes, those are likely to be only either facilitating factors or side-effects with regards to the core objective of the intervention. Instead, a project such as RESET, *primarily* aims at producing organizational and behavioral changes. Moreover, it has been widely noticed in the literature on gender mainstreaming, that those behavioral and organizational changes required by advancing gender equality are not neutral, but deeply affect social norms, organizational patterns, power distribution and cultural values (Lombardo, Meier and Verloo 2009). Evidencing gender biases can thus produce specific resistances which a process of change should also aim to address.

Hence, the following statement made by Vogel (2012: 5) about the assumptions elaborated in a theory of change acquires an even greater validity in the context of a structural change project as RESET: "The quality of a theory of change process rests on 'making assumptions explicit' and making strategic thinking realistic and transparent. Practical experience highlights that this is not straightforward to do, as these tap into deeper beliefs, values, worldviews, operational 'rules of thumb' and analytical lenses that all individuals in development bring to their work. It takes time and dialogue to be able to challenge assumptions. Power relations, both in the program's context and within organizations, limit the ability to challenge established ways of working". The latter dimension - power relations -, is of crucial relevance in the case of structural changes aimed to advance gender equality, as planned activities should be partly directed to changing heavily gendered power relations within the organization. In the case of research and academic institutions, power relations are blurred by notions of merit, excellence and scientific objectivity, which makes necessary to deconstruct each of these notions to unravel the gender biases they contain (Marx Ferree and Zippel 2015, Van den Brink 2012), which strongly affect both the position of women within research and higher education organizations and the status of gender sensitive knowledge across disciplinary fields (Albenga, 2016). It remains that these organizations are structured by a number of statutory and symbolic hierarchies that need to be disentangled if ones intend to produce *real* organizational and behavioral changes. This aspect is of specific relevance to RESET, which overarching goal is precisely to redesign the notion of scientific excellence in the prospect of greater inclusiveness, as tackled under WP6.

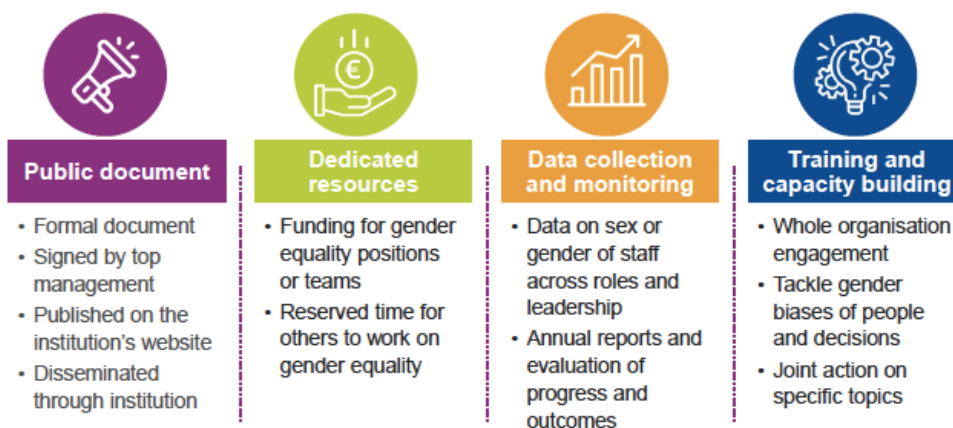
The now abundant literature on gendered institutions and how to mainstream gender equality in organizations, offer fundamental insights to put research and higher institutions in motion. Those point out the importance of addressing not only individuals, but also – and primarily – processes,

procedures, institutional settings, written and unwritten rules and established ways of doing things (EIGE 2013, Jacquot 2010). Another key input gained from these scholarships, is that changing institutions is fundamentally a *political process*, which requires a clear strategic framing at every step, evidenced-based arguments, the involvement of various stakeholders and occasionally to engage in power relations and institutional politics so as to secure allies, support and resources. A change theory adapted to the purpose of RESET, should thus encapsulate resources to better grasp institutional contexts, actors' dynamics and key necessary steps to be taken and resources to be secured to produce expected outcomes and impact.

Along with these generic specificities of gender mainstreaming and structural change implementation, EU-funded projects as RESET present fully specific features inherent to the way they have been devised and they are being assessed by the European Commission since the EU 7<sup>th</sup> Framework Programme, which also deserve to be considered. Sciences Po identified a major challenge in monitoring and evaluating a project of this nature: although a Coordination and Support Action, these projects are assessed by the European Commission and its mandated external reviewers, through the same lens as the one applied to research-oriented projects. This applies to the typology of knowledge to be mobilized by partners, and to the typology of criteria applied to assess project's performance and outreach (Forest and Mergaert 2015).

Another ambivalence results from the fact that these projects are expected to lead involved universities, research performing and funding organizations to set up and implement GEPs. Therefore, a dichotomy usually exists, between the management and implementation structure of projects as such, and the management and implementation structure of the GEPs, which are core to the execution of these projects. This was evidenced in a comparative study commissioned by the European Institute of Gender Equality in 2016 as a preliminary step to designing GEAR online tool (EIGE 2015a, 2015b). It was also found that although setting a new quality standard for the implementation of gender equality strategy in RFOs and RPOs, EU-funded structural change projects had not yet constituted a full guarantee for GEPs to be endorsed and fully enforced in participating institutions. The new requirements introduced in latest calls under which GEP projects are funded, as well as the standards promoted by the EC for GEP content and process under new Horizon Europe requirements, can be read in light of those findings.

**Fig. 4 Mandatory GEP process requirements under Horizon Europe (2021)**



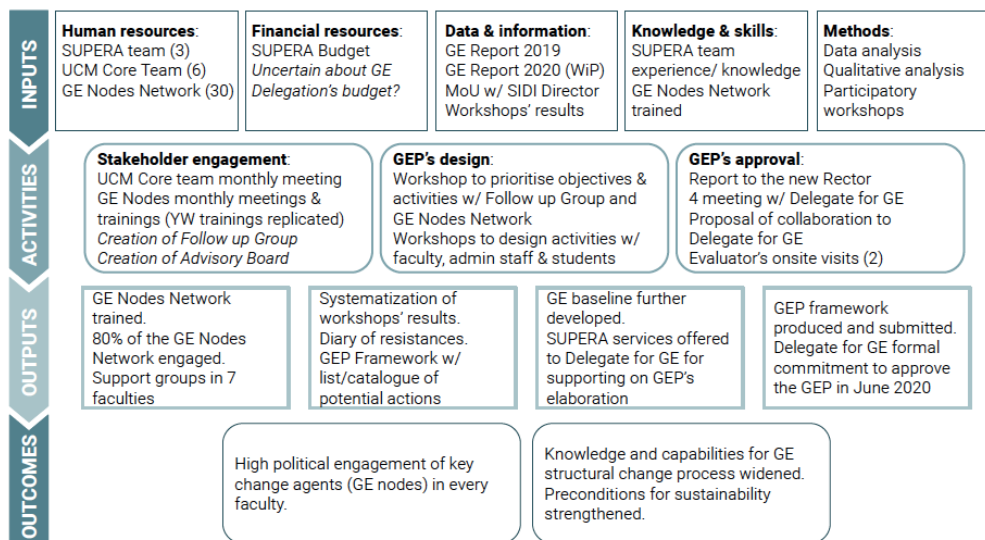
To prevent this dichotomy already from the project design phase, Sciences Po made recommendations to ensure that RESET project’s milestones and deliverables would be framed as key elements ensuring the timely and sustainable execution of the project and the wise use of its resources. This intention is explicit in the WP structure of RESET, where WP1 to WP7 are directly aimed at supporting GEP design, implementation, and monitoring. The use – and appropriation by the partners, of the notion of theory of change, also responds to this necessity, as it encourages GEP implementing institutions always to prioritize the overarching goals of the project when delivering activities and planning actions.

A few EU-funded GEP projects have already demonstrated the usefulness of a theory of change: GENOVATE’s *Guidelines for Evaluating Gender Structural Change* (Espinosa, Bustelo and Velasco, 2016), based on the continuous refinement of the project specific theory of change, provide useful inspiration and recommendations in relation to designing an evaluation approach, which pay a great deal of attention to the specificities of a “gender structural change” meant to transform institutions. Based on Tichy’s framework reinterpreted by Navarro (2007) and Mukhopadhyay et al. (2006), they embrace three key areas of change: ideas, structures, and people, underlying that institutional culture, organizational structure and

### 3.2. Process for elaborating a theory of change under RESET

First applied to the SUPERA project, our M&E approach elaborating upon the notion of theory of change, has demonstrated its value for GEP implementing partners to design their pathways to change, considering their own organizational and domestic contexts. First mobilized during a Capacity-building session delivered 12 months after the project started, it was updated towards the mid of the project, to reflect changing conditions for the implementation of structural change. It was further assessed at partners’ level as part of the Interim Monitoring Report of the project.

**Fig.5 Updated impact pathway, University Complutense of Madrid, SUPERA project (April 2020)**

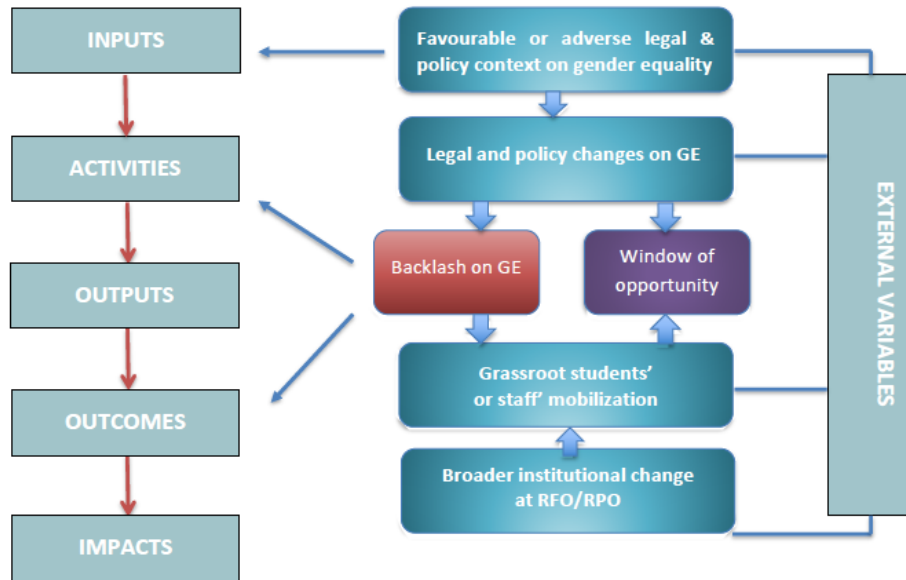


Source: courtesy of SUPERA UCM team

But this process also evidenced the need for GEP implementing organizations, to build capacities for self-monitoring and evaluation as early as possible in a project. Therefore, for RESET, a **first**

**step** consisted in introducing the notion of theory of change prior to the submission of this deliverable, on the occasion of a capacity-building session targeted at project teams and Gender Equality Boards' members on June 15<sup>th</sup>, 2021.

**Fig. 6 Impact pathway including external variables, presented to RESET partners (06.2021)**



This deliverable thus constitutes a **second step** in the process of developing and appropriate theory of change to support our monitoring and evaluation work. In order to build a valid theory of change adapted to the context(s) of implementation and objectives of a project, a participatory process is required. Partners will thus be invited as a **third step**, to elaborate from the components outlined in this document and the first insights gained through the capacity building session, focusing on the following issues:

- ➔ The data and information required to draw upon a robust diagnosis of gender biases and inequalities at play
- ➔ The key aspects of the functioning of their organization (how decisions are being actually taken, who are the key stakeholders, where are the potential bottlenecks or accelerating factors?)
- ➔ What are their specific priorities, as they were put during the proposal phase or as they emerged from the extended diagnosis phase after the project's start?
- ➔ Their assumptions about how the change should happen in their own institution, provided points 1) and 2), and potential risks or facilitating factors

This work will be coordinated by Sciences Po through bilateral online consultations over the months of September to November, to deliver refined, partner-specific theories of change based on a similar blueprint provided in section 3.3. **A fourth step** will consist in the collective validation of these partner-specific theories of change, summarizing key assumptions and pathways of change on the occasion of a RESET project meeting. A dedicated timeslot will be included to the program for that purpose.

The objective of this iterative process is to ground partners' impact pathways into the knowledge of their own organizational settings to support (self-)monitoring and evaluation work. They will be tested over the first monitoring and evaluation round of RESET (M6-M24), thus informing the interim monitoring report submitted on M24 (December 2022). The latter deliverable will include an updated version of the theory of change of RESET, reflecting the pathways of change, external variables and potential resistances faced over the first half of project's implementation, thus informing possible correction measures until its completion. Partners' updated impact pathways will also be mobilized for the Final Evaluation report due on M48 (December 2024), supporting the analysis of how RESET actually delivered its outcomes and impact, identifying unintended (positive or negative) effects, strategies of adaptation as well as the most effective strategic framings used during the project. Along with continuous project monitoring, these two rounds of evaluation are intended to let GEP implementing partners enough time to a) design their GEPs and put them in motion, and b) implement a first set of GEP actions even if those are meant to continue after project ends.

## 4. Proposed Monitoring & Evaluation plan

### 4.1. Evaluation questions

Before attempting to formulate core evaluation questions in relation to the above-sketched change theory, it is useful to introduce their role and how they should be ideally designed and which dimensions they are meant to address. Again, GENOVATE's *Guidelines for evaluating gender equality action plans* provide relevant insights to that respect. At a general level, when it comes to evaluating the design of a GEP, evaluation questions should address how the plan tackles the different areas and dimensions of structural changes for gender equality. The evaluation of implementation should be guided by questions concerning processes of coordination, (human and financial resources') management, stakeholders' participation or communication. The evaluation of results should concentrate on the extent to which the GEP is contributing/has contributed to the changes in the different areas it covers. These basic questions, as formulated below, can be considered generic and only constitute a starting point.

**Tab. 1 Generic evaluation questions on GEP**

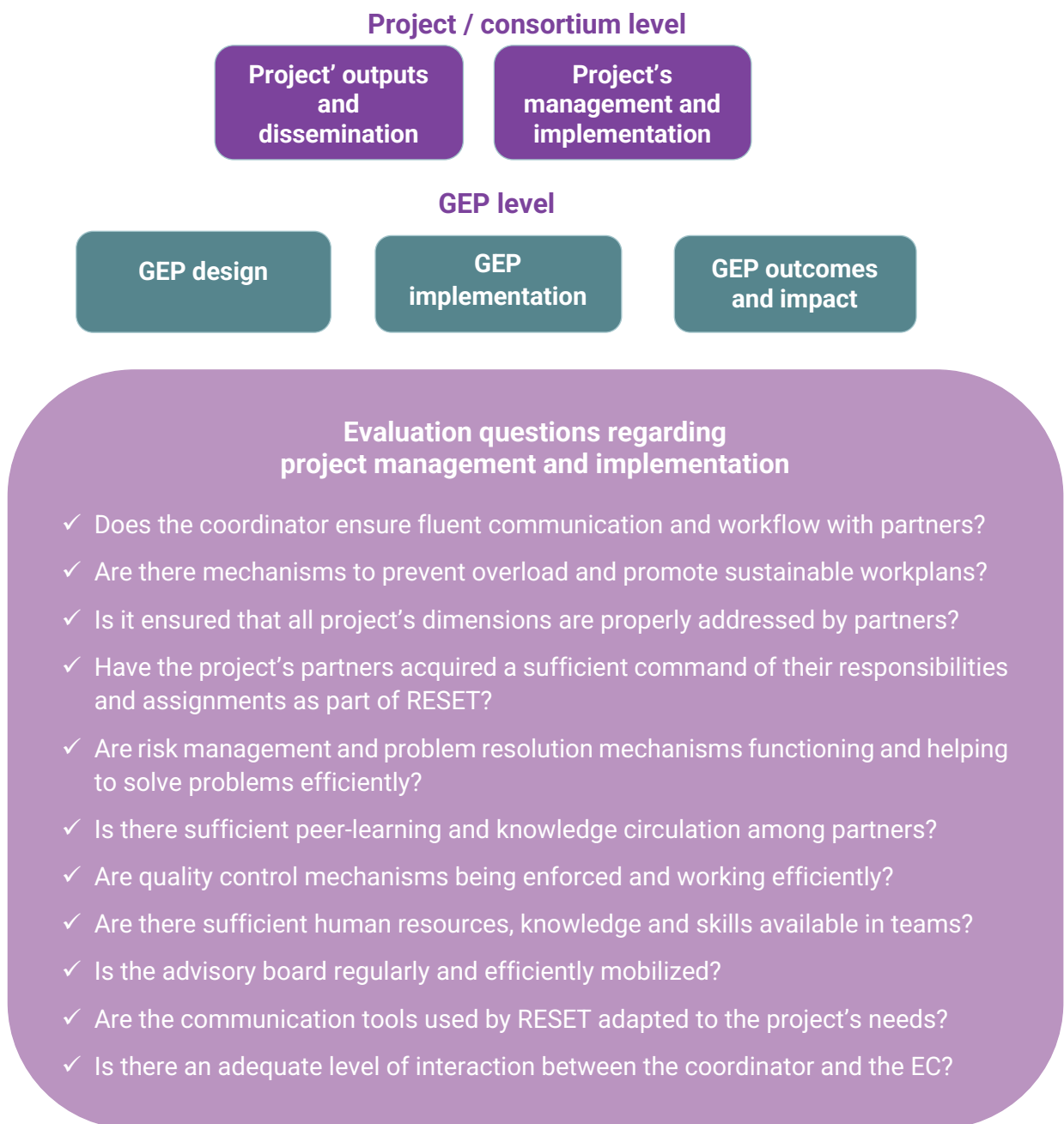
Evaluation focus	Evaluation question
Design	To what extent does the GEP include specific actions related to the areas and dimensions of gender change?
Implementation	To what extent are GEP processes (coordination, management, communication, knowledge transfer, etc.) facilitating gender change in these specific areas and dimensions?
Results	To what extent is the GEP contributing to gender change in these areas and dimensions?

Source: Evaluation focus and "inspiring" evaluation questions, in Espinosa, Bustelo, Velasco (2017: 25)



Evaluation questions need to be further contextualized, in accordance with the theory of change developed by the intervention, its context(s) of implementation but also the different areas of structural change as those identified by Navarro (2007) and Mukhopadhyay et al. (2006): Ideas, Structures and People. Fundamentally, their function is to determine what needs to be evaluated, and to trigger answers which contribute to explaining how change eventually happens. Since our monitoring and evaluation plan should integrate the two levels of implementation of the RESET project, that is the project/consortium level and the GEP level, our guiding evaluation questions should not solely focus on the latter but address these two dimensions. It should be reminded that the evaluation questions suggested hereafter should be only understood as a basis for further discussion and refinement, in line with the theories of change to be developed at partners' level on the basis of the one outlined in this deliverable.

**Fig. 7 Dimensions to be evaluated**



### Evaluation questions regarding GEP co-design

- ✓ Is the design of the GEP based on an in-depth organizational gender analysis? (the mandatory in-depth analysis draws on sufficient data and knowledge)
- ✓ Is the GEP based on a clear, context specific strategic framing of issues at stake?
- ✓ Does the GEP design involve the contribution or consultation of key stakeholders?
- ✓ Is the GEP articulated with existing gender equality strategies or bodies?
- ✓ Is the (co-)design process sufficiently transparent? (information is communicated to relevant categories of stakeholders, such as social partners, student's organizations and decision-making bodies of the organization)
- ✓ Does the GEP adequately cover the four areas of actions?
- ✓ Is it sufficiently holistic not only in terms of issue coverage, but also through addressing people *and* structures?
- ✓ Does the GEP consider intersecting inequalities and disadvantaged groups?
- ✓ Are the tasks' responsibilities formulated and clearly ascribed?

### Evaluation questions regarding GEP implementation

- ✓ Does the GEP implementation relies upon sufficient stakeholders' mobilization?
- ✓ Has the GEP been endorsed officially at the appropriate level of responsibility?
- ✓ Is it backed at the highest level of the organization?
- ✓ Has the GEP been made public and disseminated and do people know about it?
- ✓ Are the foreseen enforcement mechanisms working efficiently?
- ✓ Are the Gender Equality Boards and Co-design workshops serving their intended purposes?
- ✓ Are the targeted audiences being engaged with?
- ✓ Is it responsive to encountered resistances and windows of opportunity?
- ✓ Is the implementation process as holistic as intended in the GEP design?
- ✓ Does the GEP strive towards the sustainability of the most structural actions?
- ✓ Are there any specific external variable(s) serving or constraining GEP's implementation and how are they addressed?



### Evaluation questions regarding project's outputs and dissemination

- ✓ Does RESET actively communicate with sister projects?
- ✓ Does RESET engage with policy stakeholders at EU, national, sub-national levels?
- ✓ Does the project produce outputs – such as tools, guidelines or recommendations – which are relevant to other projects or interventions?
- ✓ Is there a proper communication and dissemination plan and is it holistic in terms of audiences and dissemination channels?
- ✓ Are partner institutions communicating efficiently about RESET outside their organizations, towards other RPOs and RFOs, professional associations, equality mechanisms, companies involved in research and innovation and other categories of stakeholders?
- ✓ Did RESET established contact and cooperation with non-EU organizations and does it have an international outreach?
- ✓ Have any of RESET's outputs been uptaken by other organizations?

### Evaluation questions regarding results and impact

- ✓ Did the GEP contribute to integrate GE in strategic documents of the organization?
- ✓ Did the GEP contribute to mainstream GE objectives in decision-making?
- ✓ Did it contribute to mainstream GE objectives in recruitment/career management?
- ✓ Did the GEP contribute to mainstream gender-sensitive data collection?
- ✓ Has the GEP led to measurable changes in decision-making functions and bodies?
- ✓ Has the GEP led to greater gender balance in senior academic positions, especially in those fields where imbalanced were particularly salient?
- ✓ Has the GEP contributed to trigger attitudinal change towards gender equality?
- ✓ Did the GEP contribute to integrate the gender dimension in research and teaching?
- ✓ Has the GEP contributed to mainstream gender knowledge across disciplines?
- ✓ Did it support the institutionalization of gender equality policies?
- ✓ Are there specific actions of the GEP that have been uptaken elsewhere?

## 4.2 Methods and tools

Prior to establishing the core components of our Monitoring and Evaluation plan and the timelines and milestones for their implementation, we will provide below a brief critical review of the main methodological instruments available to our evaluation approach, in line with the principles stated in section 1 of this deliverable. This review will help us to consolidate our methodology and select the most appropriate evaluation instruments to be further combined so as to test our theory of change through answering the evaluation questions outlined above. Obviously, it is not our purpose to carry out a comprehensive review, readily available in the extensive literature on the methodology of program evaluation (see, among others, Prosavac, 2015 for the latest edition) and complemented by field specific methodological reviews and handbooks for policy or experimental evaluation.

### Indicators

Indicators – both qualitative and quantitative – do not constitute a *method* of monitoring and evaluation, but a basic *instrument* inherent to M&E work. They can be informed by a variety of techniques of data collection as those outlined below. It is yet important to carefully design the set of indicators to be used, as those should help answering the evaluation questions posed at the start of the intervention. Indicators are first required for monitoring the planning phase, with view to check that all required components of the project are in place – for instance, that GEP design appropriately cover the four areas of action of RESET.

Performance indicators will be further used to monitor project and GEP implementation:

- ➔ Quantitative indicators related to implementation will focus on volumes, such as the number of trainees in a session or participants in a co-design workshop, the overall number of activities carried out in respective areas of action, the level of interaction on an online platform or the number of people reached in each stakeholders' categories.
- ➔ Qualitative indicators will particularly target how activities are being carried out in terms of timing, preparation, communication, stakeholders' involvement and with regards to their very objective. Qualitative indicators for implementation should inform whether the activities carried out as part of the project are reflecting its four core principles intended to tackle the main barriers of implementing GEPs in research organizations: cumulateness; innovation; inclusiveness and sustainability. Qualitative indicators are also meant to focus on people's experiences and perceptions. They are particularly important to measure attitudinal and organizational resistances or support to the project's objectives.

Provided the nature of RESET, performance indicators on project's outcomes and impact, should primarily aim at measuring institutional change, for which both quantitative and qualitative indicators are required. As stated in section 2, due to the specificities of structural change and the fact that it usually takes more time than the one allowed by project's funding, quantitative indicators should be used and selected wisely, so that they can capture key areas of change and document the achievement of RESET's intermediate and longer terms goals, and a broader set of

qualitative indicators and data collection methods should be adopted, to understand how change happens, and how far it goes.

Several typologies have been developed, to help projects' stakeholders and evaluators to design meaningful indicators. Those relates both to the purpose of indicators, and to the process for their design and implementation, as shown on Tables 2 and 3 below.

**Tab. 2. SPICED indicators**

SPICED (when using indicators)
<b>Participatory:</b> indicators should be developed and used together with end beneficiaries (citizens)
<b>Interpreted and communicable:</b> indicators need to be explained or interpreted to different stakeholders
<b>Cross-checked and compared:</b> to increase the validity of indicators
<b>Empowering:</b> the process should allow stakeholders to reflect critically on their changing situation
<b>Disaggregated:</b> indicators enable break downs for different groups, such as gender, income, etc/

**Tab. 3. Typology of indicators**

<b>Outputs indicators</b> help you to monitor whether you are doing what you planned (outputs). For example, number of capacity building training organised, number of participants, etc. However, they do not give you an idea of the effects brought about by these outputs.	<b>Impact Indicators</b> are indicators that measure the long-term impact of a project. For example, the increase proportion of women in decision-making positions, or the integration of a Gender+ approach considering other inequality factors in career management policies.
<b>Outcome Indicators</b> measure the medium impacts of a project. For example, number of participants to a capacity building training seminar that are applying their new knowledge	<b>Process indicators</b> are those indicators that are used to measure project processes or activities. For example, the number of people appointed in different department as gender focal persons.

### Documentary analysis

Document analysis can be used at any stage of the evaluation, but appears to be especially useful at an early stage, when the evaluator attempts to grasp basic contextual features of the intervention and at a later stage, when the intervention itself has led to the adoption or enactment of specific framework documents (an action plan, a strategy, a charter...) or when it has led to the review of pre-existing documents (a mission statement, a pluriannual strategy, an organization's statute, a code of ethics...). According to Caulley (1983), the primary aim of a document analysis is to gather facts. Yet, as pointed out by Bowen (2009), analysed documents are interpreted by the researcher to give voice and meaning around an assessment topic, which makes it an especially valid instrument for policy analysis and evaluation.

Documentary analysis involves the selection of facts from documents, which can be done with different degrees of sophistication, from formulating broad-based questions and hypotheses, to applying much more complex selection devices as coding. Analysing documents then incorporates coding content into themes similar to how focus group or interview transcripts are analysed (Bowen,2009). In the field of gender policy analysis, the most sophisticated methodological framework developed for document analysis is Critical Frame Analysis (CFA) designed for the comparative QUING project. Meant to be both analytical and evaluative, the project required a discursive approach, "to map and explain the similarities, differences and inconsistencies in gender+ equality policies in Europe, and second, to assess the quality of these policies". The starting point of the research was that gender equality is a contested concept that can be filled with a variety of meanings that can be unfold through a systematic and critical frame analysis, applying a specific coding to analysed documents (Lombardo, Meier and Verloo 2009).

While this methodology is resource-intensive, it can yet inform, through a delimited use, the understanding of discursive dynamics at play around GEP implementation<sup>7</sup>.

### *In-depth individual interviews with key stakeholders/informants*

In-depth interviews are a qualitative research method indicated to explore the perspectives of a small number of key informants or stakeholders about a particular program, project or situation. They are useful to provide in-depth, first-hand information about the context of an intervention, contributing to a finer-grained picture. The information gathered through individual interview is much more detailed than what can be gathered through other data collection methods such as surveys or even focus groups. Yet, as pointed out for instance by Boyce and Neale (2006), in-depth individual interviews are prone to bias. Those can be of different nature, reflecting for instance the involvement or position of the respondent in the evaluated intervention, personal values or a hierarchy between different categories of knowledge. Albenga (2016) highlighted that the gender knowledge demonstrated by researchers carrying out interviews as part of structural change projects, was often dismissed by respondents granted with a different scholarly capital. It is thus important to control for such bias through carefully designing the framework and questionnaire of the interview, as well as through a control-check of the information received via alternative sources such as policy documents, databases or surveys. Besides, interviews are necessarily time and resource intensive, as they require from the evaluator specific skills to conduct, transcribe and analyse each interview, possibly onsite, without providing sufficient grounds for generalizations due to very limited samples.

### *Focus groups*

Powell et al. define a focus group as “a group of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research” (1996: 499). Yet, the emphasis is not laid on answers to questions posed by the researcher(s), but on interaction within the group, based on topics prompted by the researcher. It is thus distinctive of focus groups, to provide data and insights gained through participants’ interactions. To achieve their goals, focus groups require careful preparation, including the adequate selection of participants in terms of knowledge, diversity of perspectives, gender balance and relation to the program or intervention. Predetermined research questions should be explored, following a tailored interview guide so as to trigger an appropriate level of interaction directly relevant to the evaluation. More specifically, Gibbs (1997) underlines that “the main purpose of focus group research is to draw upon respondents’ attitudes, feelings, beliefs, experiences and reactions in a way in which would not be feasible using other methods, for example observation, one-to-one interviewing, or questionnaire surveys (...). Compared to individual interviews, which aim to obtain individual attitudes, beliefs and feelings, focus groups elicit a multiplicity of views and emotional processes within a group context”. Focus groups therefore require specific skills from the evaluator, grounded in social research, and careful preparation. Within the frame of RESET, they can be held during onsite visits at implementing RFOs and RPOs, as well as during project meetings with core teams and/or advisory committee members, with view to document for instance encountered resistances and difficulties during GEPs’ and project’s implementation, as well as strategic framing, the degree of knowledge about

---

<sup>7</sup> Sciences Po is experienced in applying CFA. For its application to a structural change project, see: Albenga, 2016.

a GEP among specific categories of stakeholders, etc. Their result, however, can hardly be generalized as it results somehow from an artificially created situation of interaction (Espinosa, Bustelo and Velasco, 2017: 45)

### *Participatory workshops*

Along with focus group, it exists a variety of other participatory formats such as world café, workshops based on social and interaction design or Group Model Building (GMB) that can also be mobilized for the purpose of program or policy evaluation. Also relying on group dynamics, they do not necessarily require to be led by researchers and often mobilize non-academic participatory techniques. Workshops based on social design, for instance, are derived from the sequences and principles applied to product and service design. Their objective is often to help diagnosing a situation of departure (a service, a policy) which is not working well, and to design collectively potential solutions and improvements from a user-centered, experience-based perspective. These techniques are increasingly used in policy design and evaluation, as they are thought to be efficient, offering the possibility to describe and analyse the initial situation, to co-design potential solutions, to prototype and pilot-test it over a relatively short period of time and without investing much resources (see: Forest, Bastard and Denis, 2015). Although their use for program evaluation is not consolidated yet, they could nonetheless be punctually mobilized for RESET, due to the expertise shared by the evaluator and WP4 leader in mobilizing such techniques. Other techniques serve more specifically the purpose of reaching a common consensus or understanding of differences, as Group Model Building. Its application to gender equality interventions has been theoretically and methodologically grounded by Bleijenbergh, Benschop and Vennix (2008) and tested within the frame of the STAGES, GARCIA and EGERA projects. Based on system dynamics, GMB is presented as especially applicable to “messy” problem situations for the resolution of which the creation of consensus and commitment is crucial. GMB can be a valuable, but demanding tool as part of responsive program evaluation.

### *Surveys*

Large surveys are the most widely used quantitative-based method in program evaluation. Surveys enable to collect statistically significant data about key aspects of the evaluated intervention. They can be used as a simple monitoring instruments, to document the participation of stakeholders to program’s activities, their knowledge about the intervention, or to measure attitudinal change over time. They can also serve to document and analyse the situation of departure with regards to gender bias and imbalances in a given organization (diagnoses) or the positions of relevant categories of stakeholders with regards to potential solutions (prognoses). They are usually resource intensive as they require the participation of several competent researchers for their design, implementation and analysis, the access to a large survey sample and to appropriate communication channels, as well as a specific attention to a number of potential bias likely to undermine the robustness of the survey, hence its usefulness for monitoring and evaluation. As they are involving a high number of respondents and processing a considerable amount of data, surveys are also entailing issues with regards to data protection. This is notably the case for surveys covering issues and data related to sex and gender identity, inequalities and discrimination. For these reasons, their use for the purpose of M&E should be targeted and delimited, their design should involve several partners and different categories of stakeholders so as to anticipate potential bias and implementation-related problems. Ethical and data protection issues should be also properly addressed at every stage of the design and

implementation processes. In the case of structural change projects, surveys were for instance used in EGERA, to monitor the patterns of gender equality culture in partner institutions, with significant differences in terms of outreach in respective partner organizations.

### **Evaluation questionnaires**

Any evaluation work usually entails the use of questionnaires to monitor and evaluate specific activities such as project's meetings, interactive workshops, conferences, trainings or awareness-raising sessions. Questionnaires can be submitted to participants ex-ante, to check their degree of knowledge and familiarity with the subject and/or the methodology, their readiness to engage in the interaction, their expectations and potential fears, etc.; ex-post, they will serve to measure their satisfaction or discontent with the proposed concepts and methods, the degree and quality of the interaction, as well as to provide first evidences of attitudinal change and insights for further improvements. It should be noticed that their results are often used to produce data pretending to some sort of statistical significance it does not have, which can reveal misleading. Evaluation questionnaires should instead be considered as a mere control-check and informative instruments, serving as a basis to elaborate more thorough evaluation instruments as individual interviews or focus groups.

### **Experimental evaluation**

Sciences Po evaluation team is hosted at the OFCE, a research unit primarily devoted to research in economics, with a strong specialization on policy evaluation in the fields of social cohesion, employment, tax and environmental policies. This background informs us about the evolutions in the field, where experimental economics has gained increasing relevance since the early 2000s, to the point that it can nowadays be considered dominant among economists. Based on a critique of both economic models based on pure theoretical speculations, and of economic analysis of *existing* policies or market structures and dynamics, it postulates that any intervention in the functioning of world's economies should first pass the test of scientific experimentation. It thus consists in the use of controlled, scientifically designed experiments to test theoretical predictions regarding the behaviors of economic agents. Applied to policy or program evaluation, experiments can consist in measuring the impact of an intervention by controlling other intervening variables through a valid counterfactual (for instance, a similar group or unit to which the intervention has not been applied), to assess the *actual* outcomes of the program or policy (for applications to gender, see for instance Niederle and Vesterlund, 2007; Azmat and Petrongolo, 2014, Nelson, 2015). Although it can provide key insights for understanding structural change, experimental evaluation derived from experimental economics is not immune to gender and other bias. It is also criticized for "inventing" allegedly "pure" laboratory-style experiment of which findings are hardly generalizable as they are very much dependent on the conditions of the experiment. It remains that econometrics as well as experimental evaluation can, under certain circumstances, usefully fuel our methodological design.

With regards to econometrics, which unlike experimental economics, requires the access to vast bodies of data, pioneering works have been carried out under the EGERA project by Anne Boring (2017, 2021), to measure gender bias in professors' evaluation by students or in the patterns of students' international mobility preferences, and the impact of mitigating measures. Further tested over the last year of the SUPERA project (2021-2022), the use of evaluation experiments will be expanded for RESET, for which a specific task (2.5) has been foreseen in the Grant



Agreement, specific resources have been planned, and experiments-focused GDPR provisions have been included to the project Data management plan.

Evaluation experiments are the most innovative component of RESET M&E plan. As there is a growing demand for data-driven policy design and evaluation, not only gender equality policies do not make an exception, but those are especially likely to trigger such demands, as searching for evidence of gender bias and inequalities often amounts to a form of organizational (or individual resistances). Whereas there is a point not to fall in the “data trap”, which would consist in requiring gender equality policy efforts to draw upon a more robust body of evidence than any other policy area, providing quality data and solid evidence elaborating on control experiments can certainly contribute to lowering such resistances while informing quality policy design. The considerable scholarly and non-scholarly interest<sup>8</sup> triggered by the first evaluation experiments carried out by Sciences Po in this field, indicate that such evaluation protocols can help expanding awareness on gender bias in the academia. Evaluation experiments will be summarized in a specific deliverable (D.2.4), submitted on M38.

### 4.3. Core components and timelines of execution

The main building blocks and timelines of the M&E of RESET are stated in the G.A of the project and detailed in 6 tasks and 5 deliverables. Below developments provide further information about these components and how they will be articulated, as well as to precise *intermediate* internal reports (not delivered to the EC) and timelines of execution.

As a reminder, the G.A agreement of RESET consigns six tasks with regards to M&E. Three of these tasks will run for most of the project (2.2, 2.4, 2.5), thus requiring further inputs to provide a clear overview of M&E, while two (2.1 and 2.3) will last for about half a year, partly overlap and cover roughly the first year of the project. The core components of each task and their respective timelines are summarized below.

#### Task 2.1 Design and the Monitoring and Evaluation Plan (M1-M6) - COMPLETED

Monitoring and Evaluation plan, including a methodological review and a framework for developing RESET theory of change (Sciences Po), M6 (June, 2021).

##### Process or method:

Plan developed by Sciences Po based on the insights of the SUPERA project, updated to integrate lessons learnt, address the specifics of RESET, and the capacity-building needs identified during the first M&E capacity building session delivered by Sciences Po in June 2021.

##### Next steps:

- ✓ Operationalization and further enhancement of the M&E plan by Sciences Po (M7-48)
- ✓ Implementing partners to develop their contextualized project’s impact pathways (M7-12)

---

<sup>8</sup> See for instance the paper by Boring, Ottoboni and Stark (2016), viewed over 80,000 times on [Sciences Open](#) or its sequels in, among many other media outlet, [Time Higher Education](#).

## Task 2.2 Building indicators for the operationalization of GEPs (M4-M46) - ONGOING

A set of SMART (Specific, Measurable, Achievable, Realistic and Timely) and SPICED (see Table 2) qualitative and quantitative indicators for the assessment of the operationalisation of GEPs is to be developed and proposed to RESET partners.

Foreseen indicators include (but will not be limited to):

- The degree of stakeholders' involvement in the co-design and implementation of GEPs
- The availability of resources (in terms of staff, knowledge and time)
- The mechanisms adopted to ensure accountability towards set objectives
- The support from top leadership to implemented actions
- Their impact on organizational culture and daily routines. I

### Process or method:

To be initially included in D2.2 - Checklist for GEP Monitoring and Evaluation, to be submitted on M12 (December, 2021), a first set of evaluation questions and potential indicators, elaborated for the SUPERA project, has been annexed to this Monitoring and Evaluation Plan (See: Addendum). This will allow RESET partners to build ownership over the proposed indicators and have them ready to inform the GEP design process, and to Sciences Po to collect their feedback to enrich and update the checklist to be submitted on M12 based on their needs.

### Next steps:

- ✓ Sciences Po to support implementing partners in the co-design of their GEPs-specific indicators. From the perspective of formative valuation, it will also further equip partners with key capabilities to implement structural change and operationalize the set of indicators built under Task 2.2. This will entail the delivery of at least two more thematic M&E workshops at RESET project meetings (M7-M18)
- ✓ Partners to elaborate their own GEP-specific sets of indicators (M4-M18)

As underlined in the G.A, this task will also consider the hindering factors to implementation and the responsiveness of GEPs to encountered resistances.

## Task 2.3 GEPs design assessment (M10-M18) – NOT STARTED YET

Sciences Po will thoroughly assess the GEPs drafted under WP1 and the co-design process through which they were elaborated, prior to their official endorsement. The aspects covered by this assessment will include:

- The degree of stakeholders' participation and the mobilization of supporting data (WP3)
- The adequacy of the objectives to identified challenges
- The internal (articulation between the general & specific objectives and the actions engaged in each GEP), and external coherence (relationship with EU and domestic gender equality policies, and articulation with relevant intra-institutional policies) of the GEPs
- The innovativeness of the planned actions.



In addition, specific attention will be paid to:

- The inclusion of an intersectional perspective
- The actions aimed at mainstreaming an inclusive definition of merit and excellence in relation to WP6
- The balance of short, mid- and long-term actions and d) between actions aimed at raising awareness, building capacities, fixing numbers and fixing institutions.

#### **Process or method:**

For this assessment, the analysis of the GEP-related documents submitted by RESET partners, will be contrasted with the information collected during: a) thematic bilateral M&E (online) meetings; b) WP1- and WP2- dedicated time slots during project meetings and c) interviews and/or focus groups carried out during onsite visits.

#### **Next steps:**

- ✓ Sciences Po to deliver second round of bilateral M&E online meetings (M10-11)
- ✓ Implementing partners to submit their GEP drafts (M10-M18)
- ✓ Sciences Po to provide timely and thorough GEP reviews (M10-M18)

#### **Task 2.4: Continuous GEP implementation monitoring (M18-48) – NOT STARTED YET**

The implementation of each GEP will be continuously monitored from its official endorsement to the final stage of RESET. Specific attention will be devoted to:

- Supporting partners in identifying and adapting to their respective windows of opportunities for institutional change
- Partners to adopt an intersectional focus in the implementation of GEP actions and activities
- Partners to monitor both GEP process and performance through informing GEP indicators
- Partners to anticipate GEP implementation risks and adopt mitigation measures
- Ensuring the sustainability of GEP actions in terms of available resources and skills, leadership commitment and institutionalization

In the last phase of the project, the direct outcomes of the project and of each plan will be identified at the level of the participating organizations, assessing the observed outcomes compared with the initial objectives while considering the resources available (efficiency). The findings of this assessment will be summarized in a final evaluation report, in which the sustainability of project's results sustainability will be given specific attention.

#### **Process or method:**

This monitoring will be based on the regular updates provided by implementing partners during the sessions devoted specifically to WP1 and WP2 at project meetings and through their contribution to the technical interim report.

This information will be complemented by:

- The data harvested under WP3 for the purpose of partner's self-monitoring and assessment
- The first-hand information gained through the onsite visits carried out by the M&E team at each implementing partner and the regular online bilateral M&E meetings.

#### Next steps:

- ✓ The monitoring process to be launched from the official endorsement of each GEP
- ✓ Sciences Po to dedicate onsite visits to the assessment of GEP implementation
- ✓ Recommendations for enhancing GEP's implementation and sustainability to be delivered in the Interim Monitoring Report (D2.3) and the Final Evaluation Report (D2.5), on M24 and 46.

#### Task 2.5: Experimental evaluation (M12-M38) – NOT STARTED YET

Sciences Po team will elaborate on pilot experiences led under the EGERA (2013-2017) and SUPERA (2018-2022) projects, to mobilize data-based experimental evaluation to assess the impact of relevant GEP actions. Experimental evaluation can be mobilized either to inform the design of data-driven GEP actions or to monitor the impact of those already implemented through controlled experiments contrasting for instance the performance or attitudes of different sample groups towards achieving gender equality, depending on their participation/exposure to a certain activity or communication.

#### Process or method:

Experiments will be selected with GEP implementing partners based on the following criteria:

- ✓ Data availability and user-friendliness (the data set should already exist and structured in such a way that it can be used for the purpose of the experiment)
- ✓ Relevance to the project's overarching objectives (the data should inform the design, implementation or evaluation of a GEP action)
- ✓ Respect of local and project's GDPR provisions and agreement between Sciences Po and the implementing partners concerning the condition of access to and use of the data set

#### Next steps:

- ✓ Experiments will be carried out over years 2 and 3 of RESET, in coordination with WP3 and strict data management rules will apply.
- ✓ A summary report of the findings of these experiments (D2.4) will be compiled on M38).

#### Task 2.6: M&E of management and activities at consortium level (M1-M48) – ONGOING

As consortium-level activities are key to support partners in effectively designing and implementing their GEPs, management procedures and decision-making processes in the consortium will be monitored, as well as the dissemination strategies of the project. An Interim

monitoring report (D2.3) will communicate the progress achieved until the end of the second year of the project. It will contain specific instruments for evaluating the embeddedness of GEPs in the longer term. In the last phase of RESET, the direct outcomes of each plan will be identified at the level of the participating organisations, assessing the observed outcomes compared with the initial objectives while considering the resources available (efficiency). The findings of this assessment will be summarised in a Final evaluation report (D2.5), in which the sustainability of the project's outcomes and impact sustainability will be given specific attention.

### **Process or method:**

Sciences Po will actively participate to all major project meetings, and brief online exit questionnaires will be developed, to be submitted after major project meeting or joint activities. Those will assess the satisfaction of participants with regards to: a) organizational aspects; b) Content (knowledge transfer, usefulness with regards to GEP implementation and the degree of command over the project) and c) interactions. The information collected through exit questionnaires will be thoroughly analyzed and compiled in the Interim Monitoring Report and the Final Evaluation Report on M24 and M48.

Sciences Po will carry out periodic thematic bilateral online meetings with the coordinator and each WP leader, to identify potential flaws or risks in terms of project management and delivery, and propose mitigation measures, and report about task 2.6 at each consortium and follow-up meeting. This task will also be informed by partners' contributions to interim technical reports.

### **Next steps:**

- ✓ Sciences Po provided first recommendations on project management at RESET consortium meeting, held on June 9<sup>th</sup>, 2021. Those primarily focused on highlighting good practices for the sustainability of the project, and further enhancing cross-WPs coordination, notably to facilitate M&E.
- ✓ The Interim Monitoring Report will include specific recommendations with regard to management, decision-making, risk management, problem resolution and quality assurance processes, both at the level of the consortium and at partner's level.

## Annex I. EES Evaluation Capabilities Framework

The EES framework refers to three key categories of capabilities:

- Evaluation knowledge entails the comprehension of evaluation approaches, models and theories and their implications with respect to evaluation design, purposes, practices and methods.
- Professional practice refers to the sound technical principles adopted by the evaluator(s) but also to the communications and interpersonal skills required to design and implement the evaluation.
- Dispositions and attitudes refer to more personal skills enabling evaluator(s) to operate also in adverse environments, as well as to robust ethical principles in conducting their work.

### I. Evaluation knowledge

#### *I.1 Appreciates the distinctive role played by evaluation in society*

- I.1.1 Exhibits familiarity with evaluation theories, schools and approaches
- I.1.2 Shows awareness of evaluation history and trends
- I.1.3 Appreciates the linkages between evaluation and the social sciences
- I.1.4 Understands program theory and its implications for evaluation
- I.1.5 Aims at evaluation independence and excellence in all contexts

#### *I.2 Masters the antecedents of evaluation quality*

- I.2.1 Uses appropriate evaluation concepts and correct evaluation terms
- I.2.2 Displays a capacity to identify relevant evaluation questions
- I.2.3 Knows how to engage constructively with all stakeholders
- I.2.4 Comprehends the value of diverse evaluation approaches
- I.2.5 Adapts evaluation designs and methods to specific contexts

#### *I.3 Understands the potential and limits of evaluation instruments and tools*

- I.3.1 Data collection and analysis
- I.3.2 Experimental and quasi experimental methods
- I.3.3 Qualitative, participatory and mixed methods
- I.3.4 Case studies, surveys, interviews, expert panels
- I.3.5 Indicators, rating and monitoring systems

### II. Professional practice

#### *II.1 Demonstrates capacity to manage and deliver evaluations*

- II.1.1 Responds to legitimate stakeholders' needs and concerns
- II.1.2 Assesses the evaluation context and identifies the program logic
- II.1.3 Manages resources and skills prudently so as to achieve results
- II.1.4 Gathers, uses and interprets evidence with care and judgment
- II.1.5 Reports fairly and encourages effective use of evaluation results

#### *II.2 Displays interpersonal skills*

- II.2.1 Writes fluently and communicates clearly
- II.2.2 Values teamwork and leads by example
- II.2.3 Uses sound negotiating and conflict resolution skills
- II.2.4 Demonstrates gender awareness and cultural sensitivity
- II.2.5 Nurtures professional relationships

### III. Dispositions and attitudes

- III.1 Upholds ethical standards and democratic values in the conduct of evaluations
- III.2 Reaches out to clients and stakeholders
- III.3 Evinces independence of mind and appearance
- III.4 Displays self-awareness and pursues continuous professional development
- III.5 Contributes to the evaluation community

## Annex II. Checklist for GEP design and implementation

The checklist below has been developed for the M&E of the SUPERA project. It thus provides a relevant benchmark element for RESET to build capacities for (self) Monitoring and Evaluation at an early stage of project implementation. It will be further enriched and fine-tuned over the next six months from the submission of the present deliverable, and its updated version will be submitted by M12 (December, 2021) as a stand-alone deliverable.

### Why a checklist?

#### *Purpose (and limitations) of checklists for action*

Checklists have become a popular tool to support the implementation of plans or strategies in a number of domains. Yet, this popularity partly results from a misconception about their usage: for the crew of an aircraft, a checklist is not meant to replace the necessary knowledge about avionics, mechanics and safety procedures, but to offer practical guidance and a support for mobilizing such knowledge at key moments such as prior to take off. Similarly, for a team in charge of driving changes in a complex organization, no checklist can possibly be substituted to the knowledge about this organization gathered by the team members – for instance through a baseline assessment, nor it can replace careful planning, self-reflexivity and learning collectively from experience along the journey.

Confronted with the challenge of auditing and transforming large organizations such as universities, research institutions and funding agencies, involved stakeholders are often much in demands of ready-made instructions or manuals to undertake this endeavor. This is particularly the case when those challenges are to be faced with limited resources or gender expertise, and largely explains why guidelines, toolkits and checklists are being developed in practically all EU-funded initiatives in this realm, and why such tools constitute a significant portion of the good practices collected in resource instruments such as the GEAR tool developed by EIGE<sup>9</sup>.

Beyond this particular context, checklists are also relatively commonplace in program evaluation, where they are designed to support quality monitoring and evaluation (see, for instance, Scriven, 2013; Sufflebeam, 2004 and 2007). Taking this into account, the purpose of the checklist presented in this deliverable is limited to the following aspects:

- To offer a brief digest of the most widely shared criteria for the setting up and implementation of a GEP in research and the academia
- To provide guidance with regards to the fundamental resources to be mobilized for a successful GEP implementation
- To offer practical recommendations with respect to key aspects such as securing top-leadership support, involving stakeholders and building indicators for the monitoring of planned measures
- To highlight the necessity to plan and implement actions in a sustainable way, with view to impact organizations as they actually work, through daily routines, ways of doing things and existing procedures or schemes.

---

<sup>9</sup> <https://eige.europa.eu/gender-mainstreaming/toolkits/gear>

Hence, *it is not* a step-by-step guide for GEP design and implementation, as it is acknowledged that institutional and organizational contexts matter and that setting-up and carrying out a GEP should always be framed according to local existing regulations, needs, practices, available resources and, most important, windows of opportunity for change. Additionally, such step-by-step guides already exist (see, for instance, EIGE toolkit and GENOVATE, from which it is possible to draw inspiration useful for tailor-made action. It is not either and for the very same reasons, a *vademecum* of actions to be potentially implemented.

### *Purpose of a checklist for GEP implementation as part of RESET*

This checklist is conceived as a tool for the monitoring of GEP design and implementation and the formative evaluation of RESET. By definition, a formative evaluation is two-dimensional, as it focuses both on implementation (how the project is being conducted, in relation to planned activities, milestones and outputs), and progress towards set objectives, the logics being that the former conditions the latter. Progress evaluation requires to collect data allowing to benchmark participants progress towards set objectives, but also information about the factors of success or hindrance for achieving these objectives, as well as regarding the impact of carried out activities. This checklist focuses on the second pillar of a formative evaluation (self-assessment and monitoring to improve the program) and is meant to inform the design and implementation of the GEP at partners' level, but also to allow for collecting relevant information on partners' performance in achieving their own targets in terms of audience, participation and pursued institutional changes as part of GEP implementation.

#### Evaluation questions regarding GEP design

- ✓ Is the design of the GEP based on an in-depth organizational gender analysis? (the mandatory in-depth analysis draws on sufficient data and knowledge)
- ✓ Is the GEP based on a clear, context specific strategic framing of issues at stake?
- ✓ Does the GEP design involve the contribution or consultation of key stakeholders?
- ✓ Is the GEP articulated with existing gender equality strategies or bodies?
- ✓ Is the design process sufficiently transparent? (information is communicated to relevant categories of stakeholders, such as social partners, student's organizations and decision-making bodies of the organization)
- ✓ Does the GEP adequately cover the four areas of actions?
- ✓ Is it sufficiently holistic not only in terms of issue coverage, but also through addressing people *and* structures?
- ✓ Does the GEP consider intersecting inequalities and disadvantaged groups?
- ✓ Are the tasks' responsibilities formulated and clearly ascribed?



### Evaluation questions regarding GEP implementation

- ✓ Does the GEP implementation relies upon sufficient stakeholders' mobilization?
- ✓ Has the GEP been endorsed officially at the appropriate level of responsibility?
- ✓ Is it backed at the highest level of the organization?
- ✓ Has the GEP been made public and disseminated and do people know about it?
- ✓ Are the foreseen enforcement mechanisms working efficiently?
- ✓ Are the Gender Equality Boards and Co-design workshops serving their intended purposes?
- ✓ Are the targeted audiences being engaged with?
- ✓ Is it responsive to encountered resistances and windows of opportunity?
- ✓ Is the implementation process as holistic as intended in the GEP design?
- ✓ Does the GEP strive towards the sustainability of the most structural actions?
- ✓ Are there any specific external variable(s) serving or constraining GEP's implementation and how are they addressed?

### Review of available checklists

Checklists are likely to be developed for any sort of project aiming at triggering changes in a given organization. More specifically, they have been designed as part of EU-funded initiatives supporting structural changes for gender equality and integrating a gender perspective in research and the academia through the mean of GEPs. In these particular projects, checklists have been included to broader capacity building instruments, such as toolkits and guidelines for GEP implementation, some of which are collected in EIGE's Gender in Academia and Research (GEAR) online tool. The overarching objective of these tools is similar, although their purpose is of different ranges: providing practical support to RPOs and RFOs undertaking a process of change towards greater gender equality and inclusiveness.

To achieve this objective, guidelines usually provide step-by-step, process-related recommendations as well as examples of actions that can be implemented to address a number of issues related to structural change. Checklists are then either annexed, as a merely practical instrument referring to the broader information contained in the guidelines, or subsumed in the latter, as for instance in the Contextualised Guidelines for Universities and Research Organisations<sup>10</sup> developed under GENOVATE, where the overall content of the document refers to a limited number of items to be taken into account for GEP implementation, thus to be assimilated to a checklist. Toolkits, instead, typically combine step-by-step guidelines covering the life cycle of a GEP, from its design and planning to its ex-post evaluation, with additional, issue

<sup>10</sup> <http://www.genovate.eu/resources/contextualised-guidelines/>

specific resources offering inspiration to design actions addressing particular aspects of structural change, as well as to tackle individual and organizational resistances. Such toolkits have been developed for instance under INTEGER<sup>11</sup> and EGERA<sup>12</sup> projects and might also include references to good practices databases<sup>13</sup>.

### *GEAR steps and impact drivers*

Arguably, one of the most comprehensive checklists developed so far in relation to GEP planning, design, implementation and evaluation in research and the academia, are those annexed to the six steps approach defined in EIGE GEAR action toolbox<sup>14</sup>.

Based on an extensive review of initiatives, programs and policies led across the EU with the objective of transforming research performing organizations and universities towards greater gender equality and gender sensitivity (EIGE, 2016a, 2016b) GEAR addresses structural change through the mean of transformative Gender Equality Plans as a process following six steps.

- ➔ **Step 1: Getting started**
- ➔ **Step 2: Assessing the state of the play in the institution**
- ➔ **Step 3: Setting-up a GEP**
- ➔ **Step 4: Implementing a GEP**
- ➔ **Step 5: Monitoring progress and evaluating a GEP**
- ➔ **Step 6: What comes after a GEP?**

For each step, a series of pre-requisite jointly constitute a rather comprehensive checklist covering all stages of the life cycle of a plan, from planning (step 1 & 2) and design (step 3), implementation (step 4), to evaluation (step 5), while step 6 addresses the issue of sustainability of implemented measures. Whereas the full description of each step is available p.17-29 of the GEAR toolbox, pre-requisites are summarized hereafter in form of a checklist.

### ➔ **Step 1: Getting started**

- ✓ **Understand the context** with regards to aspects such as size, structure, governance, history (also in terms of implementing GE policies), organizational culture...
- ✓ **Find support** through a mapping of available (gender) expertise, potential allies, funding opportunities and alliances to be built beyond the institution
- ✓ **Understand the Gender mainstreaming cycle**, also gathering knowledge about the methods attached to each stage of the cycle, as shown in the fig. below:

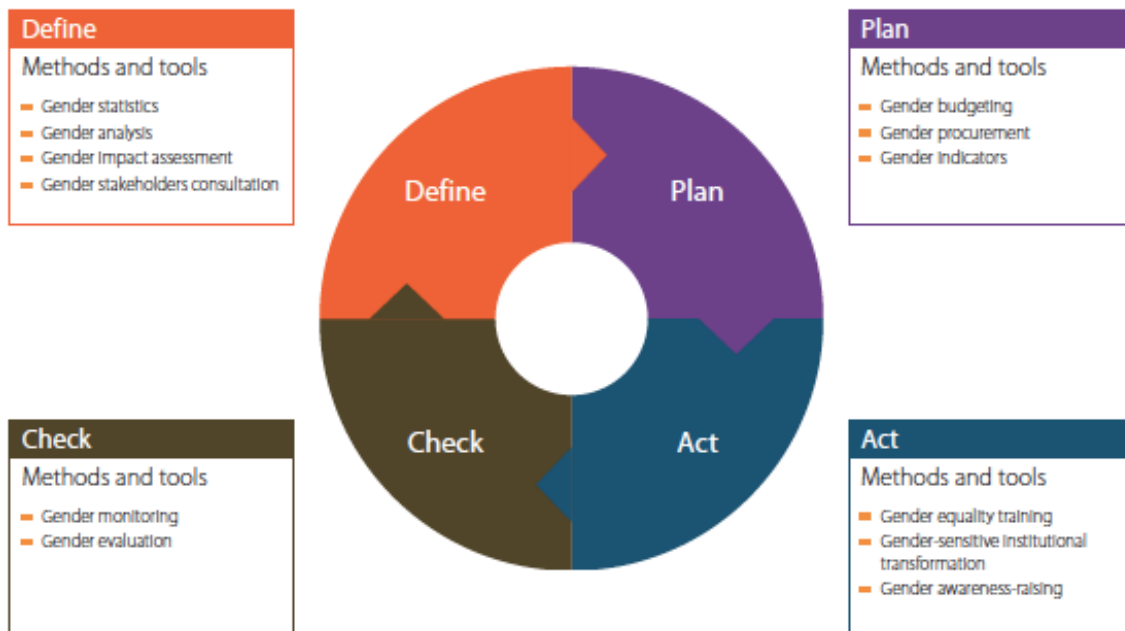
<sup>11</sup> <http://integer-tools-for-action.eu>

<sup>12</sup> [https://www.egera.eu/fileadmin/user\\_upload/Deliverables/D.7.5\\_Structural\\_ChangeToolkit.pdf](https://www.egera.eu/fileadmin/user_upload/Deliverables/D.7.5_Structural_ChangeToolkit.pdf)

<sup>13</sup> [https://www.egera.eu/fileadmin/user\\_upload/Deliverables/D64\\_Database\\_of\\_selected\\_good\\_practices\\_for\\_gender\\_sensitive\\_research\\_81604.pdf](https://www.egera.eu/fileadmin/user_upload/Deliverables/D64_Database_of_selected_good_practices_for_gender_sensitive_research_81604.pdf)

<sup>14</sup> <https://eige.europa.eu/gender-mainstreaming/toolkits/gear/action-toolbox>

**Fig. 8 Gender Mainstreaming cycle**



Source: GEAR toolbox, p. 18

➔ **Step 2: Assessing the state of the play in the institution**

- ✓ **Review relevant legislation and policies national level** (and sub-national, where appropriate), as those can offer support to designing and implementing a GEP.
- ✓ **Gather and analyze sex-disaggregated data about staff and students**
- ✓ **Identifying existing measures to promote gender equality** at the level of the organization itself.

➔ **Step 3: Setting-up a GEP**

- ✓ **Promote the participation of actors** of all levels when defining GEP measures/actions
- ✓ **Draw inspiration** from measures implemented in other organizations or contexts
- ✓ **Define SMART objectives and measures for the GEP**, that is measures which are:
  - Specific
  - Measurable
  - Attainable
  - Realistic
  - Time-related
- ✓ **Identify and use existing resources**

## → Step 4: Implementing a GEP

- ✓ **Be accountable and create accountability to the plan**, planning regular information and dissemination activities with relevant stakeholders to inform them about implemented actions, enhance motivation and maximizing impact of actions
- ✓ **Give visibility to the GEP** by tailoring communication messages for all categories of stakeholders and users, reporting about actions and progresses and communicating beyond the institution.
- ✓ **Be aware that further adjustments will be needed** to adapt to reality and to unplanned needs or problems
- ✓

## → Step 5: Monitoring progress and evaluating a GEP

- ✓ **Consider monitoring and evaluation and integral part of the change process**
- ✓ **Use the baseline assessment as a tool for ex-ante evaluation** in order to accurately measure the impact of implemented measures
- ✓ **Monitor implementation** to provide assistance where necessary and ensure that lessons are learnt and can further improve GEP implementation
- ✓ **Consider monitoring and evaluation as key instruments for sustainability**

## → Step 6: What comes after a GEP?

- ✓ **Take into consideration the lessons from previous experience(s)**
- ✓ **Benchmark what other organizations have done or are currently doing**
- ✓ **Continue to engage (new) stakeholders**
- ✓ **Think about how to make your measures sustainable**

The checklist attached to each of the six steps is complemented by “**impact drivers**” defined at the organizational level and at the process level. Those also contribute to set the stage for effective structural change through GEP design and implementation, by defining some sort of basic requirements for success. They are nonetheless more dynamic than the standard prerequisites, resources or actions to be undertaken, which are usually listed in a checklist.

### At organizational level:

- Explicit and clearly communicated support from senior management level, for greater legitimacy of implemented actions
- A Gender equality body located at the appropriate level, responsible for the gender mainstreaming strategy and equipped with sufficient resources and expertise

- Sex-disaggregated data availability, for evidence-based GEP design and follow-up
- Fostering cooperation between different categories of stakeholders
- Embedding into existing structures and management procedures, to ensure the sustainability of pursued changes

At process level:

- Stakeholders' participation in planning and implementation, to ensure that a diversity of experiences is taken on board, and to create ownership
- Setting clear targets and objectives for each area of action and planned measures, in order to assess failures and successes, with clearly ascribed responsibilities for their attainment
- Flexibility and resilience, to adapt to experience and tackle potential resistances
- Competence development, so as to transfer and enhance knowledge for the greatest possible number of stakeholders
- Monitoring and evaluation instruments available

*Checklists developed as part of sister projects*

Other types of checklists have been developed under EU-funded initiatives, as part of guidelines of toolkits for structural change. Those usually embrace only part of the process of change or pursue narrower objectives than the one of informing the design and implementation of fully-fledged strategies, either focusing on self-assessment (INTEGER), a particular issue such as recruitment, promotion and career progression (GENOVATE) or integrating the gender dimension in research projects (Toolkit for Integrating Gender in EU-funded research).

For instance, INTEGER Guidelines for the Self-Assessment of Transformational Gender Action Plans set up in Higher Education and Research *Institutions* provide, among other recommendations, a check list to ensure the sustainability of GEPs adopted under the project, which highlight the need for:

- Basing GEPs' design upon evidence
- Setting out measurable success indicators
- Assigning responsibility for individual actions to key actors
- Securing endorsement from high management level and governance bodies
- Planning resources beyond project's, and/or GEP's timeline

Under GENOVATE, the core guidelines developed for the teams in charge of GEP's coordination, include the following checklist, which is both thematic (issue coverage) and process-related:

1. Work with senior management/leadership.
2. Take a participatory and inclusive approach.
3. Support female academics and researchers directly through career progression or development programmes such as training and mentoring.
4. Improve recruitment and promotion procedures for more gender-equal outcomes.
5. Set gender targets for senior positions.

6. Develop measures to increase proportions of women in very male-dominated disciplines.
7. Seek to influence national policy on academic recruitment/promotion processes.
8. Take advantage of opportunities for synergies with national-level or EU-level developments
9. Ensure accountability of all actions through ongoing monitoring.

The toolkit for integrating gender in EU-funded research (EC, 2011, 2015), of which two versions have been developed to cover FP7 and H2020, includes a useful checklist which intends to provide guidance to those – researchers or not – involved in designing and implementing a research project, for ensuring a balance participation of men and women and integrating a gender perspective in research content. To the difference of those meant to support more holistic strategies, this checklist reflects the different stages of a project's life cycle and focus on questions such as whom the outputs of this project will benefit, does it involve people for experiments, is sex a relevant variable for analysis, are there any sources in the literature pointing out gender specificities of the subject or can its hypotheses, data collection methods or results be influenced by gender, etc., so as to ensure that adequate gender expertise is mobilized to tackle the subject

With a similar objective, evaluation and research funding agencies are also developing their own checklists, as the one released in January, 2018 by the Agency for Evaluation and Prospective of the Valencian Autonomous Community, Spain<sup>15</sup>. Built upon a strong gender expertise, this checklist is arguably the most comprehensive and finest grained elaborated so far with this purpose, and has been soon uptaken by the regional funding agency for assessing the integration of the gender dimension in submitted research proposals. Meaningfully, the checklist opens with a list of items related to assessing the gender component in research content, asking about the inclusion of gender to the literature and theoretical review, as well as to the hypotheses, chosen methods, ethical review and the language used for drafting the proposal. The composition of the research team and ascribed responsibilities are also covered, as well as the gender perspective in the dissemination of outputs and their transfer to society.

Although of a more restricted application, these checklists can be valuably mobilized or adapted for monitoring part of the activities contemplated in a GEP.

---

<sup>15</sup> <https://avap.es/wp-content/uploads/2018/10/LISTA-DE-COMPROBACION-PERSPECTIVA-DE-GENERO.pdf>

## Proposed checklist for GEP evaluation

### *Implementing principles for structural change*

The checklist initially developed for SUPERA project, listed five core principles to guide structural change processes aimed at achieving gender equality and integrating the gender perspective in the academia. Those, listed below are undoubtedly relevant to RESET:

- ➔ The project is meant to be **CUMULATIVE**, building on previous projects and knowledge.
- ➔ It also intends to be **INCLUSIVE**, targeting the whole academic community through participatory and collaborative methodologies.
- ➔ To achieve its goals, RESET also aspires to be **INNOVATIVE**, being action-oriented, experimental and based on transformation design and its user-centered approach
- ➔ It also intends to be **SUSTAINABLE**, maximizing accountability, stakeholders' involvement and aiming at long-term institutionalization of the planned actions
- ➔ Last but not least, it will be **HOLISTIC**, addressing simultaneously in their multi-layered dimensions three thematic and one cross-cutting areas of action: a) recruitment, retention and career progression; b) access to decision-making; c) integrating the gender dimension in research and d) fighting gender stereotypes and bias

The checklist below there subscribes to these core principles.

- ✓ It intends to be cumulative, drawing upon – and avoiding duplications to the greatest possible extent, with existing checklists.
- ✓ It pays attention to RESET's strive for innovation, referring to Gender Equality Boards (or however they will be called in practice) and Co-design workshops as two important implementation mechanisms.
- ✓ It highlights aspects related to stakeholders' inclusion, support and participation.
- ✓ Finally, it primarily aims at ensuring the sustainability of the measures, actions or initiatives initiated through RESET GEPs implementation.

Yet, in the context of RESET, additional elements are to be considered, to respond to the innovative dimensions of this project, and more specifically:

- ✓ The overarching objective of the project, to redesign the notion of academic and scientific excellence to steer it into a more inclusive direction, by which the value of gender diversity is acknowledged and promoted, and gender bias are debunked and prevented through awareness raising and redesigned policies and practices.
- ✓ The intensive use of co-design, as a methodology for building context specific, participatory actions while supporting stakeholders' engagement and inclusion?



- ✓ The focus placed on intersecting inequalities and discrimination, in order to promote a gender+ approach to structural change and facilitate innovations in addressing intersectionality.

All those three core RESET dimensions will be reflected in the updated and expanded checklist to be submitted as D.2.2 on M12. Therefore, the checklist below should be understood as a mere benchmark exercise and a starting point apt to inform the GEP design phase.

### *About resources*

#### **Why it matters**

Setting-up and implementing a GEP requires both qualitatively and quantitatively sufficient resources to be mobilized. Those include dedicated staff, (gender) expertise, time, knowledge about organizational practices and arrangements, facilitation, training, self-assessment as well as negotiation skills. A wise management of these resources is also required. Experience tells that staff mobilized on GEP implementation can experience difficulties when dealing with resistances, obstructions or isolation and that their contribution to put the organization in motion, is not always fully acknowledged.

As made possible by the checklist designed under WP4, an ex-ante self-assessment of the availability, scarcity or absence of requested capacities, allows for identifying gaps and defining needs in terms of resources. Lacking some of the resources to be put into GEP design and implementation should of course not prevent from action but lead to develop and use them carefully. While WP4 will assist implementing partners in developing resources, this checklist shall support them to wisely use and mobilize them.

This is especially important as RESET is short-lived, resource intensive and primarily draws upon EU-funding, which per se, is not sustainable. Hence, when it comes to resources, it will be especially relevant to GEP implementation, to make the most of those available and to project oneself beyond the timeline of the project.

#### **→ How to proceed?**

- ✓ Assess and further monitor the availability of knowledge and skills for structural change within the core team and performing also a screening in terms of attitudes towards facilitating structural change for gender equality.

- ✓ Monitor whether the financial resources made available by the project are wisely and consistently used with respect to your project's and GEP's objectives and timelines: underspending can reflect that opportunities for action or for increasing knowledge and skills are being missed. Overspending can indicate that your effort might not be sustained over the full duration of RESET. A differential consumption of financial resources on respective areas of action should draw attention over your priority agenda, with which it has to be consistent.
- ✓ Human resources are key: ensure that responsibilities for GEP implementation and monitoring are well distributed and can be sustainably exerted over the course of the project and beyond.
- ✓ Make sure that the contribution of all those involved in GEP implementation is duly acknowledged, not only within the GEP team or the consortium, but also towards the whole organization, and that their experience in driving structural change can be valued in terms of professional trajectory.
- ✓ Provide regular opportunities for staff to exert self-reflexivity and creativity, and to share about encountered difficulties, and ensure that their experience is taken into account to support GEP implementation.
- ✓ Along with HR, time is also key: ensure that for each GEP related task, time is made available, especially for permanent staff involved but exerting other responsibilities.
- ✓ Use Gender Equality Boards as fora to discuss about resources allocation, raise concerns and explore additional resources that can be made available to GEP implementation

## → Potential indicators?

### Knowledge/skills

- The regular monitoring of available resources indicates that those are growing steadily and that a broader spectrum of knowledge and skills is covered after 24, 36 and 48 months
- Internal/consortium/external opportunities for training and capacity building are actively pursued and used
- Activities are being carried out internally to disseminate knowledge relevant to GEP implementation

### Financial resources

- Funding resources are spent consistently with the intensiveness of GEP related activities
- Funding resources are distributed accordingly to the priority agenda set in the GEP
- Significant deviations with respect to planned use are reported to project management
- Opportunities for sustainable (internal or external) funding of the GEP are actively pursued

## Human resources

- Tasks and responsibilities are clearly ascribed and well distributed among team members
- GEP implementation does not exclusively falls upon staff contracted on project's payroll
- Senior, permanent staff is involved in the daily implementation of the GEP
- Fora are created – such as co-design sessions or focus groups, to share about GEP implementation and exert self-reflexivity
- The contribution of each team member is explicitly valued within and outside the team

## *About stakeholders' inclusion*

### Why it matters

It is certainly the most widely shared impact driver or factor of success for effective structural change: all relevant categories of stakeholders – depending on context, senior and middle management; supporting, administrative and research staff; students, are to be engaged with GEP design and implementation.

Involving stakeholders in an inclusive way, can help lowering resistances to adopted measures, increase participation to planned activities and enhance ownership and accountability. At the design phase, it is key to reflect their experience of gender imbalances, biases or discrimination but also their conception about their own work. During implementation, it is crucial for activities to reach their targeted audiences, for measures to deliver their full impact and for data and information to be collected in order to monitor the GEP and inform potential changes.

Participation is not only about representation, but also about co-creation: bringing together people with different views and experiences can help designing better solutions, identify neglected issues and anticipate on potential deviations. It also leads to increase creativity and contribute to build and expand a true community of practices around the common goal of achieving structural change.

RESET intends to be methodologically innovative with regards to stakeholders' participation: Co-design workshops are foreseen as shorter-lived co-creative spots through which specific problems and solutions can be worked out, with the participation of a variety of stakeholders.

### → How to proceed?

- ✓ Ensure that stakeholders with a rich knowledge of the organization and how it actually works, are engaged in the process. Those can help identifying hidden mechanisms and saving time. They can also assist in adequately framing solutions.

- ✓ Identify the appropriate communication channel for each stakeholders' category you wish to engage as there is probably no one fit for all. Usually, the most appropriate are those actually being used by a specific category for its daily practice or communication, such as newsletter, mailing lists or social media account.
- ✓ Deliver tailor-made messages which are both consistent with the overall project's objectives and adapted to their target. Some problems highlighted in your baseline assessment, are more likely to draw their attention than other. Once involved, a broader picture can be delivered.
- ✓ Use participatory techniques, privileging those for which capacities are being built under WP4 of RESET. Co-design techniques are adequate to enhance participation, lowering (and dealing with) resistances and create ownership. Co-design workshops have been devised as an innovative forum to apply such techniques under RESET. Although there is some learning cost attached to their facilitation, they also constitute a true opportunity for involving stakeholders.
- ✓ Make the most of what exists: stakeholders should not only be mobilized on an ad-hoc or thematic basis, but also engaged through existing platforms, such as on-the-job training schemes, regular executive meetings or with the support of social partners. Identifying those platforms is deemed important as it is one of the best way to ensure that targeted groups are reached, and that appropriate inputs is received.
- ✓ Creating new platforms, such as a network of gender focal persons, can also be relevant for both GEP implementation and sustainability. Such platforms can be rather flexible, and supported by dedicated communication channels, such as a gender equality newsletter or dedicated social media account.
- ✓ Inclusiveness does not only mean to engage with constituted groups such as those determined by their functions. It also refers to address all relevant categories of users of the organizations, including those on temporary/external contracts or belonging to less visible groups. Adopting an intersectional focus or agenda is thus important to prevent other biases in the adopted measures.

## → Potential indicators?

### Ensuring that knowledge about the organizations is on board

- Stakeholders representing core departments or units are regularly involved through Gender Equality Boards or other GEP-implementing structures, including HRM, central executive services, study supervision (schooling dept., for instance)
- Targeted groups are reached via tailor made communication, and engaged via the appropriate platforms, as evidenced by response or participation rates
- The loss of organizational knowledge is prevented by ensuring that departing stakeholders are replaced by people with a similar organizational background, and by transferring knowledge to the GEP implementation team.

## Participation

- Stakeholders are engaged both through ad-hoc, issue-specific fora (Co-design workshops) and more institutionalized ones (departmental meetings/on-the-job training schemes...)
- Participation techniques used are experience-based, so that their experience is value and that they learn through experience
- Stakeholders' participation is monitored quantitatively and qualitatively, through internal ex-ante/ex-post surveys or questionnaires
- The circle of involved stakeholders significantly broadens along the project, as evidenced by participation records and qualitative assessment of the participation

## Inclusiveness

- It is made sure that all voices can be heard, including those from less institutionalized (or more fluid) groups such as staff under temporary contract or external students
- Communication about the GEP and related activities, is inclusive in terms of written and visual languages, targeted audiences and delivered messages.
- An intersectional approach is regularly convoked, from data collection to drafting measures and monitoring their impact

## *About top leadership (enduring) support*

### Why it matters

Support from the top management usually ranks first among success factors for effective structural change. This is certainly only partially true, as change can also be driven bottom-up or externally and yet, it remains a powerful impact driver, as it can leverage greater support from the mid-management level, and place gender equality high on the strategic agenda of the organization.

However, two characteristics at least are to be met for this support to deliver its promises: it shall be explicit and long-lasting. Explicit support entails that it is made publicly, and through channels that are those usually used for communicating strategic decisions or commitments. Long-lasting entails that it is reiterated on different occasions and translated into support to adopted measures.

Whereas all projects similar to RESET are bound to collect evidences of top management commitment to the objectives of the project, this does not necessarily discard the risk of lip service, nor prevent such support to fade away over time or to be withdrawn. Hence, cultivating this support and making the most of it, are two important strategic goals to support GEP implementation.

## → How to proceed?

- ✓ Regularly report to the top management about the GEP, from its design to its implementation, so that core orientations and measures are known and explicitly endorsed.
- ✓ Involve top management in some co-creation activities, so as to increase the participation of mid-management and enhance ownership among all categories of participants. Show-casing that a GEP is about innovation and an institution in motion, can be a powerful argument to ensure top-management support.
- ✓ Generate accountability, by making the GEP itself publicly endorsed by the top-management, thus making its objectives the ones of the whole community.
- ✓ Identify communication opportunities: there are certainly key venues and events, both internally and externally, for your top management to express support to the GEP. Being informed in due time about it and preparing ready-made communication, can help making the most of these opportunities.
- ✓ Incorporate regular communication on the GEP status in the communication agenda of the top management
- ✓ Similarly, a careful analysis of the core challenges endorsed by the top management as those defining the future of the RFO or RPO, can be useful to craft a message on gender equality and the gender perspective in science as (similarly) strategic goals.
- ✓ Changes can occur in top management over the GEP implementation period: anticipating on potential risks or opportunities driven by this change is thus important.

## → Potential indicators?

### Generate accountability

- Regular briefs about GEP design and implementation are held, and the broader community is informed about this reporting activity
- The GEP is officially endorsed by the top management and this endorsement is made public
- Communication about the GEP is part of the communication agenda of the top management

### Make the most of windows of opportunity

- Key messages are crafted around GEP objectives and core measures, so that they can serve to the strategic communication of the organization, both internally and externally.
- Changes in leadership are anticipated, so that they can be converted into opportunities, or that risks for GEP implementation can be mitigated.

### Ensure that support is long-lasting

- Successes (in outreach, participation or impact) are highlighted, to increase awareness about changes being brought
- Windows of opportunity are used, to ensure that core measures of the GEP are integrated into framework documents such as mission statements or statute

### *About GEP visibility*

#### Why it matters

Although EU-funded projects on structural changes are meant to bring about the adoption and implementation of GEPs, evidences show (EIGE, 2015b) that those are not always officially endorsed and made public, which seriously undermine their status of internal policy document and considerably limits accountability.

Additionally, publicly communicating about a GEP and the process that leads to its adoption, is crucial not only for building a shared awareness about the challenges of the organization with regards to gender equality and integrating the gender perspective in research, but also to make it a valid reference for all categories of staff and stakeholders.

Public commitments can serve as a support for greater accountability, increase the knowledge of the community about the GEP, encourage certain categories of stakeholders to articulate new claims and broaden the scope of structural changes being carried out.

Publicly communicating about the GEP also enhances transparency about the diagnosis and the process that leads to adopting certain solutions, thus lowering potential resistances.

Hence, the visibility of the GEP shall be ensured early enough in the process of its adoption and endorsement, and throughout its implementation.

#### → How to proceed?

- ✓ Ensure to communicate widely about the participatory process leading to GEP design and adoption, and to inform the community about the expected steps/timelines of the process.
- ✓ Provide regular information about GEP related activities, beyond their targeted audiences and through appropriate channels such as social media.
- ✓ Engage with the audiences on social media to further interact with all categories of users.
- ✓ Make sure that once adopted, the GEP is posted online in a reader-friendly version on the official website of the organization, so that it can be referred to at any moment by any



interested staff/student. Location on the website is deemed important, as it also evidence the official character of the document and its ranking in terms of priority and should therefore be carefully devised. Multiple access (for instance from the Intranet) is recommended.

### → Potential indicators?

- A communication roadmap about GEP design, adoption and implementation
- Regular briefs about GEP related activities
- Impressions of and engagement with posts on social media regarding the GEP
- Evidence (such as external requests for information) shows that the GEP is known beyond the limits of the organization, which in return contributes to increase ownership internally.
- GEP available from the website of the organization and also accessible via the intranet
- GEP referred to in other official documents issued by the organization, such as annual reports, mission statements, pluriannual strategies...

### *About GEP sustainability*

#### Why it matters

Sustainability largely derives from above-addressed aspects. Appropriately used resources, stakeholders' engagement, top-management support and visibility do largely contribute, altogether, to the sustainability of a GEP and of the commitments that underpin it.

However, from an institutionalist perspective that is, a perspective that takes institutions seriously, other steps can be taken to ensure that the efforts put into the design and implementation of a GEP will be long lasting, just as the impact of the adopted measures.

Although there is no one best way to sustainability - as context and opportunity structures matter, paying attention at every step to incorporate core actions to a) existing regulation, decision making or training frameworks and b) to daily routines and practices, can enhance the resilience of the actions contained in the GEP.

Sustainability can be pursued through institutionalization (of a gender equality office, of part of the mechanisms and solutions adopted under the GEP, of its funding...). It can also be achieved by incorporating gender equality and the gender perspective in research to the mission statement of the organization or to the core messages it delivers about itself.

It can also be pursued through establishing long-lasting mechanisms for data collection, stakeholders' consultation on gender issues and making gender expertise more readily available throughout the organization. In all cases, it should be pursued at every stage of the project to ensure that the GEP will deliver its full potential for structural change, beyond the timeline of the project.

### → How to proceed?

- ✓ For each contemplated measure, consider which framework document could possibly accommodate the proposed solution, and what would be the process to incorporate it.
- ✓ Although one-shot actions might be occasionally necessary, foresee planned measures as to be integrated to regular procedures or routines, with appropriate capacity building activities for those agents in charge of implementing them.
- ✓ Make the most of opportunity structures: carefully monitor review processes meant to lead to the update or drafting of relevant documents (such as statutes, mission statements, internal regulations or processes...) and offer to involve gender expertise in the process.
- ✓ Ensure that one-shot actions are reported about so that they can be inspiring and possibly duplicated or institutionalized.
- ✓ Actively pursue other funding opportunities for GEP activities.
- ✓ Ensure that in-house gender expertise and knowledge is being built, and will be transferred.

### → Potential indicators?

- Gender training to be incorporated to on-the-job training scheme of the organization
- Gender equality mechanisms reinforced at the institutional level, in terms of mandate, staff and resources
- Gender mainstreaming instruments – such as data collection or gender auditing systems, a network of focal person or dedicated communication channel, are put in place
- Gender knowledge disseminate within the organization, for instance through resources centres, on-the-job training modules, new courses or curricula
- A process of drafting and adoption for a new GEP is initiated or planned prior to the completion of the RESET project
- The existence of a GEP, its period of validity and ideally, its process of adoption, are enacted in a formal institutional document
- The GEP is annexed to the statutes, mission statement or any other regulatory or planning document of the organization

## References

Albenga V. (2016): "Between knowledge and power: triggering structural change for gender equality from inside in higher education institutions", In Bustelo, M., Ferguson L. and M. Forest (eds.) *The Politics of Feminist Knowledge Transfer: Gender Training and Gender Expertise*, Abingdon and New York: Palgrave MacMillan: 139-156.

Alonso A. and M. Forest (2012): 2012. "Is Gender Equality Soluble into Self-Governance? Regionalizing and Europeanizing Gender Policies in Spain", in Forest M. and E. Lombardo (eds.), 2012 *The Europeanization of Gender Equality Policies: A Sociological-Discursive Approach*, Basingstoke: Palgrave, Gender and Politics Series: 192-213

Azmat, G. and Petrongolo, B. (2014) 'Gender and the labor market: What have we learned from field and lab experiments?', *Labour Economics*. Elsevier B.V., 30, pp. 32–40.

Boring A. and Philippe A. (2021) "Reducing Discrimination in the Field. Evidence from an Awareness Raising Intervention Targeting Gender Biases in Student Evaluation of Teaching", *Journal of Public Economics*, 193.

Boring, A. (2017) 'Gender biases in student evaluations of teaching', *Journal of Public Economics*. Elsevier B.V., 145, pp. 27–41.

Boring, A, Ottoboni, K., & Stark, P. (2016). "Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness" *ScienceOpen Research*.

Benschop Y. and M. Verloo (2011) "Gender Change, Organizational Change and Gender Equality Strategies" in Jeanes E., D. Knights and P. Y. Martin: *Handbook of Gender, Work and Organization*, Wiley, 278-290.

Bleijenberg, I.; Benschop, Y. and J. Vennix (2006): "Group model building as a Gender Mainstreaming Instrument", [https://www.researchgate.net/publication/228838098\\_Group\\_model\\_building\\_as\\_a\\_Gender\\_Mainstreaming\\_Instrument](https://www.researchgate.net/publication/228838098_Group_model_building_as_a_Gender_Mainstreaming_Instrument) [accessed Aug 30 2018].

Boussaguet L. Jacquot S. and P. Ravinet (2006) *Dictionnaires de Politiques publiques*, Paris: Presses de Sciences Po.

Bowen G. (2009) "Document Analysis as a Qualitative Research Method", *Qualitative Research Journal*, Vol.9(2): 27-40.

Bustelo, M. (2003) "Evaluation of Gender Mainstreaming. Ideas from a Meta- evaluation Study". *Evaluation*, Vol.9(4): 383-403.

Bustelo, M. (2016). "Evaluation from a gender+ perspective as a key element for (re)gendering the policymaking process". *Journal of Women, Politics & Policy*, Vol. 38:1, 84-101.

Caulley, D. (1983): "Document analysis in program evaluation", *Evaluation and Program Planning*, Vol: 19-29.

Chen, H. T. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective* (2nd ed.). Thousand Oaks, CA: Sage.

Dahmen, J. and Peterson, H. (2017a) *Monitoring Handbook*, GenderTime project edited.

Dahmen, J. and Peterson, H. (2017b) Trying to do the impossible: Monitoring gender equality action plans in seven European Institutions. In Godfroy, A-S. and Pourrat, Y. (eds.) *Transferring, Implementing, Monitoring Gender Equality in Research Careers*. Paris: EDP Sciences.

European Commission, (2011, 2015): *Toolkit Gender in EU funded Research*, Brussels: Directorate-General for Research and Innovation.

EIGE (2015) *Opinion Paper on the Benefits of Integrating Gender Equality and Gender Mainstreaming in Research Performing Organizations*, Vilnius: European Institute for Gender Equality (EIGE) and Brussels: DG Research, European Commission, unpublished work.

EIGE (2015) *Analytical Paper on the Integration of Gender Equality and Gender Mainstreaming in Research Performing Organizations in the EU28*, Vilnius: European Institute for Gender Equality (EIGE) and Brussels: DG Research, European Commission, unpublished work.

EIGE (2013) *Institutional mechanisms in favour of gender mainstreaming, Comparative Intermediate report*, Vilnius: European Institute for Gender Equality (EIGE), unpublished work.

Equality Challenge Unit (2014) *Measuring progress on equality: qualitative evidence*. London: ECU.

Espinosa J., Bustelo M. and M. Velasco (eds.) (2016) *Evaluating Gender Structural Change. Guidelines for Evaluating gender Equality Action Plans*, Madrid: Complutense University of Madrid (UCM).

Niederle, M. and Vesterlund, L. (2007) 'Do Women Shy Away from Competition? Do Men Compete Too Much?', *the quarterly Journal of Economics*, Vol.122(3): 1067–1101

Funnell, S. C. and Rogers, P. J. (2011) *Purposeful Program Theory. Effective use of theories of change and logic models*. San Francisco: John Wiley & Sons.

Jacquot S. (2010) "The Paradox of Gender Mainstreaming: Unanticipated Effects of New Modes of Governance in the Gender Equality Domain", *West European Politics*, Vol.33(1): 118-135.

Krizsan, A. and Lombardo, E. (2013) "The quality of gender equality policies: A discursive approach". *European Journal of Women's Studies*, Vol. 20(1): 77-92.

Lombardo E. and L. Mergaert (2016) "Resistances to Gender Training and Mainstreaming Processes" in In Bustelo, M., Ferguson L. and M. Forest (eds.) *The Politics of Feminist Knowledge Transfer: Gender Training and Gender Expertise*, Abingdon and New York: Palgrave MacMillan: 43-61.

Lombardo E. and L. Mergaert (2013) "Gender Mainstreaming and Resistance to Gender Training: A Framework for Studying Implementation, *NORA - Nordic Journal of Feminist and Gender Research*, Vol. 21: 4, 296-311.

Lombardo E., Meier P. and M. Verloo (eds). (2009) *The Discursive Politics of Gender Equality. Stretching, Bending and Policy-making*. London: Routledge (2nd paperback ed. 2012)

Mayne J. (2015) "Useful Theory of Change Models", *Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme*, Vol. 30(2): 119–142.

Mayne, J., & Johnson, N. (2015). Using theories of change in the Agriculture for Nutrition and Health CGIAR research program. *Evaluation*, Vol.21 (4): 407–428.

Mergaert L. and M. Forest (2015) "Incorporating Gender and Diversity", in Dingwall, R., MacDonnell Mary B. (eds), *The SAGE Handbook of Research Management*, NY: Sage Publishers.

Marx Ferree M. and K. Zippel (2015): "Gender Equality in the Age of Academic Capitalism: Cassandra and Pollyanna Interpret University Restructuring" *Social Politics: International Studies in Gender, State & Society*, Vol.22 (4): 561–584.

Mergaert, L. and E. Lombardo (2014): "Resistance to implementing gender mainstreaming in EU research policy", in: Weiner, E. and H. MacRae (eds): "The persistent invisibility of gender in EU policy" *European Integration online Papers (EIoP)*, Special issue 1, Vol. 18, Article 5.

Mukhopadhyay, Maitrayee, Steehouwer, Gerard and Wong, Franz (2006) *Politics of the possible. Gender mainstreaming and organizational change. Experiences from the field*. Amsterdam: Royal Tropical Institute.

Navarro, N. (2007) *Desigualdades de género en las organizaciones: procesos de cambio organizacional pro equidad*. San Salvador: UNDP.

Nelson J a. (2015) "Not-So-Strong Evidence for Gender Differences in Risk Taking". *Feminist Economics* Vol. 57(1): 1–29.

NSF (2002) *The 2002 User-Friendly Handbook for Project Evaluation*, Washington: NSF.

Prosavac E. (2015) *Program Evaluation: Methods and Case Studies*, New-York: Routledge.

Rogers, P. (2014). "Theory of Change", *Methodological Briefs: Impact Evaluation 2*, UNICEF Office of Research, Florence.

Sainsbury D. and C. Bergqvist (2009) "The Promise and Pitfalls of Gender Mainstreaming", *International Feminist Journal of Politics*, Vol. 11(2): 216-234

Scriven, Michael (2013): Key evaluation checklist, available at: [http://www.michaelscriven.info/images/KEC\\_3.22.2013.pdf](http://www.michaelscriven.info/images/KEC_3.22.2013.pdf)

Sharpe G. (2011) "A Review of Program Theory and Theory-Based Evaluations", *American International Journal of Contemporary Research* Vol. 1(3): 72-75.

Stein, D., & Valters, C. (2012). *Understanding 'theory of change' in international development: A review of existing knowledge*. Asian Institute and the Justice and Security Research Programme.

Sudman, S. (1976). *Applied Sampling*. New York: Academic Press.

Stufflebeam, Daniel L. (2004): Evaluation Design Checklist, available at: [http://www.wmich.edu/evalctr/archive\\_checklists/evaldesign.pdf](http://www.wmich.edu/evalctr/archive_checklists/evaldesign.pdf)

Timmers, T.M., Willemsen, T.M. & Tijdens, K.G. (2010): "Gender diversity policies in universities: a multi-perspective framework of policy measures", *Higher Education* 59: 719.

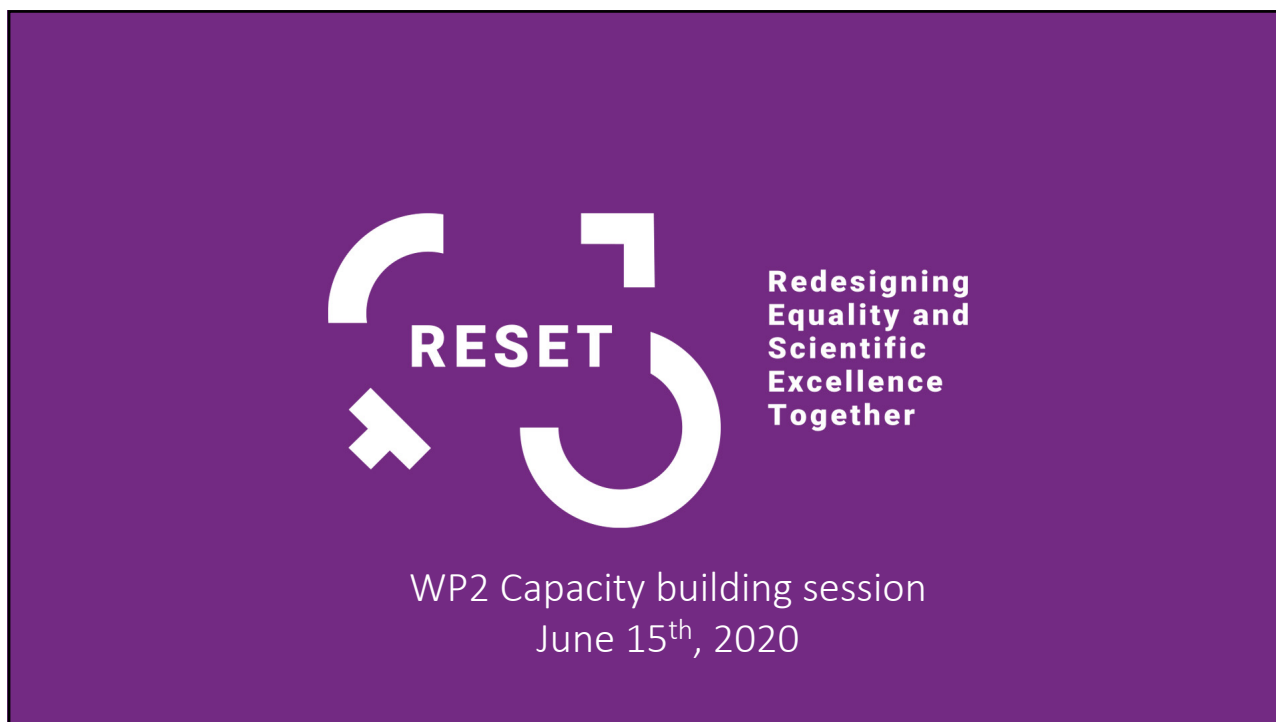
Van den Brink, M. (2015) "The politics of knowledge: the responses to feminist research from academic leaders", *Equality, Diversity and Inclusion: An International Journal*, Vol. 34(6): 483-495.

Van den Brink M. and Y. Benschop (2012) "Slaying the Seven-Headed Dragon: the Quest for Gender Change in Academia", *Gender, Work and Organizations*, Vol. 19 (1): 71-92.

Van den Brink M. and Y. Benschop (2012): "Gender practices in the construction of academic excellence: Sheep with five legs", *Organization*, Vol. 19(4): 507-524.

Vogel I. (2012) 'Review of the use of "Theory of change" in international Development', *Review Report*, UK Department for International Development, See [http://www.theoryofchange.org/wp-content/uploads/toco\\_library/pdf/DFID\\_ToC\\_Review\\_VogelV7.pdf](http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/DFID_ToC_Review_VogelV7.pdf).

Wenger, E. (1998). "Communities of practice: Learning, meaning and identity", *Journal of Mathematics Teacher Education* Vol.6(2): 185- 194.



1



2



RESET Project

## Gender inequalities and bias in research and the academia: the status quo

3

RESET Project

**SEX** refers to the biologically determined characteristics of men and women in terms of reproductive organs and functions based on chromosomal complement and physiology. As such, sex is globally understood as the classification of living things as male or female. Although it is rather fixed, sex cannot be fully encapsulated in this binarity...

**GENDER** refers to the social construction of women and men, of femininity and masculinity, which varies in time and place, and between cultures. As a concept, gender is thus more fluid than sex, although changes in the definition of gender roles usually take time

4

RESET Project

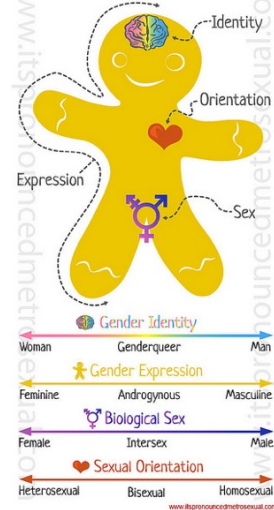
**Biological sex** (objectively measurable organs, hormones and chromosomes)

**Sexual orientation**  
(who I am attracted to)

**Gender identity**  
(how I think about myself)

**Gender expression**  
(how I demonstrate my gender)

Genderbread Person



5

RESET Project

**GENDER EQUALITY** A situation where individuals are **free to develop their personal abilities and make choices** without the limitations imposed by **strict gender roles**.

Although gender equality often refers primarily to the respective situations of women and men, attention should also be brought to preventing discriminations on grounds such as **sexual orientation**, gender identity or expression, and to the situation of **non-binary individuals**.

6

RESET Project

**INTERSECTIONALITY**

Every person is a mosaic of identities: woman/man, old/young, rich/poor, married/single, with a migratory background, etc.



7

RESET Project

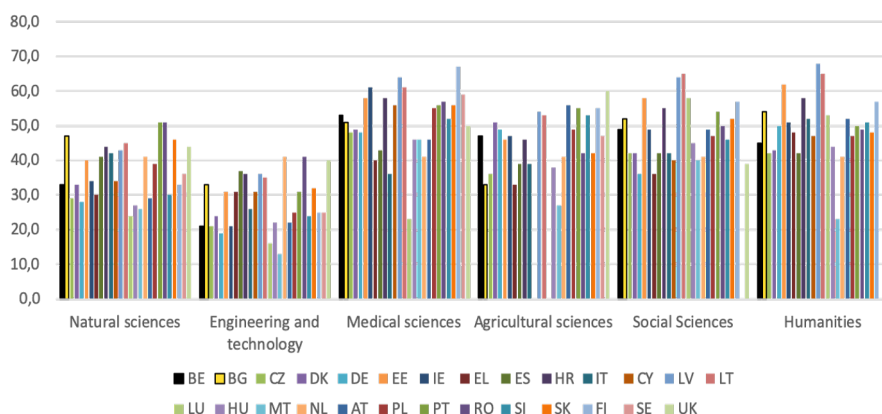
**Building capacities for sustainable  
change and self-monitoring**

8

RESET Project

## Horizontal segregation

Percentage of female researchers in the higher education sector in the EU28 by scientific field, 2012 (adapted from 'She Figures 2015')

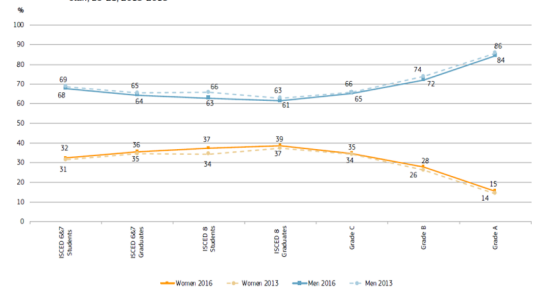


9

RESET Project

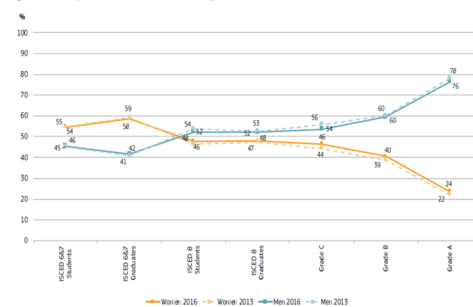
## Vertical segregation

Figure 6.2 Proportion (%) of men and women in a typical academic career in science and engineering, students and academic staff, EU-28, 2015-2016



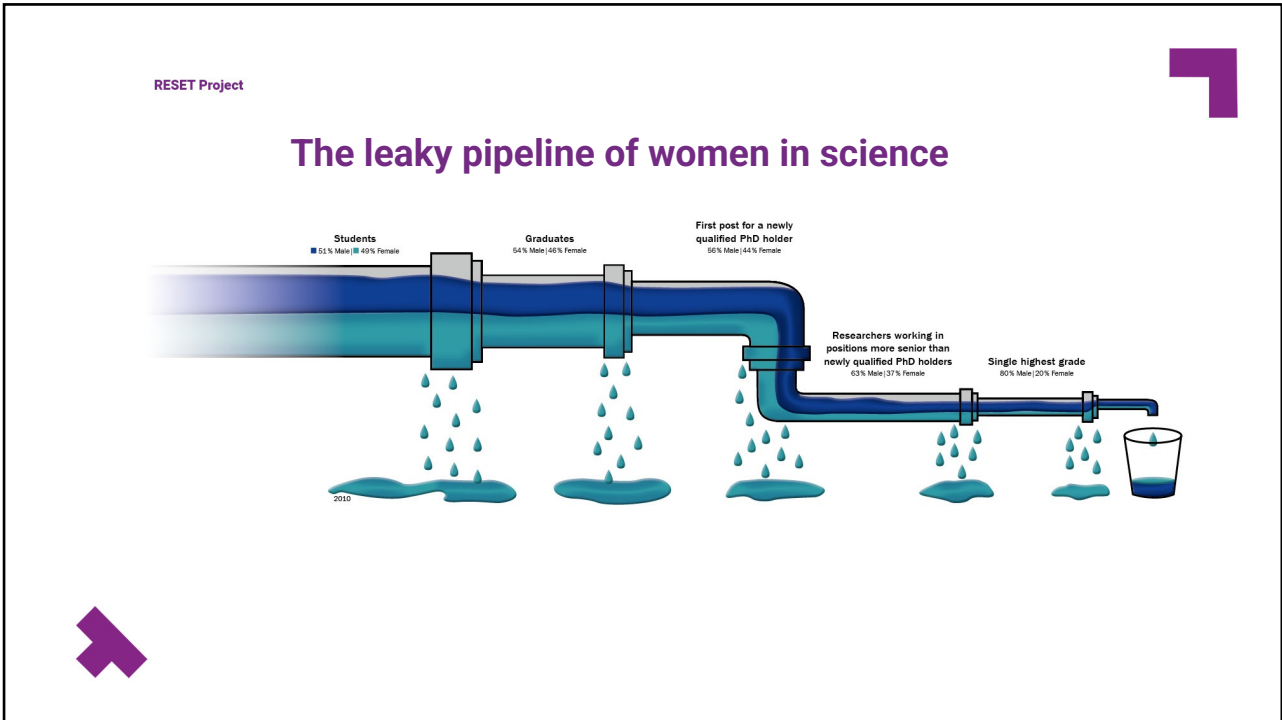
Notes: Reference years for Eurostat data: 2012-2016. Exceptions to the reference year for WS data: HR: 2014-2017; LU: 2015-2016; UK: 2014-2016; CY: AT: SI: SE: 2013-2015; MT (Malta College for Arts, Science and Technology): 2017. Eurostat data unavailable for: PL (ISCED 8 graduates) 2012; MT (Women ISCED 8 graduates) 2012; NL (ISCED 8 students and graduates) 2016; WS data unavailable for: BG, CZ, EE, IE, PL, LT (2015), LV, MT (2015), RO. Others: Data are in headcounts (HC). Break in time series: DE (Grades B-C): 2016; ES: 2015; UK: 2014. Data rounded to nearest multiple of 5. UK: The same person may be counted in several grades. IE (French speaking community). SE: The same person may be counted in several fields. SE: Data do not include persons of unknown sex. PL: Grade C data include some persons with MSc, only LT, UK: Eurostat data for 2015. ISCED 6A7 corresponds to ISCED 5A of ISCED-97; ISCED 8 corresponds to ISCED 8 of ISCED-97. The base reference population of WS data is that of 'Researchers' as defined in the Frascati Manual (OECD, 2015), with the exception of the following countries which used Academic staff based on the UOE Manual (UNESCO/OECD/Eurostat, 2017): BG, DE, EE, EL, FI, LV, LT, NL, SI, SK, SE. Source: Women in Science database, DG Research and Innovation; Eurostat - Education Statistics (online data codes: educ\_emr5, educ\_grad5, educ\_use\_emr03, educ\_use\_grad02).

Figure 6.1 Proportion (%) of men and women in a typical academic career, students and academic staff, EU-28, 2015-2016

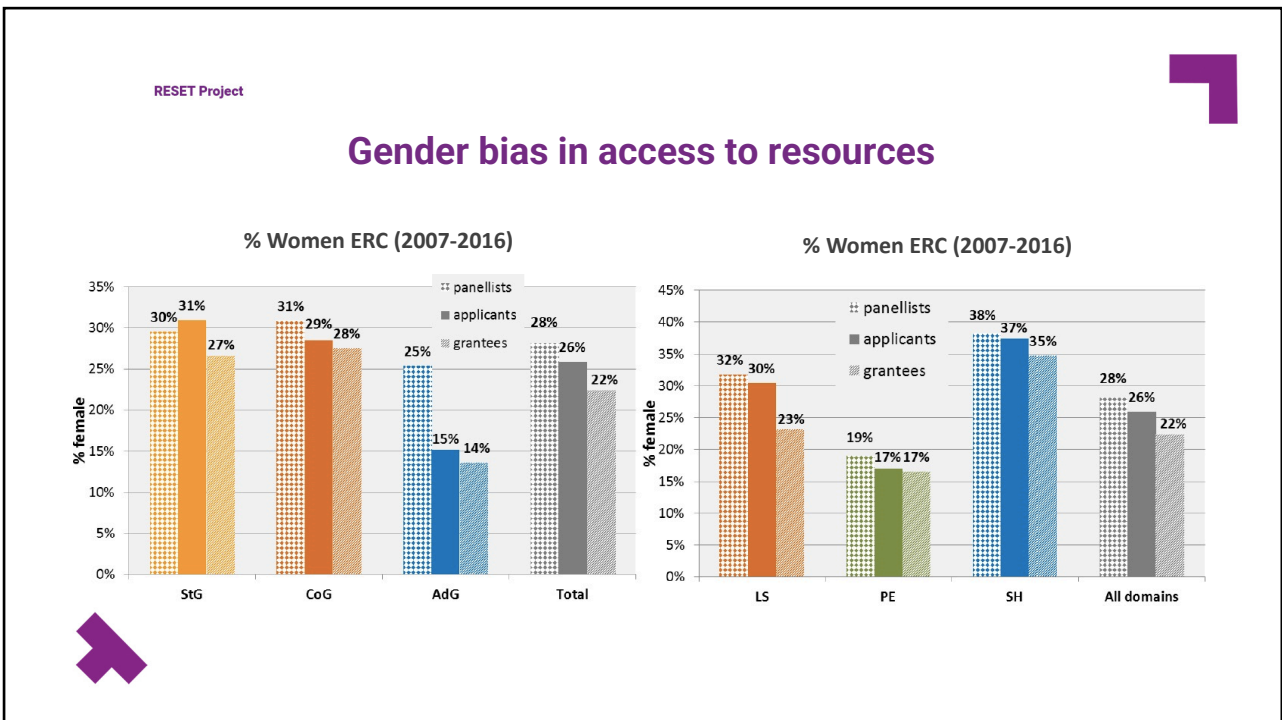


Notes: Reference years for Eurostat data: 2012-2016. Exceptions to the reference year for WS data: CZ (Grade A); SE (Grade A): 2014-2015; FR: 2012-2015; HR: 2014-2017; LU: 2015-2016; RO, UK: 2014-2016; EE, CY, NL, AT, SI, SE: 2013-2015; MT (Malta College for Arts, Science and Technology): 2017. Eurostat data unavailable for: NL (ISCED 8 graduates) 2016; WS data unavailable for: LT (2015), MT (2015), IE (Grade D); Eurostat data for 2015. ISCED 6A7 corresponds to ISCED 5A of ISCED-97; ISCED 8 corresponds to ISCED 8 of ISCED-97. Others: Data are in headcounts (HC). Break in time series: DE (Grades B-C): 2016; ES: 2015; UK: 2014. Data rounded to nearest multiple of 5. UK: The same person may be counted in several grades. IE (French speaking community). SE: Data do not include persons of unknown sex. PL: Private colleges and other smaller institutions are not included. SE: Grade C data include some persons with MSc, only LT, SK: The base reference population of WS data is that of 'Researchers' as defined in the Frascati Manual (OECD, 2015), with the exception of the following countries which used Academic staff based on the UOE Manual (UNESCO/OECD/Eurostat, 2017): BG, DE, EE, EL, FI, LV, LT, NL, SI, SK, SE. Source: Women in Science database, DG Research and Innovation; Eurostat - Education Statistics (online data codes: educ\_emr5, educ\_grad5, educ\_use\_emr03, educ\_use\_grad02).

10



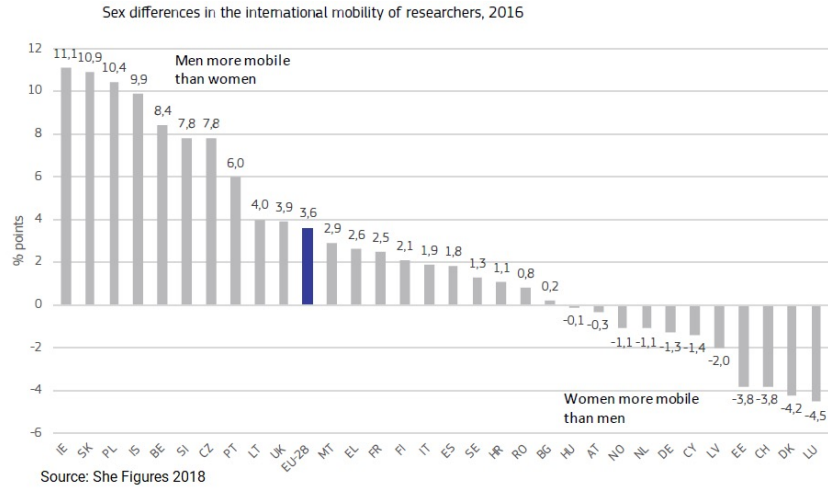
11



12

RESET Project

## Gender bias in access to research mobility



13

RESET Project

## Invisibilizing women in research and the academia

### MATHILDA EFFECT

Coined by Margaret Rossiter, refers to the systematic invisibilisation of women's contributions to science.

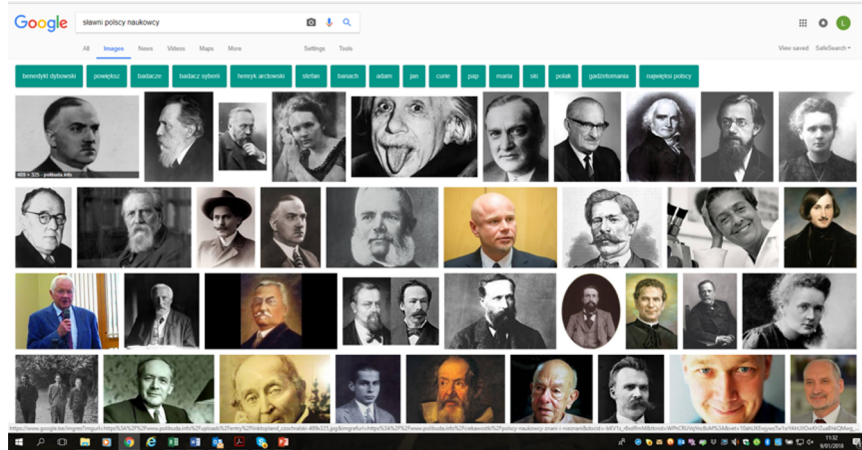


Johnson calculated the flight path for the first NASA mission to space. Her skills were so highly valued that when real computers were brought on the scene, NASA made it her job to verify the computer's results! She co-authored over 26 papers and received credit for only one. Source : <http://blackwomenincomputing.org>

14

RESET Project

## The masculine image of science



15

RESET Project

## About gender blind and gender biased research

16



RESET Project

## Gender blind and gender biased research



It shouldn't come as a surprise, that very gendered and segregated knowledge production settings can produce gendered knowledge

**GENDER BLINDNESS** Failing to take into account potential sex differences as well as the gendered roles and conducts of women and men in society - ultimately leads to reproducing stereotypes, biases and inequalities. It also results in **biased research**

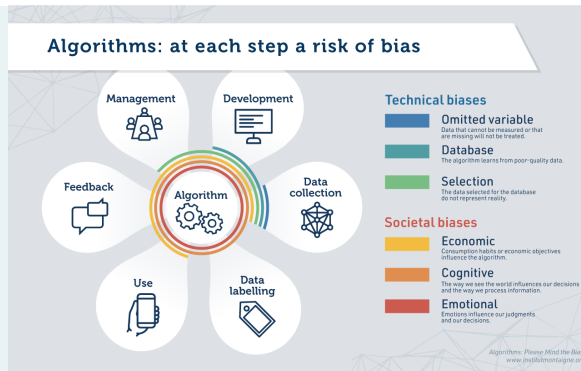
17

RESET Project

## Gender blind and gender biased research



Different symptoms for heart failure

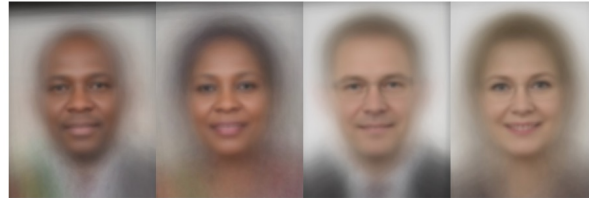


Bias in algorithms creation & use

18

RESET Project

## Gender blind and gender biased research



Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

Source: Institut Montaigne (2020): [Please Mind the Bias Report](#)

19

RESET Project

## Unconscious bias


**UNCONSCIOUS BIAS** occurs when we make judgments or decisions on the basis of our prior experience, our own personal deep-seated thought patterns or assumptions, and we are not aware that we are doing it. The irony is that prejudice are by-products of the efficiency of human cognition.



20

RESET Project

**Time for a break!**



The slide features a white background with a purple L-shaped graphic in the top right and bottom left corners. The text 'RESET Project' is in the top left, and 'Time for a break!' is in the center. A photograph of a white coffee cup on a saucer with steam is on the right.

21

RESET Project

**Gendering research programs and projects in Europe: a new paradigm?**

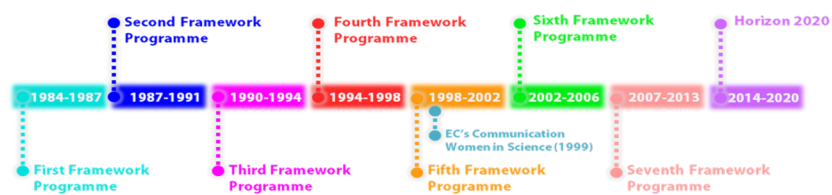
The slide features a white background with a purple L-shaped graphic in the top right and bottom left corners. The text 'RESET Project' is in the top left, and the main title is centered.

22

RESET Project

## Gender equality objectives in EU policies

- Gender equality in scientific careers.
- Gender balance in decision-making.
- Integration of the gender dimension into R&I contents



23

RESET Project

## EU Council conclusions of December 2020

- Recalls, with great concern that there continues to be a major gender imbalance preventing Europe from using the full potential of its R&I system aiming for excellence*
- Calls on the Commission and Member States for a renewed focus on gender equality and mainstreaming , including through the instrument of gender equality plans and the integration of the gender dimension into R&I content*
- Invites Member States and research funding organisations to (...) to ensure that allocation of research funding is not affected by gender bias.*

24



## Integrating gender under Horizon Europe



Gender Equality Plan: **Eligibility Criterion**



Integration of the gender dimension: **Award Criterion**



Gender balance: **Ranking Criterion**

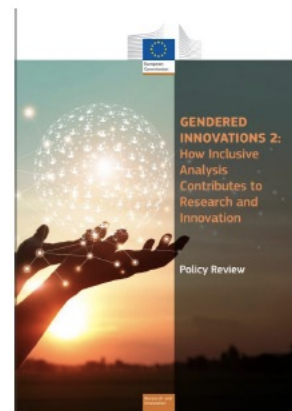
Under Horizon Europe the **integration of the gender dimension into R&I content is mandatory** unless the topic description explicitly includes a sentence such as the following:  
"In this topic the integration of the gender dimension (sex and gender analysis) in research and innovation content is not a mandatory requirement."

Source: Anne Pépin, DG Research & Innovation



## Integrating gender under Horizon Europe

- ✓ EC to recast existing tools to support RPOs and RFOs in advancing gender equality ([GEAR](#)) and to launch new ones.
- ✓ EC released [case studies](#) and fact sheets on gender in various research areas, including ICTs and [AI](#).



RESET Project



## What we know about structural change in practice: principles and key drivers for change



27

RESET Project



### Definition of a gender equality plan

As defined by the European Commission, a gender equality plan consists of a set of actions aiming at:

- Conducting audits of procedures and practices to identify gender bias.
- Identifying and implementing innovative strategies to correct any bias.
- Setting targets and monitoring progress via indicators.

European Commission Communication on 'A Reinforced European Research Area Partnership for Excellence and Growth' (COM(2012) 392 final)



28

RESET Project

## Mandatory process requirements for GEPs under Horizon Europe




**Public document**

- Formal document
- Signed by top management
- Published on the institution's website
- Disseminated through institution




**Dedicated resources**

- Funding for gender equality positions or teams
- Reserved time for others to work on gender equality



**Data collection and monitoring**

- Data on sex or gender of staff across roles and leadership
- Annual reports and evaluation of progress and outcomes



**Training and capacity building**

- Whole organisation engagement
- Tackle gender biases of people and decisions
- Joint action on specific topics



29

RESET Project

## Recommended areas of action



Work-life balance and organisational culture



Gender balance in leadership and decision-making



Gender equality in recruitment and career progression



Integrating the gender dimension into research and teaching content



Measures against gender-based violence, including sexual harassment

Essential factors for gender equality in R&I



30



RESET Project

## About institutional change

**Institutional change** is a strategy aiming at:

- ✓ Removing the obstacles to gender equality that are inherent to the research & higher education system itself.
- ✓ Adapting the practices of organisations... even if it requires to challenge **power distribution** and **ways of doing things**.

It is focused on organizational settings, practices and processes, engaging with individuals primarily as potential **agents of change**.

31

**To effectively transform the status quo, institutional change has to be:**

- ✓ Participatory
- ✓ Holistic
- ✓ Inclusive
- ✓ Visible
- ✓ Flexible & resilient
- ✓ Sustainable

June 2021

32

## Strategically framing change towards gender equality

Change agents need to strategically frame their action so that it can deliver sustainable results:

- ✓ **Identify allies, and engage with them** to agree about the diagnosis and co-design solutions
- ✓ **Identify levers for change** inside the organization and activate them
- ✓ **Identify or create windows of opportunity for change**, connecting gender equality to other crucial strategic goals for the organization

June 2021

33

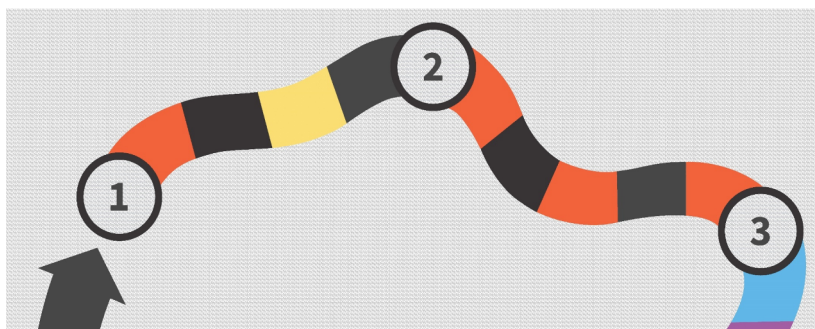
RESET Project

## Designing and setting-up a GEP: a step-by-step process

34

RESET Project

- ✓ **Getting started > stakeholder mapping**
- **Analysing and assessing the state-of-play**
- **Setting up a Gender Equality Plan**
- Implementing a Gender Equality Plan
- Monitoring progress and evaluating a Gender Equality Plan
- What comes after?



35

RESET Project

## STEP 2: Gender analysis

Gather Data on how the situation is right now

Here you diagnose the problems

Participative techniques will focus on analyzing problems and finding insights

36

RESET Project

## STEP 2: Gender analysis

**How do you observe and listen ?**

- ✓ Interviews
- ✓ Focus groups
- ✓ Workshops
- ✓ Observation (ethnography)
- ✓ Diaries
- ✓ Benchmark (for inspiration)

Use creative participatory methods

Persona's, causal diagram, lotus flower brainstorm,....

37

RESET Project

## STEP 2: Gender analysis

### Cause diagram

38

RESET Project

## STEP 2: Gender analysis

### Journey map

	Before	During	After
Timeline			

RESET Project

## STEP 2: Gender analysis

### Definition of a persona:

Personas are fictitious users (archetypes) of a product, service or policy, that represent the goals and wishes of a large group of users during the design process. (Pruitt & Grudin 2003)

**Images**

**Demographic information**

**Description**

**Quotes**

**Motivations**

RESET Project

## Lunch break



41

RESET Project

## STEP 3: Setting up a GEP

### About theory of change:

A theory of change “consists of a set of statements that describe a particular program, explain why, how, and under what conditions the program effects occur, predict the outcomes of the program, and specify the requirements necessary to bring about the desired program effects”. Sharpe, 2011

Elaborating it *ex-ante*, e.g from the **GEP design phase**, helps better structuring and planning the intervention. It is also crucial for designing meaningful evaluation questions and explaining causal relations between inputs/activities and outputs/outcomes and ultimately, impacts.

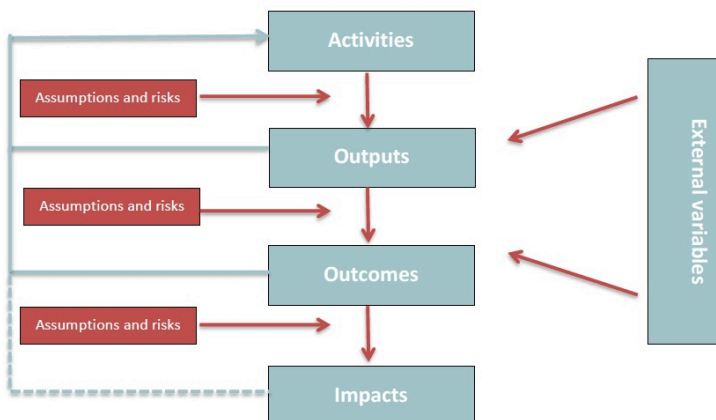
42

RESET Project

### STEP 3: Setting up a GEP

#### About impact pathways:

An impact pathway is a visual representation of the intended process of change and of how activities are meant to deliver desired outputs, outcomes and impacts



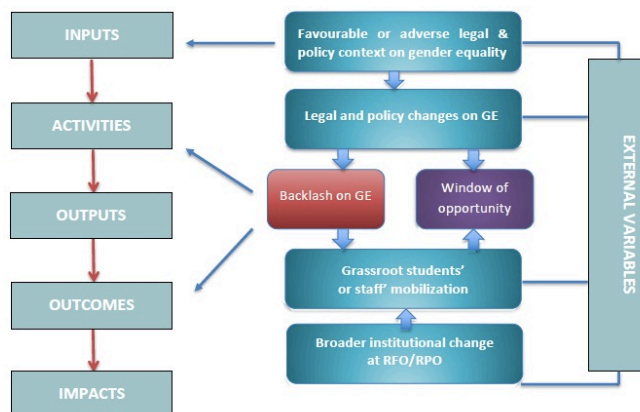
Source: Based on United Nations Children's Fund, Supplementary Program Note on the Theory of Change, Peer Review Group meeting, 11 March 2014, UNICEF, New York, 2014, p. 4. [www.unicef.org/about/execboard/files/PRG-overview\\_10Mar2014.pdf](http://www.unicef.org/about/execboard/files/PRG-overview_10Mar2014.pdf).

43

RESET Project

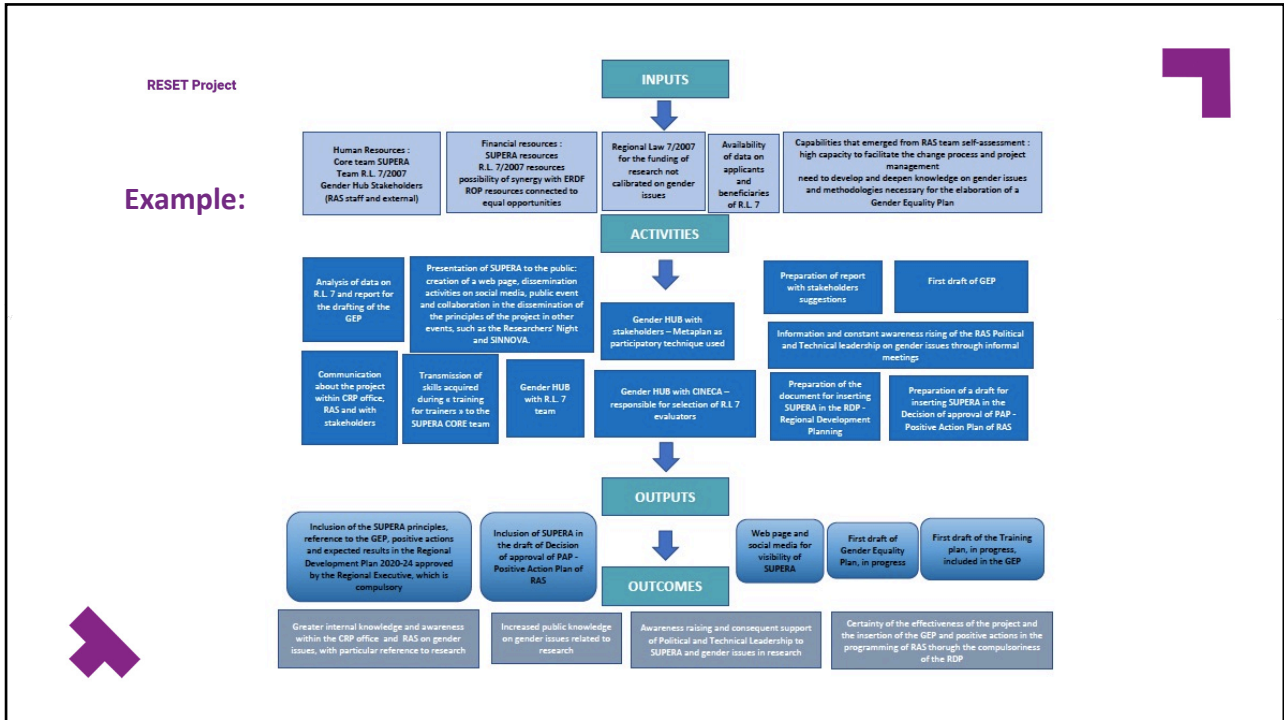
### STEP 3: Setting up a GEP

#### An impact pathway is also a lot about context:



44





45

RESET Project

## And now... your turn! Start devising your own GEP impact pathway

- ✓ **Your inputs:** resources, knowledge, people...
- ✓ **Your activities:** GEP activities (either for its design or planned as part of the GEP itself)
- ✓ **Your outputs:** the expected *immediate* result of the activities
- ✓ **Your outcomes:** the expected *intermediate* result of the activities
- ✓ **Your impact:** what will change?

46