



Work Package 5

Deliverable 5.1: TARGETED-MPI Evaluation Plan

Project Full Title:	Transparent and Resilient Gender Equality Through Integrated Monitoring Planning and Implementation
Call Identifier:	H2020-SwafS-2019-1
Type of Action:	CSA
Start Date:	01 September 2020
End Date:	31 August 2024
Project Number:	872260

Deliverable dissemination level *		
PU	Public	X
CO	Confidential, only for members of the consortium (including EC services)	

Deliverable information	
Work package	5
Work package leader	BUL
Deliverable number	D5.1
Deliverable name	5.1
Deliverable type *	Public Report
Contractual delivery date	31/12/20
Partner responsible for the production of the deliverable	BUL
Author(s)	Mustafa Ozbilgin Nur Gundogdu
Reviewer(s)	ULANC

* R = Document, report; DEM = Demonstrator, pilot, prototype; DEC = Websites, patent fillings, videos, etc.; ETHICS = Ethics requirement; ORDP = Open Research Data Pilot; DATA = data sets, microdata, etc.; OTHER = Other

History of changes			
Date	Version number	Description	Persons involved
30.11.2020	v0.1	First draft	<i>Mustafa Ozbilgin</i> <i>Nur Gundogdu</i>
20.12.2020	v0.2	Second internal draft	<i>Mustafa Ozbilgin</i> <i>Nur Gundogdu</i>
21.12.2020	v0.3	Third internal draft	<i>Mustafa Ozbilgin</i> <i>Nur Gundogdu</i>

24.12.2020	v0.4	First final version for circulation to Project Partners	<i>Mustafa Ozbilgin</i> <i>Nur Gundogdu</i>
31.12.2020	v2.0	Final version after quality control for submission to EC/REA	<i>Mustafa Ozbilgin</i> <i>Nur Gundogdu</i>
21.06.2021	v2.1	Revision with minor correction	<i>Mustafa Ozbilgin</i> <i>Nur Gundogdu</i>

History of reviews		
Review date / period	Version reviewed	Reviewer details (name, organisation)
11.12.2020	V0.1	Konstantinos Zografos
22.12.2020	V0.2	Konstantinos Zografos
24.12.2020	V0.3	Konstantinos Zografos

List of Abbreviations

Abbreviation	Term Description
AHSSBL	Arts, Humanities, Social Science, Business and Law
AUB	The American University of Beirut
AUEB	The Athens University of Economics and Business
EC	European Commission
ECU	Equality Challenge Unit
EDI	Equality, diversity and inclusion
EFORTI	Evaluation Framework for Promoting Gender Equality in Research and Innovation
EQUAL-IST	Gender Equality in Information Science and Technology
ERA	European Research Area
GCI	Glass Ceiling Index
GE	Gender Equality
GEC	GE Committee
GEP	Gender Equality Plan
GSFPGE	The General Secretariat for Family Policy and Gender Equality
HE	Higher Education
HEI	Higher Education Institution
ILO	International Labour Organisation

LGBTQ	Lesbian, Gay, Bisexual, Trans and Queer
NIHR	National Institute for Health Research
OECD	Organization for Economic Cooperation and Development
PILOTINA	Promoting Gender Balance and Inclusion in Research, Innovation and Training
RCGE/KETHI	Research Center for Gender Equality
SAU	South African University
SDG 5	Sustainable Development Goal 5
SDR2020	Sustainable Development Report 2020
SSE	Stockholm School of Economics
STEMM	Science, Engineering, Mathematics and Medicine
STG	Sustainable Development Goals
TARGET	Taking A Reflexive Approach to Institutional Transformation
Targeted MPI	Transparent and Resilient Gender Equality through Monitoring, Planning and Implementation
ULANC	Lancaster University
UN	United Nations
UNDP	United Nations Development Programme
VAW	Violence Against Women
VUB	The Vrije Universiteit Brussel
WEF	World Economic Forum

Contents

Executive Summary	8
1 Introduction.....	9
1.1 Aims of GEP Evaluation Methodologies	11
1.2 Methods of GEP Evaluation Methodologies.....	12
1.3 Dimensions of GEP Evaluation Methodologies	12
2 GEP Evaluation Methodology: Definitions and Method of Research.....	16
2.1 Definitions	16
2.2 Method of Research	16
3 Parameters of GEP Evaluation Methodologies	18
4 GEP Evaluation Indicators and Targets.....	21
5 GEP Evaluation Methodology: Considerations for Research Partner Contexts	33
5.1 The Athens University of Economics and Business (AUEB).....	33
5.2 Lancaster University (ULANC).....	33
5.3 The American University of Beirut (AUB).....	34
5.4 Stockholm School of Economics (SSE)	34
5.5 The Vrije Universiteit Brussel (VUB)	35
References.....	36
Appendix A	40

List of Tables

Table 1: World Economic Forum-Global Gender Gap Report 2020 Sub-indexes and Indicators	18
Table 2: European Institute for Gender Equality's GE Index 2020	19
Table 3: Sustainable Development Goal Report 2020 Gender Equality Indicators	20
Table 4: UNDP Gender Inequality Index 2020	20
Table 5: Athena SWAN Ten Key Principles	21
Table 6: Generic GEP Evaluation Indicators and Targets.....	24

Executive Summary

In this report, we examine methodologies that are developed in order to evaluate Gender Equality Plans (GEPs). Adopting a realist approach to evaluation methodology for GEPs, we explore the what, how, where and for whom (beneficiaries) in designing the GEP evaluation methodology. Realist approach allows us to consider multilevel, multidimensional, multicultural, multi-contextual, multidisciplinary, process relational, embedded, and critical aspects of the evaluation methodology.

Drawing on a systematic literature review, we examine the growing body of literature on GEP evaluation methodologies from both academic sources and research and policy reports. In the introduction section, we first investigate the GEP evaluation methodologies, identify the levels of analyses (i.e., international, supranational, national, sectoral, organisational), explore the range of rationales and aims for which evaluation methodologies have been developed, review the qualitative and quantitative methods used for evaluation methodologies, and finally identify the dimensions of the evaluation methodologies. The introduction section ends with a summary of a range of dimensions in evaluation methodologies and identifies three distinct aspects which should be considered in the future evaluation methodologies: First, the significance of values such as human rights and equality that inform the development of evaluation methodologies. Second, the growing significance of intersectionality, which remains silent previous evaluation methodologies. Third, the critical impact of the Covid19 pandemic on gender relations and the need for developing an evaluation methodology that is sensitive to this impact.

In the second section of the report, we explain our own methods for reviewing the extant academic and policy literature. We provide in this section operational definitions of our terms of reference such as evaluation methodology and gender equality. We specify in this section, the criteria by which we collated the academic and grey literature and our particular methods of review. We also outline how we identified the dimensions of the evaluation methodology drawing on these reviews. Section three identifies the parameters of GEP evaluation method at three different levels: macro, meso and micro level parameters. The report then explores indicators and targets of the GEP evaluation methodologies in section four. We conclude with a custom made evaluation methodology tailored to fit with this TARGETED-MPI project and the unique contexts of its partner organisations in section five.

1 Introduction

The objective of this report is to develop a methodological framework for assessing the performance of the GEPs that will be implemented in each TARGETED-MPI-participating organisation (Task 5.1). We developed a methodology and procedure document, outlining the evaluation methodology and procedures, including the evaluation indicators and targets (Targeted MPI, 2020; p.24).

In this report, we review the extant literature and research reports on evaluation methodologies for Gender Equality Plans (GEPs). Mobilising a realist approach to evaluation methodology for GEPs, we outline the content, context, methods and beneficiaries of the GEP evaluation methodologies. The choice of a realist approach allows us to consider GEP evaluation methodologies from a multilevel and multidisciplinary perspective.

This section defines, reviews and provides insights into aspects of evaluation methodologies for GEPs. In order to achieve this, we reviewed the extant literature on evaluation methodologies and identified common patterns and specific approaches based on level of analyses, aims of the evaluation methodology, methods of evaluation and dimensions of the GEP evaluation methodology. We conducted a systematic literature review, and analysed academic studies, grey literature and smart practices from GEP related funded projects to determine in which levels evaluations occurred. We adopted a realist approach to GEP evaluation methodologies. In line with this approach, we identified that evaluations methodologies differentiate in terms of level, aim, scope, content and data availability.

In a positioning paper, Marra (2020) explains that the interest in evaluation of GEPs has gained an international character in recent years although the regulatory requirements and methodologies have been far from being widespread, universal or standardised. While the GEP evaluation methodologies are discussed internationally, their adoption is contingent upon careful crafting and tailoring that captures local regulatory, cultural and institutional conditions. The riches of the GEP evaluation methodologies have been their multi-faceted, multi-layered and multilevel expansion from international organisations to the organisations of science. GEP evaluation methodologies have been developed in an interrelated way across macro, meso and micro levels. International and supranational organisations such as the UN with Sustainable Development Goals (Sach et al., 2020), OECD with its report on gender and evaluation methods (OECD 2003), The World Economic Forum (World Economic Forum's Gender Gap Index, 2020), European Institute for Gender Equality (European Institute for Gender Equality's Index, 2020) and UNDP (UNDP Gender Inequality Index, 2020) have generated considerable insight into how GEP could be evaluated from a comparative and international perspective, using both qualitative measures and quantitative metrics and with multiple dimensions. These international level GEP evaluation methodologies are of import for developing national initiatives and methodologies for evaluating GEPs as most countries comply with these international reporting requirements.

In line with the promulgation of GEP evaluation methodologies at the international level, there are also national level GEP evaluation methodologies that are developed. Equality

Challenge Unit (ECU) in the UK provides insights into national level methods for evaluating GEP in the higher education sector (ECU 2018). One national level initiative in the UK is the Athena SWAN project, which has a gender equality charter for the UK universities (Athena Swan 2020). Research support organizations can be a catalyst in reducing gender equality. In support for the take up of Athena Swan charter across universities in the UK, UK's National Institute for Health Research (NIHR) has linked the research funding process and results to universities' Athena SWAN performances since 2011. The study by Ovseiko et al. (2020) compares 2006, 2011 and 2016's data to show how this funding incentive increased the ratio of women academics whose grant proposals were funded. Among institutions that received NIHR grants, there were only seven silver awards in 2011 and this number approximately increased to 10 times and became 69. The Athena SWAN incentive had a positive effect on the number of female subject leads but did not have any effect on the number of female research directors. Athena SWAN needs cooperation and participation of stakeholders, enough time, finance, skills and knowledge to be successful. Athena Swan as a national evaluation methodology provides an expansive approach that captures both qualitative and quantitative aspects of women's work in the higher education sector. Athena SWAN is underpinned by the value and belief that gender equity in research can boost social development, sustainability of academic workforce and quality. There are two other examples of national level GEP evaluation methodologies. For example, Bustelo (2003) provided a meta-evaluation of 11 GEPs conducted at the national and regional level between 1995 and 1999 in Spain, which highlighted the political nature of GEP evaluation methodologies. Second, Bühler et al. (2020) evaluated Germany's two flagship programmes named the "Women Professorship Programme " and the "Pact for Research and Innovation", which aimed to increase gender equality in higher education institutions (Bühler et al, 2020; p.7) .

At the organisational level, there are also methodologies developed in order to cater for specific institutional drivers for GEP evaluation. Depending on the particular international and national system in which they are located, each organisation develops specific custom made GEP evaluation methodologies in order to abide by their own bylaws, organisational GEPs, and to comply with the national/international level reporting requirements (Ovseiko et al., 2017; Ovseiko et al., 2020). Loots and Walker (2016) reviewed gender equality policies at South African University (SAU) at sectoral and organizational level. Another organizational level analysis was done at University of California Irvine (Monreo et al, 2008), where 80 women academics participated in the interview to explain how they were experiencing gender equality in the university.

At the organisational level, a number of studies focused on organisational and school level analyses. Ovseiko et al. (2017) explored the medical science department of Oxford University staff's experiences, and perceptions about the impacts of Athena SWAN. Drawing on data collected from a survey and interviews, authors conducted a thematic analysis of 2407 staff (%63) responses to the survey, 523 (%22) responses to free-text comments and 37 interviews. Only 59 of them mentioned the Athena SWAN project in the free text section, which suggests a low level of recognition among staff. Results of the survey and thematic analysis show the positive impact of Athena SWAN, its limitations and some recommendation to improve it. Participants highlighted significant structural and cultural changes that support women

academics' career development. They reported that gendered nature of care responsibilities was considered, and discrimination and bias reduced as a result of this consideration. The study also addressed staff perceptions about some of the Athena Swan's limitations. There was concern about the gender pay gap and power imbalances between genders, intervention for work-life balance, positive discrimination and sustainability of positive changes. This evaluation at the organisational and school level data showed that the GEPs inspired by Athena SWAN presented a good starting point. However, the GEP evaluation also highlighted that the sustainability of gender equality remains a challenge and, organizational and societal changes are needed. The organisational level analyses is often nested in the national context in line with the customs, traditions, regulatory and normative pressures for the national setting as well as the specific requirements and values of the organisation. The emergent nature of the evaluation methodologies at the organisational level suggests that organisations have developed GEP evaluation methodologies to fit with their specific requirements.

1.1 Aims of GEP Evaluation Methodologies

GEP evaluation methodologies serve different purposes and have varied aims including evaluating the effectiveness of a particular GEP, identifying its fit for purpose, documenting the change that has happened through the implementation, investigating the inputs, processes, outcomes and impact of GEPs, and revealing particular gender impact on different stakeholder groups. There are a number of European level funded projects that include a transparent set of aims for their GEP evaluation methodologies. EFFORTI is a EU funded project which evaluates three interventions which integrated gender dimension to higher education's teaching and research contents. EFFORTI evaluation framework was developed in order to compare designs, practices and outcomes of three organizations' (two are from Austria and one is from Spain) gender equality programmes (Palmen et al., 2020)

In the case of South African University (SAU), Loots and Walker (2016) aimed to expose the student values related to gender equality and the interventions that they would need. Ovseiko et al. 's (2017) gender equality evaluation at University of Oxford targeted to expose Athena SWAN's implications, limitations, the staff's expectations and future recommendations. Evaluation of German's two gender equality programmes investigated the impacts of these programs to especially on the number of women academics as a researcher and leadership as well as their publications and citations (Bührer et al., 2020). The meta-analysis of 11 GEP evaluations revealed what are important in evaluation structure and implications (Bustelo, 2003). It also exposed the importance of acknowledging the political nature of evaluation; resource investment; stakeholder participation; institutional structure; timeliness; evaluation criteria; and communication during the evaluation process (Bustelo, 2003). Tatli et al. (2011) question the impact of national level equality policy in higher education institutions and examine the role and power of equality officers in implementing the national level policy. Their study evaluates the trail and challenges of the way national level policy to its institutional implementation.

GEP evaluation methodologies are often custom made with divergent aims inspired by their desires to meet specific institutional demands for auditing, accountability, compliance with

institutional norms and legal regulations. Aims of GEP evaluation include a wide range of rationales such as evaluating the effectiveness of the GEP, its efficiency, fit for purpose, changes that are induced by the GEP, accounting for the antecedents, correlates, processes and outcomes and impacts of the GEP on constituent groups. In the case of the higher education sector, the international, supranational, sectoral conditions and the specific organisational ethos shapes the evaluation methodologies that are developed.

1.2 Methods of GEP Evaluation Methodologies

Evaluation methodologies for organisational and management programmes has been discussed since the 1970s (Kilmann and Herden, 1976). Yet, there had been a paucity of interest in developing evaluation methodologies specifically for GEPs. Development of techniques and methods for GEP evaluation methodologies have been gaining traction in the last 20 years. Realist evaluation is one of the theory-based evaluation approaches and tries to reveal impacts of interventions on outcomes and answer the question “what work, for whom, in what circumstances” (Pawson and Tilley, 1997). The realist evaluation approach was used in Caffrey et al. 's (2016) study with multi method qualitative case studies to understand interventions, implications and outcomes of Athena SWAN programme.

In order to understand GEP evaluation methodologies, it is important to make a distinction between counting and accounting. While counting and quantitative metrics are used for simple headcounts to representation at positions of power and authority and pay gaps, more sophisticated accounting techniques have been developed in order to account for gender inequality (Ozbilgin et al. 2016) using qualitative techniques including policy analyses, interviews, observations, and discourse, theme and content analyses. GEP evaluation methodologies draw on qualitative, quantitative and mixed methods, using a range of data gathering techniques. Drawing on a documentary review and interviews were conducted with 17 head of women agencies and implementers of gender evaluation, and analyses of 11 GEP evaluations, Bustelo (2003) explains that the method of evaluation depends on the role, content, timing and agent of evaluation. A GEP evaluation can be summative (focusing on the outcome) or formative (focusing on the process or development); can evaluate design, process or results (content); can be ex-ante (before intervention) or ex-post (after intervention); can be internal vs external (agent).

1.3 Dimensions of GEP Evaluation Methodologies

GEP evaluation methodologies are multifaceted and complex. While in their early form at the European level, GEP evaluation methodologies contained three key thematic dimensions such as equal treatment, positive action and gender mainstreaming (Rees 2001) programmes, they have recently become more sophisticated in dealing with systemic aspects of gender inequality and empowerment. GEP evaluation methodologies include varied dimensions depending on the specific focus on the GEP and the requirements of the context. Muller et al.'s (2011) extensive review of evaluation of GEPs in the context of higher education institutions show that three central themes were evident: promoting women's scientific

careers, reforming science and management, and gender proofing pedagogy and education. Since the 2010s, evaluation methodologies for GEP have gained further new and challenging aspects.

There are illustrious examples of recent studies which include complex set of dimensions in their GEP evaluation methodologies. The EFFORTI project of the European Union and the UK's Athena Swan project have complex sets of dimensions in their evaluation methodologies. Schmidt and Graversen (2020) show the process of EFFORTI project from the literature review to the final stage of the conceptual model. The project aimed to model a road map for gender equality interventions at team, organizational and system levels at higher education institutions. Gender equality is one of the European Research Area (ERA) goals and the EFFORTI framework was designed to support ERA's three objectives: "increase the number of women in research and innovation; increase the number of women in leadership positions in academia; and integrate the gender dimension in research content and curricula" (Schmidt and Graversen, 2020; p.2). Final conceptual evaluation framework consists of five main categories: "personnel, working conditions, professional capabilities, structural features, research and innovation" (p.9), twenty-one dimensions and forty-six sub-dimensions.

In the EU Horizon 2020 funded project TARGET, Palmén et al. (2020) focused the evaluation methodology on the dimensions of resources allocated for the GEP, resource allocation and definition of gender concept in the design stage of the GEPs. Authors argue that how the interventions are defined and operationalised by the GEPs, have an impact on the outcomes and dimensions that need to be considered as a result. In the implementation stage, managerial level commitment, gender competence, and lack of academic recognition are identified as critical dimensions. Outcomes and impacts of the GEPs were evaluated in terms of increased gender awareness and competence, commitment to teaching and research, structural change and representation of genders, the effect on the professional accreditation process. The study shows that the definitions and operationalization of the GEPs vary within and across interventions. Their findings support the need to tailor make evaluation methodologies around certain principles and values rather than fixed dimensions.

At the national level, the UK's Athena Swan project has a complex set of evaluation dimensions. Ovseiko et al. (2020) detail the evaluation dimensions of the Athena SWAN project across "(a) key career transition points including recruitment, induction, promotion, research excellence, (b) career development, including, training, appraisal and development review, support for career progression, support for research grant applications; (c) flexible working and managing career breaks, including cover and support for maternity and adoption leave; maternity return rate; uptake of paternity, shared parental, adoption, and parental leave; flexible working; a transition from part-time back to full-time work after career breaks and (d) organisation and culture, including culture; human resources policies; the proportion of heads of school, faculty, or school by gender; representation of men and women on committees; participation on influential external committees; committee workload; policies, practices, and procedures; workload model; timing of meetings and social gatherings; visibility of role models; outreach activities" (p.1). The dimensions that Athena Swan have are applied and interpreted by the participating organisations which apply to its charter. Examination of

organisational submissions to Athena Swan suggest varied interpretations of the Athena Swan Charter and dimensions.

GEP evaluation methodologies are highly context specific. Depending on the central focus and partners/stakeholders involved in the GEP, a custom made GEP evaluation methodology should be designed. For example, Loots and Walker (2016) developed a capabilities-based approach, drawing on a mixed-method and intersectional study in order to develop a conceptual framework for gender equality on the base of capabilities. It highlights the importance of gender equality based on social justice. Interviews were conducted with 57 South African University's students, and qualitative data was obtained from 843 students' surveys. In the case of Germany, the scholars defined most relevant indicators from systematic literature and smart practices to create the evaluation framework. They classified the indicators "as input, output, outcomes and impacts at micro, meso and macro-level" (Bührer et al., 2020; p.4). It is also interesting to note that at different national contexts, different dimensions are legitimated drawing on dominant theoretical frameworks and perspectives.

We explore the dimensions of evaluation methodologies further in the following sections. Yet, there are a number of general considerations for development of custom-made evaluation methodologies. Scholars such as Monroe et al. (2008) show that gender equality needs systemic, rather than individual redress. Monroe et al. (2010) also note that what often is presented as a pipeline problem, i.e., absence of qualified and talented women in science, is in fact a problem of systemic gender inequality in construction of merit and talent and treatment and rewarding of women in academia. Although there are a large number of dimensions for GEP evaluation methodologies, we should give more emphasis on the systemic and institutional dimensions rather than individual and numerical representation aspects alone. An excessive focus on individual career development may lead to a deficit understanding of gender inequality (Bourne and Ozbilgin 2008) and an evaluation based on quantitative representation and head counts may omit significant aspects of cultural biases that shape gender inequality (Sayce and Ozbilgin 2014). Espinosa (2013) identifies that it is not only the dimensions of an evaluation methodology that is important but also the values such as human rights and equality which should underpin such an evaluation methodology. She poses the following questions in order to humanise evaluation methodologies: How can we make evaluations more human rights and gender-sensitive? How can we define human rights and gender-sensitive criteria, evaluation questions and indicators? What kinds of evaluative methodologies and techniques help to measure the results of gender mainstreaming strategy? Instead of offering an exact evaluation methodology, Espinosa (2013) proposes that the evaluation team needs to engage with all stakeholders during the evaluation process, and team members should be creatively engaged in finding ways to align evaluation methodologies with values of human rights and gender equality.

Two additional and yet significant considerations are not covered by the current evaluation methodologies. These are the significance of intersectionality in gender equality and the impact of Covid 19 on evaluation methodology. GEP evaluation methodologies are often drawn with limited attention to intersectional character of gender at work. When etc

categories such as age, ethnicity, sexual orientation, socio-economic class, disability and other emic categories that emanate from a deep understanding of relations of power in local settings are excluded from gender analyses, certain gendered forms of oppression and disadvantage remain intact (Tatli and Ozbilgin 2012). Sang et al. (2016) identify that the evaluation of intersections of gender and ethnicity in higher education points to the vulnerable position of migrant women in academia. Therefore, it would be a good idea to consider intersectionality as a dimension in development of future evaluation methodologies for GEPs.

The second elephant in the room is the context of the Covid19 pandemic and the absence of evaluation methodologies which consider the impact of this crisis on gender relations in this period. Recent research on the pandemic shows that it had varied gendered impacts (Alon et al. 2020). As there was not a precedent to the Covid19 epidemic and the crisis response that the higher education institutions have shown, the current formulations of evaluation methodologies for GEPs should be customised to capture the gender impacts of the Covid19 crisis. There are some remarkable dimensions which could help evaluate the gendered impact of the epidemic. These include the work from home, work-life balance, working with virtual teams, changes in pedagogy and teaching context, involvement of women in new forms of research, management, administration, teaching and supervision activities that emerged as a result of the pandemic.

2 GEP Evaluation Methodology: Definitions and Method of Research

2.1 Definitions

GEP evaluation methodology could be defined as a range of qualitative, quantitative and mixed methodologies which are used to evaluate the content, process, and effectiveness of gender equality programmes. GEP evaluation methodologies have been developed by international organisations such as the UN and the OECD, nation states with gender equality agendas, and organisations for their specific gender equality programmes and plans.

At the heart of the GEP evaluation methodologies is the concept of gender equality. For this report, we adopt the following definition of *gender equality*: “the enjoyment of equal rights, opportunities and treatment by men and women and by boys and girls in all spheres of life. It asserts that people’s rights, responsibilities, social status and access to resources do not depend on whether they are born male or female.” (International Labour Organization [ILO], 2000; 91).

A *Gender Equality Plan (GEP)* is a plan of activities, actions, protocols, strategies, dimensions and targets to achieve gender equality in a particular setting. According to the European Commission, *Gender Equality Plan* is defined “as a set of action aiming at: Conducting impact assessment / audits of procedures and practices to identify gender bias; identifying and implementing innovative strategies to correct any bias; and setting targets and monitoring progress via indicators” (European Commission, 2012; p13).

We use the term *evaluation methodology* to include a wide range of qualitative, quantitative and mixed methods and approaches that are commonly used or tailor made to serve specific projects for evaluating a social phenomenon, such as the gender equality plans in this specific case.

2.2 Method of Research

In order to conduct the review that informs this report, we adopted a systematic literature review methodology (Adams et al., 2016), which allowed us to identify, read and review the extant academic and grey literature. To identify evaluation methods and criterias of gender equality, a systematic literature review was conducted. We searched Web of Science, Ebscohost, Google Scholar and JSTOR databases with search terms “evaluation” AND “gender equality”; “assessment” AND “gender equality”; “review” AND “gender equality” in title, abstract and keyword areas. We assessed the titles and abstracts of all records and obtained only full text of related studies. English language articles published between January 1990 and October 2020 were included. In total, we identified 24 articles that contained all the required key words and have direct relevance to the topic.

We extended the review of grey literature by searching the EU Horizon 2020 project databases. We searched the EU Horizon 2020 funded, completed and ongoing projects’ and

published reports which relate to gender equality. We reviewed the projects we identified through this search in terms of their level of methodologies, aims, methods and dimensions. We found the following projects of specific interest to the GEP evaluation methodology: BALTIC GENDER, EFFORTI, EQUAL-IST, PILOTINA and TARGET projects were taken into consideration. We have read and reviewed these projects in detail in terms of their GEP evaluation methodologies.

We have also supported the development of the themes in this report with other supporting literature based on independent searches and exploring reference lists of the identified sources. Once we identified common patterns, levels, methods and dimensions in GEP evaluation methodologies, we have brought in further examples from the broad literature.

3 Parameters of GEP Evaluation Methodologies

We reviewed parameters of some of the often cited significant GEP evaluation methodologies. Below we present in tabular form the key parameters that these GEP evaluation methodologies use. At the macro level, there are a number of international and supranational organizations which have developed parameters for GEP evaluation methodologies. The below parameters draw on different data sources and have different priorities, methodologies and indicators. We outline below the parameters of the GEP evaluation methodologies that they have developed and review the dimensions that they use.

The World Economic Forum’s Global Gender Report is one of the GEP evaluation methodologies that is frequently cited. The report was introduced in 2006. Rather than women’s empowerment, it focuses on the gender equality gap. 153 countries’ data comes from different international bodies as statistics and executive opinion surveys. The Global Gender Gap Report compares countries and regions according to 4 sub-indexes and 14 indicators by scaling them between 0 to 1 (See Table 1):

Table 1: World Economic Forum-Global Gender Gap Report 2020 Sub-indexes and Indicators

Table 1: World Economic Forum-Global Gender Gap Report 2020 Sub-indexes and Indicators
<p>Economic Participation and Opportunity</p> <p>Labour Force Participation Rate (% females to males ratio) Wage Equality For Similar Work (1-7 scale females to males ratio) Estimated Earned Income (int. \$ females to males ratio) Legislators, Senior Officials And Managers (% females to males ratio) Professional And Technical Workers (% females to males ratio)</p>
<p>Educational Attainment</p> <p>Literacy Rate (% females to males ratio) Enrolment in Primary Education (% females to males ratio) Enrolment in Secondary Education (% females to males ratio) Enrolment in Tertiary Education (% females to males ratio)</p>
<p>Health And Survival</p> <p>Sex Ratio at Birth (% females to males ratio) Healthy Life Expectancy (Years females to males ratio)</p>
<p>Political Empowerment</p> <p>Women in the Parliament (% females to males ratio) Women in Ministerial Positions (% females to males ratio) Years With Female Head of State, (Share of Tenure Years females to males ratio)</p>

Adapted from World Economic Forum’s Gender Gap Report, 2020, p. 48.

European Institute for Gender Equality’s Gender Equality Index 2020 measures gender equality among European countries. The index uses different European Union statistics such

as Eurostat and Eurofound and evaluates each country on a scale from 1 to 100. The index is based on 6 domains, 14 sub-domains and 31 indicators (See Table 2):

Table 2: European Institute for Gender Equality’s GE Index 2020

Table 2: European Institute for Gender Equality’s GE Index 2020
Work Participation Segregation and Quality of Work
Money Financial Resources Economic Situation
Knowledge Attainment and Participation Segregation
Time Care Activities Social Activities
Power Political Economic Social
Health Status Behaviour Access

Adapted from European Institute for Gender Equality’s GE Index 2020, p. 2.

Gender equality is one of the Sustainable Development Goals (SDGs) of the United Nations and labeled as SDG5. The Sustainable Development Report 2020 (SDR2020) has measured participating countries performance and trends for 17 SDGs since 2016 (Sachs et al., 2020). Though SDGs were accepted by 193 countries, in SDR2020 the data from official and unofficial data sources are available just for 166 countries. SDR2020 compared gender equality globally, national, regional, sub-regional and income groups and scaled their performance between 0 and 100. The following table shows 6 indicators of SDR2020 for gender equality (See Table 3).

Table 3: Sustainable Development Goal Report 2020 Gender Equality Indicators

Table 3: Sustainable Development Goal Report 2020 Gender Equality Indicators
GENDER EQUALITY-SDG5
Demand for family planning satisfied by modern methods (% of females aged 15 to 49 who are married or in unions)
Ratio of female-to-male mean years of education received (%)
Ratio of female-to-male labor force participation rate (%)
Seats held by women in national parliament (%)
Gender wage gap (% of male median wage)
Gender gap in time spent doing unpaid work (minutes/day)

Adapted from Sustainable Development Report 2020, (Sachs et al., 2020; p. 76)

UNDP's Gender Inequality Index 2020 uses data obtained from international organizations for 189 countries and compares them by countries, human development groups and regions. The index measures gender inequalities with three dimensions and 5 indicators (See Table 4).

Table 4: UNDP Gender Inequality Index 2020

Table 4: UNDP Gender Inequality Index 2020
Reproductive Health
Maternal Mortality Ratio (deaths per 100.000 live births)
Adolescent Birth Rate (births per 1000 women ages 15-19)
Empowerment
Share of Seats in Parliament (%)
Population with at least some secondary Education (% ages 25 and older)
Labour Market
Labour Force Participation Rate (% ages 15 and older)

Adapted from UNDP Gender Inequality Index 2020, p. 316-319.

In terms of parameters of GEP evaluation methodologies, Athena Swan is one of the inclusive and structured evaluation methodology at university and business school level. It provides a framework which has been used globally at university and school level to support women academics careers. Athena SWAN helps universities and schools to change their policies and practices to increase gender equality. The following table indicates its ten key principles (See Table 5).

Table 5: Athena SWAN Ten Key Principles

Table 5: Athena SWAN Ten Key Principles
We acknowledge that academia cannot reach its full potential unless it can benefit from the talents of all.
We commit to advancing gender equality in academia, in particular, addressing the loss of women across the career pipeline and the absence of women from senior academic, professional and support roles.
We commit to addressing unequal gender representation across academic disciplines and professional and support functions. In this we recognise disciplinary differences including: <ul style="list-style-type: none"> - the relative underrepresentation of women in senior roles in arts, humanities, social sciences, business and law (AHSSBL) - the particularly high loss rate of women in science, technology, engineering, mathematics and medicine (STEMM)
We commit to tackling the gender pay gap.
We commit to removing the obstacles faced by women, in particular, at major points of career development and progression including the transition from PhD into a sustainable academic career.
We commit to addressing the negative consequences of using short-term contracts for the retention and progression of staff in academia, particularly women.
We commit to tackling the discriminatory treatment often experienced by trans people.
We acknowledge that advancing gender equality demands commitment and action from all levels of the organisation and in particular active leadership from those in senior roles.
We commit to making and mainstreaming sustainable structural and cultural changes to advance gender equality, recognising that initiatives and actions that support individuals alone will not sufficiently advance equality.
All individuals have identities shaped by several different factors. We commit to considering the intersection of gender and other factors wherever possible.

Source: Athena SWAN (2020)

4 GEP Evaluation Indicators and Targets

The literature identifies a large number of indicators and targets for GEPs at macro, meso, micro levels of analyses. Macro level indicators are the starting point of GEP evaluation indicators and targets of this project. GEP evaluation indicators are drawn from systematic literature review which uses both academic literature and grey literature. Most of the grey literature includes EU Horizon 2020 projects which are completed and ongoing. In particular, we made use of BALTIC GENDER, EFFORTI, EQUAL-IST, PLOTINA and TARGET because of the reports' availability, contents, level of analysis and method of evaluations. European Research Area (ERA) has three-pronged gender equality priorities: "increasing the number of women in research and innovation (R&I), increasing the number of women in leadership positions, and integration of gender dimensions in research content and curricula" Schmidt and Graversen, 2020; p.2). We are building on these priorities with an intersectionality approach that is sensitive to local conditions. BALTIC GENDER (Kisakurek, 2019) project includes 7 quantitative

and 6 qualitative indicators for “career advancement and leadership, resources, decision making, recruitment, work and family, gender in research content” (p.1).

EFFORTI project (Kalpazidou et al., 2019) is based on the ERA priorities. It has 5 categories named “personal, working conditions, professional capabilities, structural features and responsible research and innovation” (p.27) and 21 dimensions and 46 quantitative and qualitative indicators. EQUAL-IST (Canali, 2017) has quantitative indicators for different groups of people (students, academics staffs, non-academics staffs) for “gender equality indicators, degree of work-balance, equal opportunities dedicated machineries/ committees” (p.11) areas and different degrees of comparability (departmental, institutional, national, European level). The project used interview methods to identify qualitative indicators. PILOTINA project (Schwarz-Woelzl, 2020) developed core indicators for five main dimensions: a) the governing bodies, key actor and decision making, b) recruitment, retention and career progress, c) work and personal life integration, d) researcher and research: gender equality and sex and gender perspective, e) integration of gender and sex dimension in teaching curricula. The project also defined 40 specific indicators. TARGET (Chizzola et al., 2018) defined three dimensions to take a reflexive approach to gender equality for institutional transformation: “gender-related institutional barriers to career; gender imbalances in decision making processes; and gender dimensions in research and innovation content and higher education curricula” (p.17). Appendix A provides an exhaustive table which covers parameters and dimensions of the GEP evaluation which are covered by these studies.

Taking forward from these projects and results of systematic literature review, we present at the table below the following indicators and targets distinguished by macro, meso and micro levels and qualitative and quantitative methods. Macro level represents the nation, meso level represents the university and micro level represents the business school. We identify under clusters of parameters a set of indicators in the form of questions at a specific level. We also provide an explanation of the form that the evaluation methodology would take.

In this report, we provide a generic protocol for evaluation indicators. Each partner within the Targeted MPI has very different macro, meso and micro contexts, structures, concerns and priorities for gender equality plans. Therefore, the evaluation procedures related to the implementation of the proposed methodology will gain shape in consultation with partners prior to the three phases of the GEP evaluation: ex-ante, first phase, second phase. In the same way, the evaluation targets will be tailored to the idiosyncrasies of the partner institutions and custom prepared in collaboration with the project partners. The definition of the targets will consider the macro, meso, and micro contextual factors of each partner institution. The decisions regarding the evaluation targets will be partly shaped by the ex-ante evaluation results and in line with discussions with each project partner.

In order to define evaluation targets for each institution, we will engage with the responsible staff from partner institutions and stakeholders. Having received the below indicators, each project partner will furnish the evaluation team with their insights on their specific contexts and concerns of GEP evaluation methodology, which would allow the evaluation team to craft an institution specific protocol for evaluation targets. Partner organizations have different contexts. Therefore, they should have tailored evaluation targets for this project. Table 6

indicates the full range of indicators and targets to be used during these engagements. After exchanges with project partners, we will identify different evaluation targets for each institutions' first phase and second phase evaluations. The evaluation targets of first and second phase will be the same to measure and monitor the GEP implementation effects. We will use indicators in Table 6 in order to establish targets for the ex-ante evaluation of partner institutions. The decisions regarding the evaluation targets will be informed by the ex-ante evaluation results. At the ex-ante, first and second phase of evaluations relevant data collection instruments, i.e. quantitative and qualitative measures and templates will be designed in collaboration with partner institutions and delivered to them after the identification of the specific indicators and targets for each phase of the evaluation process in each partner institution.

Data analysis of these three phases will draw on the aggregate data that is provided by partner institutions. As the BUL has clearance for analyses of aggregate data alone, the partner institutions will provide the qualitative and quantitative data in response to indicators and targets set. Analyses will use content analyses techniques for the qualitative data, and statistical techniques in order to analyse quantitative data. Once, the evaluation targets are set for ex-ante, first and second phases of the project, depending on the choice of targets for each institution, we shall apply a relational analysis technique in order to explore the interface between macro, meso and micro levels. Such an exploration will help the evaluation methodology to specify the level which is responsible for gender equality. As each level of analyses, relates to a different level of responsibility for gender equality. For example, a national level macro policy change could easily alter the meso university level and micro the school level, policy and GEPs. Thus, we indicate the level of analyses in order to highlight the approach that we take to the level of our analyses. Therefore, the macro, meso and micro separation relate to the level and the source of the data that is collected rather than the level of the GEP intervention.

We have the following generic parameters for the GEP evaluation methodology: Policies, structure, numerical representation, gender equality data, decision making/leadership, careers, human resources, research, investment, communication, change, curriculum, innovation, impact, and pay gap. As each partner institution has different organizational structure and temporal arrangements, depending on the specific requirements and temporal data gathering practices in each institution a different level could be agreed at the ex-ante, first and second phase evaluations. For example, some institutions may be gathering annual data, others could be gathering quinquennial or more frequent data on different aspects of GEP. Similarly, some institutions may have department level data and others may only have faculty or school level data. There are also terminological as well as definitional differences in terms of data gathered. The evaluation practice needs to ultimately capture these institutional differences. Therefore, the generic evaluation indicators need to be tailored further to institutional arrangements in partner organisations. We added three new parameters to the evaluation methodology, based on a review of what appears to be missing in the previous GEP evaluation methodologies. These are the values of gender equality and human rights, intersectionality and the gendered impact of Covid19 (See Table 6).

Table 6: Generic GEP Evaluation Indicators and Targets

Level of Indicators	Method/ Type of Indicator	Indicator	Explanation
POLICIES			
Macro-Meso	Qualitative	What are the main national gender equality policies pertaining to the HE sector and business schools?	Details
Macro-Meso	Qualitative	Are the national policies implemented with the university?	Yes/No
Meso	Qualitative	Is the GEP a central concern in the institutional review board of the university?	Yes/No Details
STRUCTURE			
Meso	Qualitative	To what extent GEP is mainstreamed in the structure of the university?	Details
Meso	Qualitative	Is there a specific unit at the university responsible for the implementation of the GEPs?	Yes/No Details
Meso	Qualitative	What are the gender equality policies and structures at the business school?	Details
NUMERICAL REPRESENTATION			
Meso-Micro	Quantitative	What is the degree of vertical sex segregation at the university/business school?	Numbers
Meso-Micro	Quantitative	What is the degree of horizontal sex segregation at the university/business school?	Numbers

Meso-Micro	Quantitative	What is the proportion of women among permanent and temporary academic staff at the university /business school?	Numbers
Meso-Micro	Quantitative	What is the degree of change by gender diversity in numbers based on the institutional monitoring system of your university/business school?	Numbers Detail of time interval of institutional monitoring of gender diversity. Preferably quinquennial.
Meso-Micro	Quantitative	What is the current Glass Ceiling Index (GCI) for the university/business school? $\text{GCI}(\text{She Figures}) = \frac{(\text{WGrade A+B+C}) / (\text{W} + \text{M})}{\text{Grade A} / (\text{W} + \text{M})}$ $\text{GCI}(\text{B-A}) = \frac{(\text{WGrade A+B}) / (\text{W} + \text{M})}{\text{Grade A} / (\text{W} + \text{M})}$ $\text{GCI}(\text{C-B}) = \frac{(\text{WGrade B+C}) / (\text{W} + \text{M})}{\text{Grade B} / (\text{W} + \text{M})}$ <p>where W = number of women, M = number of men</p>	Numbers
GENDER EQUALITY DATA			
Meso-Micro	Qualitative	Which forms of gender data are collected at the university/ business school?	Details
Meso-Micro	Qualitative	Who is responsible for gathering GE data?	Details
Meso-Micro	Qualitative	What is the purpose of collecting the specific gender data?	Details

Meso-Micro	Qualitative	To what extent is GE data used in the university/business school?	Details
DECISION MAKING / LEADERSHIP			
Meso-Micro	Quantitative	What is the proportion of women in governing bodies of the university/business school??	Numbers
Meso-Micro	Quantitative	What is the proportion of women in leadership roles and change over time at the university /business school?	Numbers
Meso	Quantitative	What is the proportion of women in university committees?	Numbers
Meso-Micro	Quantitative	What is the level of gender representation in advisory bodies at the university and the business school?	Numbers
Meso-Micro	Quantitative	What is the proportion of female and male staff in key decision-making roles of the business school/university?	Numbers
Meso-Micro	Quantitative	What are the number and proportion of female lead research projects/grants at the university /business school?	Numbers
CAREERS			
Meso-Micro	Qualitative	To what extent GEPs are considered in the professional development process?	Details
Meso-Micro	Qualitative	To what extent network building activities are developed in line with GEPs?	Details
Meso-Micro	Quantitative	What is the proportion of women within the recruitment pools?	Numbers

Meso-Micro	Quantitative	How has the recruitment of talented women changed over time?	Numbers Preferably quinquennial change.
Meso-Micro	Quantitative	What are gender differences among academic staff based on employee data from the HR department, disaggregated at the university / business school level? if any, a) In job satisfaction, b) In motivation, c) In psychological contract, d) respect, e) recognition, f) harassment, g) bullying, h) Mobbing i) Organizational justice /equal treatment	Numbers
Meso	Qualitative	To what extent the promotion and tenure criteria are transparent, unbiased, inclusive and flexible at your university?	Details
Meso	Qualitative	How does the university build confidence for promotion and responsible positions to support gender equality?	Details
Meso	Qualitative	How does the university support gender equality in order to advance research careers?	Details
Meso-Micro	Qualitative	To what extent academics feel that they are equally treated (based on social justice, and psychological contract metrics) by gender at the university	Details
HUMAN RESOURCES			

Meso-Micro	Qualitative	Does the university/ business school have provision of fair and equal allocation and use of workspace/facilities?	Yes/No
Meso	Qualitative	Does the university have flexible working arrangements which support gender equality?	Yes/No
Meso	Qualitative	To what extent does the university promote compatibility of work and life in support of gender equality?	Details
Meso	Qualitative	Does the university have Work-Life Balance Policies for <ul style="list-style-type: none"> a) Parental Leaves b) Tele-Working c) Flexible Working Hours Arrangement d) Child Care Services e) Summer Camp 	Yes/No
Meso	Qualitative	To what extent Work-Life Balance policies is used at the university <ul style="list-style-type: none"> f) Parental Leaves g) Tele-Working h) Flexible Working Hours Arrangement i) Child Care Services j) Summer Camp 	Details
Meso-Micro	Quantitative	To what degree gender equality is considered in design and delivery of training/human capital development?	Details
RESEARCH			
Meso-Micro	Qualitative	To what extent is gender analysis used in research design and management at the university/business school?	Details

Meso-Micro	Quantitative	What is the gender proportion of researchers involved in funded and coordinated projects at the university/business school?	Numbers
Meso-Micro	Quantitative	Are there any gender differences in scientific outputs at the university / business school?	Yes/No
Meso-Micro	Quantitative	What is the gender distribution of participants in academic networks?	Numbers
Meso-Micro	Qualitative	To what extent gender issues are considered in research and innovation capacities/excellence ?	Details
Meso-Micro	Qualitative	Is gender equality considered in setting research priorities and outcomes?	Yes/No
Meso-Micro	Qualitative	To what extent are GEPs considered in research project plans and implementation ?	Details
Meso-Micro	Quantitative	What is the share of scientific papers including sex/gender variables and dimensions?	Numbers
Meso-Micro	Quantitative	To what extent is gender equality considered in research processes?	Details
Meso-Micro	Quantitative	To what extent is gender equality considered research quality (integration of a gender dimension/perspective in research and content, in research projects, patents, and agreements)?	Details
Meso-Micro	Qualitative	Does the university/business school consider gender equality in providing open access?	Yes/No Details
INVESTMENT			
Meso-Micro	Qualitative	What is the level of funding allocated to promote GE?	Numbers

Meso-Micro	Qualitative	What is the level funding to achieve structural transformation along GE lines?	Numbers
COMMUNICATION			
Meso	Qualitative	Does the university have guidelines on gender sensitive language?	Yes/No
Meso	Qualitative	To what extent are gender equality issues considered in the university's internal and external communications?	Details
CHANGE			
Meso-Micro	Qualitative	Do you have any evidence to show that there has been a decrease of GE barriers at the university?	Yes/No Details
Meso-Micro	Qualitative	Do you have any stories of significant organisational/cultural change with regard to GE?	Yes/No Details
CURRICULUM			
Meso	Qualitative	Do you have guidelines/initiatives for gender sensitive curriculum at the university/ business school?	Yes/No Details
INNOVATION			
Meso-Micro	Qualitative	To what extent gender equality is considered in innovation indicators ?	Details
Meso-Micro	Qualitative	To what extent gender equality is considered in diffusion of innovation in products, services, processes?	Details
Meso-Micro	Qualitative	To what extent is gender equality considered in entrepreneurial activities?	Details
IMPACT			

Meso-Micro	Qualitative	To what extent gender equality is considered in economic impacts?	Details
Meso-Micro	Qualitative	To what extent gender equality is considered in social impacts?	Details
Meso-Micro	Qualitative	To what extent gender equality is considered in environmental impacts?	Details
Meso-Micro	Qualitative	To what extent is gender equality considered in jobs, growth & competitiveness of stakeholders?	Details
Meso-Micro	Qualitative	Does the university /business school consider gender equality in their public engagement?	Details
PAY GAP			
Meso-Micro	Qualitative	What is the level of gender pay gap at your university?	Details
Meso-Micro	Qualitative	To what extent has the gender pay gap closed over the years?	Details Preferably quinquennial
Meso-Micro	Qualitative	What measures does the university take in order to address the gender pay gap?	Details
Meso-Micro	Qualitative	Is gender pay gap data collected at the university?	Yes/No Details
Meso	Qualitative	How does the university measure the gender pay gap?	Details
VALUES			
Meso-Micro	Qualitative	To what extent does your organisation/business school strategy is informed by the value of gender equality and human rights?	Details

Meso-Micro	Qualitative	Are the values of gender equality and human rights explicitly mentioned in values, principles, vision of the university/business school?	Details
INTERSECTIONALITY			
Meso-Micro	Qualitative	To what extent does your university/business school collect data on multiple categories of diversity?	Details
Meso-Micro	Qualitative	To what extent does your university/business school use multiple categories of diversity in designing GEPs?	Details
Meso-Micro	Qualitative	To what extent does your university/business school consider multiple categories of diversity in implementation of GEPs?	Details
COVID19 IMPACT			
Meso-Micro	Qualitative	Does your university/business school have a GEP which considers the Covid19 impact?	Details
Meso-Micro	Qualitative	To what extent does your university/business school monitor the gendered impact of Covid19 on students and staff?	Details
Meso-Micro	Qualitative	What specific measures, if any, has your university/business school taken in order to combat the adverse effects of the pandemic.	Details
Meso-Micro	Qualitative	To what extent does the Covid19 crisis had a gendered impact on work and life of academics at your university/business School?	Details

5 GEP Evaluation Methodology: Considerations for Research Partner Contexts

In this section, we outline our findings on specific and idiosyncratic considerations for research partner contexts. This section is developed based on some highlights from State of Practice Synthesis Reports of partner organisations. These reports provide overviews of National GE Legislation /Policies, GE Policies/Regulations regarding the institution and examples of National GE Studies. In this section, we outline the key aspects of these reports in terms of macro-national, meso-institutional regulatory arrangements and interventions.

5.1 The Athens University of Economics and Business (AUEB)

GE regulations have been on the agenda for Greece since 1983 with the revision of Family Law, which had significant changes made in 2019. The General Secretariat for Family Policy and Gender Equality (GSFPGE) is the primary regulatory body responsible for promoting GE. GSFPGE also leads the Research Center for Gender Equality (RCGE/KETHI). The National Council for GE under the GSFPGE monitors the progress and evaluates the plans with its national, regional and local level committees. RCGE is conducting gender-related research and projects to contribute to gender-based policies and regulations. The Law 4589/Article 33, which went into operation in 2019 is the most significant legislation to promote GE in the higher education sector in Greece. According to the legislation, every HEI in Greece is suggested to have a GE Committee (GEC). Among 25 HEIs in Greece, only 7 of them have a GE Committee. The GE Committee gives advice and support to the governance of the university and administration of schools. Athens University of Economics and Business is the premier university in the field of business and economics in Greece. The AUEB does not yet have any policy regarding GE at the institutional level. However, the newly elected rector has started to form a GE Committee. Gender equality (SDG 5) is one of the Sustainable Development Goals (SDGs), and the university has only a social responsibility plan in line with the SDGs. The university organized different types of activities, but they are limited and they need to be supported. Another critical issue that needs attention is that female academics are under-represent at professor level and decision-making bodies of the university. Structural, cultural changes and adequate resources are required to create a sustainable promotion in GE at the national and institutional level in Greece. Thus, the evaluation will mainly focus on central concerns of the GE activities and plans of AUEB, which relate to representation of women in senior academic posts, decision making mechanisms and institutional arrangements for gender equality.

5.2 Lancaster University (ULANC)

The UK has one of the most regulated legal contexts in Europe in terms of anti-discriminatory legislation. The Equality Act of 2010 combined the previous equality laws on Race Relations Act, Sex Discrimination Act and the Disability Discrimination Acts and broadened the scope with other categories of diversity such as sexual orientation, age, religion and belief among others. The central aim of The Equality Act 2010 is to protect people from discrimination based

on age, disability, gender reassignment, marital status, race, religion or belief, sex, and sexual orientation at work and in life. Most of the UK organisations have an equality, diversity and inclusion (EDI) plan in line with The Equality Act of 2010. Lancaster University (ULANC) is a research-intensive university, with a progressive approach, which is implemented under the strategy document called the EDI Vision 2020. Athena Swan is the main driver for gender equality in the higher education sector in the UK. The university has held institutional level Bronze Athena SWAN since 2008. Though the EDI Committee has promoted gender equality, the gender pay gap in the university reportedly remains over the sector average. The following categories are ULANC's main concerns to design EDI policies for all staff: Parental Leave; Work/Life Management; LGBTQ+ and Trans Equity; Bullying, Harassment, and Sexual Misconduct Policy Parental Support; and Disability, mental health, accessibility, and wellbeing. In terms of the GEP evaluation at the ULANC three central concerns are gender pay gap, intersectional (such as disability, mental health, LGBT, and trans equity) aspects of gender equality and work life balance issues.

5.3 The American University of Beirut (AUB)

At the national level, the Lebanese Labor Law prohibits all types of discrimination between men and women in the workplace except sexual harassment. Gender segregation is considered legitimate in some laws of the country, although there are attempts to remedy this. The country has improved maternity leave conditions. There is an ongoing need in Lebanon in order to decrease discrimination and enhance gender equality. Lebanon has adapted its National Strategy for Gender Equality to the Sustainable Development Goals. The American University of Beirut is one of the oldest and well-known universities in the Middle East Region. The university has some policies regarding gender equality, anti-harassment and discrimination. The Grievance Policy and Procedures, Non-Discrimination and Anti-Discriminatory Harassment Policy and Sexual Harassment Policy are implemented under Title IX Office Initiatives. The GEP evaluation for the AUB will focus on the key concerns of anti-harassment, especially sexual harassment, anti-discrimination and gender equality.

5.4 Stockholm School of Economics (SSE)

Sweden has consistently ranked high in international indices of gender equality. At the governmental level, GE has been discussed since 1993. However, the central, regional and local level governmental implementation of GE started in 2011. The country has appointed a Minister of Gender Equality to promote GE policies and practices. The Discrimination Act provides a comprehensive guide to Swedish institutions. Stockholm School of Economics (SSE) is one of the leading European Business Schools. SSE operates under Swedish law and national legislation. This means the institute follows the Discrimination Act. However, the institution uses the document as a primary and inclusive governing document regarding equality and diversity management. The institution has not had a specific gender equality policy yet. Still, it has an equality and diversity manager, an action plan for Equal Opportunities since 2016 and a guide prepared based on students and employers' perspectives. SSE is currently preparing precise and inclusive equality and diversity policy for 2020-2021 and plans to

implement it in 2021. The followings are the main concerns of the SSE in terms of gender equality: Promoting full-time employment; Violence Against Women (VAW) and its data collection; Gender mainstreaming in Research Institutions Competence development; Tools to combat domination techniques; and Commitment to gender mainstreaming. In terms of GEP evaluation, the central concerns of the SSE are gender issues in career management, glass ceiling, glass cliff and employability.

5.5 The Vrije Universiteit Brussel (VUB)

Belgium is a federal state and power is distributed into the federal, regional and community levels. The federal government designed social and labour laws for all Belgians and has taken into consideration gender equality since the 1980s. There are two institutions directly reporting to the prime minister to advise and support gender equality regulations and policies: The Centre for Equal Opportunities and Opposition to Racism and The Institute for the equality of women and men. The community level is responsible for education and research. The Flemish government's current (2019) documents do not have specific references to gender. The government has recently encouraged diversity in universities. Vrije Universiteit Brussel (VUB) is a Dutch-speaking university in the Brussels-Capital and responsible for the federal governments and the Flemish government policies. VUB developed its first Diversity Plan in 2005 and has used the Gender Action Plan (GAP) since 2014. Current GAP is approved in 2019 and primarily focuses on gender and ethnicity, but it also is applied for other discrimination factors. The GEP evaluation for the VUB would focus on the representational, curriculum and intersectional aspects of gender, ethnicity and other categories of discrimination.

References

- Adams, R., Jeanrenaud, S., Bessant, J., Denyer, D., & Overy, P. (2016). Sustainability-oriented innovation: A systematic review. *International Journal of Management Reviews*, 18(2), 180-205
- Alon, T. M., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. (2020). *The impact of COVID-19 on gender equality* (No. w26947). National Bureau of Economic Research.
- Athena Swan (2020) Athena Swan Charter and Principles, <https://www.advance-he.ac.uk/equality-charters/athena-swan-charter#principles>
- Barnard, S. (2017). The Athena SWAN Charter: promoting commitment to gender equality in higher education institutions in the UK. In *Gendered success in higher education* (pp. 155-174). Palgrave Macmillan, London.
- Bourne, D., & Özbilgin, M. F. (2008). Strategies for combating gendered perceptions of careers. *Career Development International*, 13(4), 320-332.
- Bustelo, M. (2003). Evaluation of gender mainstreaming: Ideas from a meta-evaluation study. *Evaluation*, 9(4), 383-403.
- Bührer, S., Schmidt, E. K., Palmén, R., & Reidl, S. (2020). Evaluating gender equality effects in research and innovation systems. *Scientometrics*, 1-17.
- Caffrey, L., Wyatt, D., Fudge, N., Mattingley, H., Williamson, C., & McKeivitt, C. (2016). Gender equity programmes in academic medicine: a realist evaluation approach to Athena SWAN processes. *BMJ open*, 6(9), e012090.
- Canali, C. (2017). D2. 4 EQUAL-IST Gender Audit Methodology for ICT Research Institutions. <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b20c4d67&appId=PPGMS>
- Chizzola, V., De Micheli, B., & Vingelli, G. (2018). TARGET. Gender Equality Audit Tool-GEAT (D 3.1). https://irihs.ihs.ac.at/id/eprint/4925/1/741672_TARGET_GEAT_D3.pdf
- Coe, I. R., Wiley, R., & Bekker, L. G. (2019). Organisational best practices towards gender equality in science and medicine. *The Lancet*, 393(10171), 587-593.
- European Commission (2012). Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A Reinforced European Research Area Partnership for Excellence and Growth https://eige.europa.eu/sites/default/files/era-communication_en_2012.pdf
- European Institute for Gender Equality (2020). Gender Equality Index 2020. <https://eige.europa.eu/gender-equality-index/2020>

ECU (Equality Challenge Unit) (2018). Monitoring and evaluating impact, https://www.ecu.ac.uk/wp-content/uploads/2018/03/2018-03-ECU_Research-and-data-briefing-5_final.pdf

GÜDLER, J., & REINHARDT, A. (2007). Gender equality in DFG research funding—facts and assessment. *DFG Infobrief*, (1), 1-9.

International Labour Office (2000). ABC of women workers' rights and gender equality. Geneva:ILO https://www.ilo.org/wcmsp5/groups/public/---dgreports/---gender/documents/publication/wcms_087314.pdf

Kalpazidou Schmidt, E., Bühner, S., Schraudner, M., Reidl, S., Müller, J., Palmén, R., ... & Groó, D. (2019). A conceptual evaluation framework for promoting gender equality in research and innovation, toolbox I-A synthesis report.<https://efforti.eu/sites/default/files/2018-03/EFFORTI%20D3.3%20FINAL%20report%2027032018.pdf>

Kilman, R. H., & Herden, R. P. (1976). Towards a systemic methodology for evaluating the impact of interventions on organizational effectiveness. *Academy of Management Review*, 1(3), 87-98.

Kisakürek, B. (2019). Updated handbook of gender-sensitive indicators in the Baltic Gender project. <https://oceanrep.geomar.de/48716/7/MS23-BG-Indicators%20Handbook-V3.pdf>

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

Loots, S., & Walker, M. (2016). A capabilities-based gender equality policy for higher education: Conceptual and methodological considerations. *Journal of Human Development and Capabilities*, 17(2), 260-277.

Marra, M. (2020). A few introductory notes on the special issue on the Evaluation Framework for Promoting Gender Equality in Research and Innovation (EFFORTI). *Evaluation and Program Planning*, 79, 101767.

Mergaert, L., & Minto, R. (2015). Ex ante and ex post evaluations: Two sides of the same coin?: The Case of Gender Mainstreaming in EU Research Policy. *European Journal of Risk Regulation*, 6(1), 47-56.

Monroe, K. R., & Chiu, W. F. (2010). Gender equality in the academy: The pipeline problem. *PS: Political Science & Politics*, 43(2), 303-308.

Monroe, K., Ozyurt, S., Wrigley, T., & Alexander, A. (2008). Gender equality in academia: Bad news from the trenches, and some possible solutions. *Perspectives on politics*, 6(2), 215-233.

Müller, J., Castaño Collado, C., González, A., & Palmen, R. (2011). Policy towards gender equality in science and research. *Brussels economic review*, (2-3), 295-316.

Tatli, A., Nicolopoulou, K., Özbilgin, M., Karatas-Ozkan, M., & Öztürk, M. B. (2015). Questioning impact: interconnection between extra-organizational resources and agency of

equality and diversity officers. *The International Journal of Human Resource Management*, 26(9), 1243-1258.

OECD (2003). Review on Gender and Evaluation
<https://www.oecd.org/dac/evaluation/dcdndep/31736413.pdf>

Ovseiko, P. V., Chapple, A., Edmunds, L. D., & Ziebland, S. (2017). Advancing gender equality through the Athena SWAN Charter for Women in Science: an exploratory study of women's and men's perceptions. *Health research policy and systems*, 15(12), 1-13.

Ovseiko, P. V., Taylor, M., Gilligan, R. E., Birks, J., Elhoussein, L., Rogers, M., ... & Buchan, A. M. (2020). Effect of Athena SWAN funding incentives on women's research leadership. *bmj*, 371.

Özbilgin, M., Tatli, A., Ipek, G., & Sameer, M. (2016). Four approaches to accounting for diversity in global organisations. *Critical Perspectives on Accounting*, 35, 88-99.

Palmén, R., Arroyo, L., Müller, J., Reidl, S., Caprile, M., & Unger, M. (2020). Integrating the gender dimension in teaching, research content & knowledge and technology transfer: Validating the EFFORTI evaluation framework through three case studies in Europe. *Evaluation and Program Planning*, 79, 10175

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage: London.

Rees, T. (2001). Mainstreaming gender equality in science in the European Union: The 'ETAN report'. *Gender and education*, 13(3), 243-260.

Reidl, S., Beranek, S., Holzinger, F., & Streicher, J. (2020). Gender equality regimes and evaluation regimes in Europe and their implications for policy design and evaluation. *Evaluation and Program Planning*, 83, 101860.

Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., Fuller, G., & Woelm, F. (2020). The Sustainable Development Goals and COVID-19. *Sustainable Development Report, 2020*.
https://s3.amazonaws.com/sustainabledevelopment.report/2020/2020_sustainable_development_report.pdf

Sang, K., Al-Dajani, H., & Özbilgin, M. (2013). Frayed careers of migrant female professors in British academia: An intersectional perspective. *Gender, Work & Organization*, 20(2), 158-171.

Sayce, S., & Özbilgin, M. F. (2014). Pension trusteeship and diversity in the UK: A new boardroom recipe for change or continuity?. *Economic and Industrial Democracy*, 35(1), 49-69.

Schmidt, E. K., & Graversen, E. K. (2020). Developing a conceptual evaluation framework for gender equality interventions in research and innovation. *Evaluation and Program Planning*, 79, 101750.

Schwarz-Woelzl, M. (2020). The Set of PILOTINA Indicators. https://zsi-vienna.github.io/PLOTINA/assets/data/PLOTINA_indicators.pdf

Striebing, C., Schmidt, E. K., Palmén, R., Holzinger, F., & Nagy, B. (2020). Women Underrepresentation in R&I: A Sector Program Assessment of the Contribution of Gender Equality Policies in Research and Innovation. *Evaluation and Program Planning*, 79, 101749.

TARGETED-MPI (2020). The Grant Agreement number 872260, Transparent and Resilient Gender Equality Through Integrated Monitoring Planning and Implementation Project Report, European Commission Research Executive Agency. Associated with document ref. Ares (2020) 2600829-18.05.2020.

Tatli, A., & Özbilgin, M. F. (2012). An emic approach to intersectional study of diversity at work: A Bourdieuan framing. *International Journal of Management Reviews*, 14(2), 180-200.

Ud Din, N., Cheng, X., & Nazneen, S. (2018). Women's skills and career advancement: a review of gender (in) equality in an accounting workplace. *Economic research-Ekonomska istraživanja*, 31(1), 1512-1525.

UNDP (2020). Gender Inequality Index <http://hdr.undp.org/en/content/table-5-gender-inequality-index-gii>

Unterhalter, E. (2006, November). New times and new vocabularies: Theorising and evaluating gender equality in Commonwealth higher education. In *Women's Studies International Forum* (Vol. 29, No. 6, pp. 620-628). Pergamon

Van den Brink, M., Benschop, Y., & Jansen, W. (2010). Transparency in academic recruitment: a problematic tool for gender equality?. *Organization Studies*, 31(11), 1459-1483.

Women, U. N. (2013). Gender Mainstreaming Approaches in Development Programming: Being Strategic and Achieving Results in an Evolving Development Context.

World Economic Forum. (2020). Global Gender Gap Report. http://www3.weforum.org/docs/WEF_GGGR_2020.pdf

Appendix A

Indicator	Project
What are the main national gender policies / science policies which are relevant as contextual information for the project and may reinforce the development of the GEP?	TARGET
Is the organisation in line with the current national level of awareness regarding gender equality, as manifested in national gender equality legislation and national funding provided for the implementation of gender equality initiatives?	TARGET
Does the organisation have an ethics code or code of conduct? If so, does it specifically include gender equality as a key asset?	TARGET
To what extent are gender equality issues mainstreamed in the organisation's structure (e.g., is there any dedicated personnel in charge of gender equality issues)?	TARGET
Gender equality policy and structures	PLOTINA
Scissor diagram for academic staff	BALTIC GENDER
Increased number of women in academic and other RTDI positions	EFFORTI
Assessment of GE in Academic Staffs	EQUAL-IST
What is the share of female and male staff in the different research, administrative and management units of the organisation?	TARGET
What is the overall share of female and male staff (human resources; research teams) in the organisation at different levels and if relevant in different disciplines?	TARGET
Glass Ceiling Index (GCI)	BALTIC GENDER
Assessment of GE among students	EQUAL-IST
If applicable: What is the overall share of female and male students (broken down by disciplines and ISCED levels)?	TARGET

Which kinds of gender-disaggregated data are available? Who collects them? For which purpose?	TARGET
Part-time employment	BALTIC GENDER
Sex of the chief scientist on scientific cruises	BALTIC GENDER
Increased number of women in decision- making positions	EFFORTI
Increased confidence and ability of leadership roles	EFFORTI
Women's representation in committees	BALTIC GENDER
Assessment of GE in Governance	EQUAL-IST
What are the organisation's key decision-making bodies, and what is the overall share of female and male staff in them?	TARGET
Representation in (main) governing body(ies), by gender	PLOTINA
Representation in (main) advisory body(ies), by gender	PLOTINA
Increased professional development of work skills (for career success)	EFFORTI
Improvement of network building and use	EFFORTI
Percentage of women within the recruitment process	BALTIC GENDER
Improved recruitment of talented women	EFFORTI
Job Satisfaction- appropriate respect/recognition for (academic/scientific/leadership) work	EFFORTI
Job Satisfaction-Positive individual job rating	EFFORTI
Job Satisfaction-Overall work climate	EFFORTI

Job satisfaction-Allocation of workload	EFFORTI
Transparent, non-biased and flexible promotion/tenure criteria	EFFORTI
Strengthened confidence for promotion and responsible positions	EFFORTI
Improved support to advance research career	EFFORTI
Equal workspace/facilities allocation	EFFORTI
Flexible working arrangements	BALTIC GENDER
Improved compatibility of family and career	EFFORTI
Assessment of Work-Life Balance Degree Within the Institution (Parental Leaves, Tele-Working, Flexible Hours Arrangement, Kindergarten, Summer Camp)	EQUAL-IST
Demand and supply of basic child care	PLOTINA
Provision of advanced child care services	PLOTINA
Provision of services for work and personal life integration	PLOTINA
Standard procedure for parental leave	PLOTINA
Incorporation of gender analysis in research design and management	BALTIC GENDER
Share of funded and coordinated projects, by gender	PLOTINA
Scientific outputs	EFFORTI
Networks	EFFORTI
Training/human capital	EFFORTI

Strengthened R&I capacities/excellence	EFFORTI
Incorporation of GEPs in research project plans and implementation	BALTIC GENDER
Number of scientific papers including sex/gender variables and dimensions	PLOTINA
Increased funding to promote GE	EFFORTI
Increased funding to achieve structural transformation	EFFORTI
Recommendations on gender sensitive language	BALTIC GENDER
Increased gender awareness	EFFORTI
What are the key internal and external communication instruments? To what extent are gender equality issues mainstreamed in the organization's internal and external communications?	TARGET
Gender sensitive language and images in institutional documents	PLOTINA
Decrease of GE barriers	EFFORTI
Organisational/cultural change with regard to GE	EFFORTI
Recommendations on gender sensitive didactics	BALTIC GENDER
Scientific outputs	EFFORTI
Networks	EFFORTI
Training/human capital	EFFORTI
Strengthened R&I capacities/excellence	EFFORTI
Research priorities and outcomes in terms of GE	EFFORTI

Conventional innovation indicators	EFFORTI
Diffusion of innovation in products, services, processes	EFFORTI
Knowledge about sex and gender incorporated into engineering innovation processes	EFFORTI
Economic impacts	EFFORTI
Entrepreneurship	EFFORTI
Strengthened framework conditions for R&I	EFFORTI
Jobs, growth & competitiveness of participants (incl. small and medium enterprises (SMEs))	EFFORTI
Achieved gender equality in research process	EFFORTI
Research quality: integration of a gender dimension/perspective in research and content, in research projects, patents, and agreements	EFFORTI
Contributions to strengthening gender- sensitive research	EFFORTI
Gender equality	EFFORTI
Ethics	EFFORTI
Public engagement	EFFORTI
Science education	EFFORTI
Open access	EFFORTI
RRI/governance	EFFORTI
Research priorities & outcomes in terms of GE	EFFORTI

R&I indicators	EFFORTI
Societal impacts	EFFORTI
Environmental impacts	EFFORTI