

Chapter 2

Relevance of Monitoring for a Reflexive Gender Equality Policy

Angela Wroblewski and Andrea Leitner

Abstract

The TARGET approach aims at establishing a reflexive gender equality policy in research performing and research funding organisations. Monitoring has enormous potential to support reflexivity at both the institutional and the individual levels in the gender equality plan (GEP) development and implementation context. To exploit this potential, the monitoring system has to consist of meaningful indicators, which adequately represent the complex construct of gender equality and refer to the concrete objectives and policies of the GEP. To achieve this, we propose an approach to indicator development that refers to a theory of change for the GEP and its policies. Indicator development thus becomes a reflexive endeavour and monitoring a living tool. This requires constant reflection on data gaps, validity of indicators and the further development of indicators. Furthermore, we recommend the creation of space for reflexivity to discuss monitoring results with the community of practice.

Keywords: Gender monitoring; gender indicators; gender equality policy; policy steering; reflexivity; gender equality plan

**Overcoming the Challenge of Structural Change in Research Organisations:
A Reflexive Approach to Gender Equality, 33–52**



Copyright © 2022 by Angela Wroblewski and Andrea Leitner.

Published under exclusive licence by Emerald Publishing Limited.

This work is published under the Creative Commons Attribution (CC BY 4.0) licence.

Anyone may reproduce, distribute, translate and create derivative works of this book (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode> 26th April 2021, signed by Angela Wroblewski and Rachel Palmén.

doi:[10.1108/978-1-80262-119-820221003](https://doi.org/10.1108/978-1-80262-119-820221003)

Introduction

The TARGET approach to gender equality plan (GEP) development and implementation is based on the complete policy cycle model developed by [May and Wildavsky \(1978\)](#), which emphasises the role of empirical evidence for policy development in general. The starting point for the development of gender equality policies is the gender analysis, which identifies the main gender equality problems. The results of this analysis are used to define the gender equality priorities and goals, which then form the basis for the development and implementation of concrete measures. Both the implementation of these measures and the development of the context should be closely monitored, while the measures themselves should be evaluated by an external body after a given period of time and/or during the implementation phase. This approach is in line with the expectations formulated by the European Commission (EC) in the context of the GEP requirement in Horizon Europe ([EC, 2020, 2021](#)).¹

The steps in the process outlined above hold enormous potential for reflexivity. For instance, the gender analysis is far more than the analysis of gender-segregated data such as the assessment of the representation of women and men in different areas or hierarchy levels and their access to resources. In addition, it should contain a discussion of the underlying gender concept (How is gender defined?), the gender equality objectives (What should be achieved?) as well as assumptions on reasons for gender inequalities (What are the underlying mechanisms?) within the organisation. The latter might be gender stereotypes, which influence criteria used in decision-making or the presentation of the organisation to the public (e.g. webpage, folders). Indicators for the gender analysis and monitoring can support this reflexive process if they go beyond simple sex counting. Careful checks should be made to ascertain if the data or indicators used contain some kind of gender bias or if they strengthen – unintendedly – gender stereotypes. Gender indicators should be based on an explicit gender concept, refer to at least one gender equality objective and provide a measurement that allows an analysis of the development of gender equality in the organisation.

¹The EC formulated a GEP requirement in Horizon Europe. Participants, i.e. public bodies, research organisations or higher education institutions established in a Member State or Associated Country, must have a GEP in place that fulfils mandatory process-related requirements. In concrete terms, the EC requires that (1) the GEP is a public document, formally signed by top management, (2) dedicated resources are provided for gender equality (e.g. funding of a gender equality position), (3) the GEP is based on empirical evidence and monitoring, and (4) training and capacity building are foreseen within the institution (e.g. regarding gender bias). The Commission also formulated five recommended areas to be addressed in the GEP: work-life balance and organisational culture, gender balance in leadership and decision-making, gender equality in recruitment and career progression, integration of the gender dimension into research and teaching content, measures against gender-based violence including sexual harassment.

If gender equality priorities, targets and measures are formulated on such a basis, they will doubtlessly focus not only on increasing female representation but also on eliminating gender bias from structures and processes within the organisation. Monitoring the implementation of such priorities, targets and measures also opens up opportunities for reflection by empirically analysing both the progress towards gender equality and any persistent gender differences (or even backlash), thereby providing food for thought for further discussion. Involving stakeholders in all steps paves the way for an evidence-based gender equality discussion in an organisation, thereby raising awareness and encouraging a deep reflection on both the individual and institutional levels. The results of both the gender analysis and the monitoring should therefore be used to clearly communicate the need for action and the priorities identified.

This chapter discusses the principles of monitoring and gender indicators and presents ways of developing a monitoring system for a tailor-made GEP. These will be illustrated using concrete examples taken from monitoring systems developed in the TARGET project.

Purpose and Principles of Monitoring

The main purpose of monitoring is to provide empirical evidence for the assessment of policy implementation and the reflection on current developments regarding gender equality ([International Labour Organization, 2020](#); [Wroblewski, Kelle, & Reith, 2017](#)). Usually, the monitoring builds on the empirical analysis of the status quo (gender analysis or audit) and its data sources and indicators. It is, however, more than a regular update of the gender analysis. The monitoring itself will represent a further development due to the implementation of concrete policies and possible changes in the context. Therefore, gender monitoring should be interpreted as a living tool and as such be subjected to constant reflection regarding the reliability and validity of its indicators. A measure is reliable to the extent that it produces the same results repeatedly. While no data collection is totally reliable, the aim is always to reduce measurement error as far as possible. A measure is valid to the extent that it measures what it is intended to measure. The latter is of specific relevance in the gender context, an aspect that will be illustrated in the following.

[Markiewicz and Patrick \(2016, p. 12\)](#) define monitoring as:

the planned, continuous and systematic collection and analysis of program information able to provide management and key stakeholders with an indication of the extent of progress in implementation, and in relation to program performance against stated objectives and expectations.

According to [Rossi, Freeman and Lipsey \(1999, p. 192\)](#), monitoring generally involves ‘program performance in the domain of service utilization, program organization and/or outcomes’. In concrete terms, a continuous monitoring of

policy implementation generally pursues four goals, which together support the efficient use of resources:

- Monitoring should provide an overview of current developments in the context of the policy of interest (e.g. number and gender composition of employees or students, number and gender composition of decision-making bodies). Changes in relevant context indicators might influence policy implementation and should therefore be analysed on a regular basis.
- The core function of the monitoring is to provide information about policy implementation (e.g. number of policies implemented, number of participants in training programmes and share of women, number of beneficiaries of subsidies and share of women, budget spent on specific measures).
- The monitoring aims at identifying deviations between planned and actual policy implementation, which may indicate ineffective policy implementation or unrealistic policy assumptions. If such problems are detected at an early stage, they can be counteracted by adapting the policy or its implementation.
- In an ideal scenario, the indicators used in a monitoring system also provide the basis for policy steering. For example, when performance agreements between a university and the government or within a university (e.g. between the rectorate and the faculties) contain gender equality objectives, which are related to indicators, these indicators should be formulated in a way that corresponds to specific gender equality objectives.

In general, the monitoring mainly addresses two groups, who should act on its results. The first is management, which takes monitoring data into account when deciding on the continuation, termination or adaptation of policies. The second are the people implementing the policies, who should use the monitoring results to reflect on and optimise implementation as required.

To serve its purpose, a monitoring should be tailored to the concrete context of an organisation and its gender equality policies. The aim is not to provide lots of data (data cemetery) but data that are analysed on a regular basis. Accordingly, efficient monitoring should be based on the following principles (see also [Wroblewski et al., 2017](#)):

- Monitoring systems are based on data that are available on a regular basis and easily accessible. In most cases, monitoring indicators consist of quantitative indicators that are derived from the main objectives in a policy field. However, objectives cannot always be formulated in a quantifiable manner. In such cases, qualitative indicators should be included.
- A monitoring system should include indicators that describe the context of the policy or measure, its implementation as well as the expected output or outcome.
- Indicators focusing on the implementation of policies should be derived from a logic model or programme theory that has been explicitly formulated for the concrete policy.

- Monitoring indicators should be developed with the participation of the main stakeholders. The aim is to establish an agreed set of indicators that all relevant stakeholders accept as meaningful and relevant. This agreed set of indicators should likewise be based on a data source that all stakeholders define as reliable.
- The agreed set of indicators should be analysed at regular intervals (e.g. yearly or monthly). The timing should be linked to the planned intervals for presentation and discussion of monitoring results (e.g. in the form of annual or monthly reports). Regular presentation of monitoring results will both contribute to a gender equality discourse within the organisation and provide the basis for organisational learning.

Even if monitoring provides a basis for the assessment of policy implementation, it still has to be distinguished from evaluation. Monitoring is the systematic documentation of key aspects of policy implementation that indicate whether the policy is functioning as intended or adhering to some appropriate standards. In contrast, evaluation is

the systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy. (Weiss, 1998, p. 4)

Since an evaluation usually takes place after a certain period of policy or programme implementation, it conveys an ex-post perspective. If the evaluation is performed in parallel to implementation, it is referred to as an ongoing evaluation that is characterised by blurred boundaries between monitoring and evaluation. However, while monitoring is carried out internally, evaluation aims at providing an external view on implementation. An evaluation can be commissioned by those implementing the policy or programme or by a superior authority (e.g. a state authority in the case of state-funded policies).

Monitoring and evaluation are complementary approaches. The complementarity can take different forms (Markiewicz & Patrick, 2016, p. 17): The relationship is sequential when monitoring generates questions to be answered in an evaluation or evaluation identifies areas that require future monitoring. It is informational when monitoring and evaluation draw on the same data sources but ask different questions and frame different analyses. It is organisational when monitoring and evaluation draw on the same data sources, often channelled through the same administrative unit. It is methodological when monitoring and evaluation share similar processes and tools for obtaining data. It is hierarchical when performance data are used by various hierarchies, sometimes for monitoring and sometimes for evaluation. Finally, it is integrative when both approaches are designed at one time, unified and draw on a shared monitoring and evaluation framework. Regardless of the concrete relationship, monitoring and evaluation functions are integral to the effective operation of policies and programmes and increase the overall value they create.

Gender Indicators

The monitoring of a GEP ideally contains indicators that allow the assessment of its implementation as well as its outcomes. Hence, the monitoring is composed of gender indicators. Gender indicators do not represent gender equality per se. As gender equality is a complex construct, a gender indicator can only be an approximation. As Beck (1999, p. 7) puts it:

An indicator is an item of data that summarises a large amount of information in a single figure, in such a way as to give an indication of change over time, and in comparison to a norm.

Hence, indicators differ from statistics: the latter merely present facts while the former involve comparison to a norm and interpretation. A gender indicator is thus an indicator that captures gender-related change over time.

The deviation between the indicator and the construct to be measured has to be reflected on and considered in the interpretation. In this context, the conceptualisation of gender and its equivalent in empirical evidence is of specific relevance. While gender is seen from a theoretical point of view as socially constructed (Butler, 1990; West & Fenstermaker, 1995; West & Zimmermann, 1987), it is usually coded dichotomously in administrative data (female/male). Accordingly, the variable *sex* or *gender* available in empirical data does not provide information about gender (Döring, 2013; Hedman, Perucci, & Sundström, 1996; United Nations Economic Commission for Europe (UNECE) & World Bank Institute, 2010). In addition, sex and gender interact with each other, for example, when the male body was the main reference in human medicine and clinical trials were conducted primarily by men, or when gender research in the 1960s focused mainly on women and was mainly conducted by female researchers (Stefanick & Schiebinger, 2020). Gender refers to norms, behaviours and roles associated with being a woman, man, girl or boy, as well as their relationships with one another. As a social construct, gender can change over time. Furthermore, both sex and gender produce inequalities that intersect with other social and economic inequalities. Hence, when discussing gender-based discrimination, gender intersects with other factors of discrimination such as age, socioeconomic status, disability, ethnicity, gender identity and sexual orientation (van der Haar & Verloo, 2013; Verloo, 2006; Walby, Armstrong, & Strid, 2012). To approach this complex construct in empirical analysis, the variable *sex* is differentiated by other relevant variables – if these are available. The availability of information on other relevant characteristics like disability, care responsibilities or gender identity is the exception rather than the norm. The assumption that specific characteristics like care responsibilities mainly apply to women may lead to an unintended emphasising of gender stereotypes and supports the identification of discrepancies as gender-based even though they are based on other characteristics (Degele, 2008; Stadler & Wroblewski, 2021). This problematic aspect gains additional relevance because available data might be gender biased, especially in the case of administrative data. The production of administrative data tends to overrepresent realities, which are male dominated. This becomes a problem if such data are used for analysing gender

imbalances, for example, when labour market statistics are used to analyse gendered patterns of employment because official statistics only consider paid employment (Criado-Perez, 2019; D'Ignazio & Klein, 2020; Hedman et al., 1996).

Gender indicators are not merely statistics on men and women. They highlight the contributions of men and women to society and (in our context) to science and research as well as their different needs and challenges. To depict this complex picture adequately, a set of indicators that covers all relevant aspects is required. The interpretation of one isolated indicator may be misleading. In the context of gender equality policies, the monitoring has to contain indicators, which address all three main gender equality objectives. In other words, it must contain indicators about women's representation in all fields and at all hierarchical levels, indicators that represent structural barriers for women (such as women's participation in decision-making) and indicators that display the integration of the gender dimension into research content and teaching.

Data availability differs for these three dimensions, which in turn affects the validity of indicators. It is easier, for example, to depict women's representation than it is to show the gender dimension in research content and teaching (see EC, 2018, 2019a, 2019b, 2019c, 2019d). In most cases, the availability of data on objective, gender-balanced representation in all fields and at all hierarchical levels is quite good. Education establishment knows the gender composition of students and staff in different disciplines as well as in decision-making bodies. Information on the share of women at different hierarchical levels is likewise usually available. Data availability is not so common when it comes to structural barriers for women's careers. Information on the representation of women at different stages in appointment procedures, for instance, is not available by default. The availability of data on the integration of the gender dimension into research and teaching content is generally limited.

Different data sources – such as administrative data that is electronically available (e.g. student or staff records) or project/publication repositories (to identify projects and publications with gender content) – are likewise relevant for monitoring. However, it is not always possible to extract gender-relevant information from electronic data management systems (e.g. in the context of recruitment). Hence, the development of indicators for gender analysis or gender monitoring often requires an adaptation of existing data sources, the establishment of new data collection mechanisms and specific data collection (e.g. a survey). Indicators can be either quantitative (e.g. number, percentage, ratio) or qualitative (e.g. assessment in qualitative terms). Regardless of their type, indicators should always be SMART² (Doran, 1981). Ideally a combination of qualitative and quantitative approaches will be used to compensate for the shortcomings of both approaches (e.g. Flick, 2018; Mertens, 2017).

²SMART indicators are specific (i.e. should be precise and focused, not a combination of multiple things), measurable (i.e. there should be a practical and undisputed means of measuring), achievable (i.e. should not refer to something that is beyond the means of achievement), realistic (i.e. should not be vague and hardly make sense) and time bound (i.e. should not consider the situation over an indefinite period).

The previous comments point to three key aspects of indicator development: First, it is important to use a consistent gender construct. Second, indicators should be derived from gender equality objectives and targets. Third, data collection is not an end in itself but should contribute to the purpose of monitoring. In the following, we will illustrate these aspects in reference to institutional context indicators and indicators addressing policy implementation for the three gender equality objectives.

Institutional Context Indicators

Institutional context indicators allow a description of the status quo of gender equality in the institution and provide the main information about the institution needed to interpret developments and changes properly. For a proper interpretation of these indicators, further information on the context is required (e.g. number of staff and students, number of management positions and decision-making bodies or number of new appointments). Changes in the share of female professors, for instance, should be interpreted with caution when the institution only has a few professorial positions. In such a case, one newly appointed woman or one retiring woman can have a big influence on the share of female professors. Furthermore, the interpretation of a lack of change requires information on the number of appointment procedures in the respective period. In the case of research funding organisations (RFOs), institutional context indicators refer to their core task, namely funding. These can include the number of calls or funded projects, the budgets available for funding or the number and composition of review panels.

Institutional context indicators describing the status quo of gender equality are usually also used to measure outcomes. They should represent all three gender equality dimensions addressed in the GEP. [Table 2.1](#) provides concrete examples for such indicators for research performing organisations (RPOs) and RFOs.

Table 2.1. Examples for Institutional Context Indicators for RPOs and RFOs.

	RPOs	RFOs
Gender balance in all disciplines and at all hierarchical levels	Share of women in disciplines (students, staff) and hierarchical positions	Share of women among applicants Share of female principal investigators
Decision-making	Share of women in decision-making bodies	Share of women among evaluators Share of women in RFO decision-making bodies
Gender dimension in research and teaching content	Share of research projects that address the gender dimension Share of teaching courses that consider the gender dimension	Share of research projects that address the gender dimension

Source: own research.

Indicators for Policy Implementation

Examples for indicators that focus on the implementation of policies can include the number of participants in programmes, the budget spent on programme implementation or the number of complaints addressed to an equality officer. A meaningful indicator for the monitoring of policy implementation should be derived from the concrete objective of the GEP or the concrete policy. In the course of policy development, a logic model (W.K. Kellogg Foundation, 2004) or theory of change (Funnell & Rogers, 2011) should be formulated, which explains the underlying assumptions on why the policy is expected to reach its target groups and objectives.

Following this approach, the starting point for indicator development are the objectives, activities and targets formulated in the GEP. The objective is what is to be ultimately achieved, the final form or situation we would like to see. But it also has to be clearly distinguished from a vision. A vision can be idealistic; a goal must be more realistic. An organisation will ideally have a fixed vision that does not change over time. However, it can have different objectives and targets that are periodically adjusted to the vision.

In most cases, and given their different purposes, it makes sense to differentiate between monitoring and evaluation targets. The targets formulated in the GEP relate to a strategic level or in evaluation terms to the impact. Monitoring targets generally refer to the implementation level, that is, the desired outputs of policies or measures (e.g. 100 employees should receive gender competence training in a specific year). They also need to be formulated for time spans that are covered by the monitoring (data collection dates/frequencies, e.g. annual, biannual). Evaluation targets, in contrast, refer to the impact or level of outcome. Indicators for this level cannot be measured in short frequencies (e.g. monthly or even biannually), and it is therefore of no practical use to set such short evaluation intervals. Targets at each level should be set at the same frequency/period as was planned for their measurement. Accordingly, targets at outcome level (for evaluative purposes) should ideally be set at three- or five-year intervals.

The dimensions which monitoring indicators should represent also apply to the outcome or evaluation level. However, achieving the desired outputs does not necessarily result in achievement of the expected outcomes. Although this should logically be the case, assumptions that the measures should work can prove to be wrong, or unexpected circumstances can arise, which might affect outputs or outcomes.

The assumptions as to why interventions should lead to their expected outcome are usually formulated in a theory of change or programme theory.

A program theory is an explicit theory or model of how an intervention, such as a project, a program, a strategy, an initiative or a policy, contributes to a chain of intermediate results and finally to the intended or observed outcomes. (Funnell & Rogers, 2011, p. xix)

The formulation of a theory of change allows lessons to be learned from failure and success and by referring to monitoring results. Reflections on policy or programme implementation based on monitoring can lead to an adaption of

Table 2.2. Examples for Visions, Objectives and Targets.

Visions	Objective	Evaluation Targets at Impact Level	Monitoring Targets
Structural barriers for women's careers are abolished	To foster equality in recruitment practices	Increase the share of women among newly appointed professors up to the share of women among applicants	Increase the share of women among newly appointed professors to $X\%$ by Y (date)
Women and men are equally represented in decision-making	To foster gender balance in decision-making committees and boards	Increase the share of women in decision-making committees and boards	Increase the share of women in board X to $X\%$ by Y (date). Increase the share of gender-balanced committees to $X\%$ by Y (date)
All research projects consider the gender dimension in content in all stages of the research process	To promote the integration of the gender dimension into research and innovation	Increase the share of research projects that consider the gender dimension in their content Increase the share of reviewers with gender competence or expertise	Fund X (#) research projects that consider the gender dimension in their content per year $X\%$ of all reviewers received gender training in year Y

Source: own research.

objectives or the implementation framework. The theory of change defines the central processes or drivers by which change is expected to come about for the organisation or the target group. The assumptions on which the theory of change is based could be derived from a formal research-based theory or an unstated, tacit understanding about how things work. A simplified representation of a theory of change is the logic model.

The program logic model is defined as a picture of how your organization does its work – the theory and assumptions underlying the program. A program logic model links outcomes (both short- and long-term) with program activities/processes and the theoretical assumptions/principles of the program. (W.K. Kellogg Foundation, 2004, p. III)

The logic model is merely a simplified representation of mechanisms that lead to the expected outcome and impact because it does not consider feedback loops

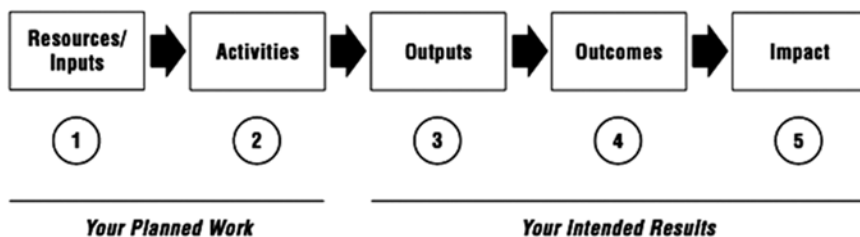


Fig. 2.1. Logic Model (Source: W.K. Kellogg Foundation (2004, p. 1)).

or nonlinear relations. However, referring to a theory of change when developing policies and monitoring indicators forces responsible stakeholders to think carefully about the concrete objectives and targets of an intervention and be realistic about the expected outcome given a specific input. Table 2.3 provides example input and output indicators for the three gender equality dimensions.

Referring to a logic model supports the formulation of consistent and coherent policies and reduces the risk of failure due to unrealistic expectations that implementation cannot meet. It also provides criteria for the success and failure of policies (Engeli & Mazur, 2018). To illustrate this, we will now look in more detail at how the logic model can be applied to quotas for decision-making bodies.

Table 2.3. Examples for Implementation Indicators.

Policy/Programme Aim	Input Indicator	Output Indicator
Abolishment of structural barriers for women’s careers	Share of job advertisements that are formulated in gender-sensitive language Share of selection committee members who participated in anti-bias training	Share of women among newly appointed staff in relation to the share of female applicants
Gender balance in decision-making	Number of gender competence training measures for members of decision-making bodies	Share of women in newly established decision-making bodies
Integration of gender dimension into research content and teaching	Share of researchers who participated in awareness-raising or training measures focusing on the gender dimension in research content	Share of research projects that formulate gender-specific research questions (self-assessment)
	Share of teachers who participated in training measures focusing on gender-sensitive didactics	Share of courses with literature focusing on relevant gender issues in in their syllabus

Source: own research.

Example: Logic Model for Quotas for Decision-Making Bodies

Gender equality policies in academia have long been based on the critical mass theory formulated by [Kanter \(1977\)](#), in which it was assumed that cultural change will take place when women's representation in an organisation exceeds a certain benchmark (the so-called critical mass). Experience has shown, however, that this does not automatically take place: women's underrepresentation in top positions in particular remains unchanged. Hence, specific instruments have been introduced to support women on their path to top-level positions. Quotas have proved, for example, to be an efficient instrument in increasing women's representation in decision-making in academia ([Lipinsky & Wroblewski, 2021](#); [Voorspoels & Bleijenbergh, 2019](#)). [Table 2.4](#) shows a logic model for a quota regulation for decision-making bodies to increase women's representation in decision-making.

At first sight, quotas look like an intervention with a clearly defined objective: They aim at increasing the representation of the underrepresented sex in a specific group like a decision-making body. However, a second look reveals another, often implicit objective: Quotas should also lead to less gender-biased or more women friendly decisions. This implicit assumption has led to critique of the implementation of quota regulations and their effects (e.g. [Guldvik, 2008](#); [Meier, 2008](#); [Sacchet, 2008](#); [Storvik & Teigen, 2010](#); [Törnqvist, 2008](#); [Voorspoels & Bleijenbergh, 2019](#)). [Childs and Krook \(2008\)](#) suggested differentiating between numeric (share of women in decision-making bodies) and substantive (considering women's concerns in decision-making and abolishing a gender bias in decision-making procedures) representations of women. Hence, if a quota regulation pursues both objectives and addresses them both with targeted measures, two logics will need to be formulated to achieve a meaningful monitoring. [Table 2.5](#) shows a logic model for specific anti-bias training for members of decision-making bodies.

Interpretation and Further Development of Monitoring and Indicators

The indicators integrated into the monitoring should be interpreted regularly, for example, on an annual basis. Ideally, the interpretation intervals will be compatible with the policy cycle, for example, the policy implementation period. When interpreting an indicator, it is necessary to define its underlying norm. This normative element allows the identification of failure or success. The share of women in decision-making bodies alone does not provide any information if the concrete value has to be interpreted as positive or negative. It is possible to define several benchmarks and, in most cases, multiple perspectives on indicators are relevant. First, the value can be interpreted over time, so the focus lies on developments since the last measurement. Second, the value of a specific group can be compared with a relevant comparison group (e.g. the situation of female PhD students is compared with that of male PhD students). Third, the interpretation of an indicator refers to an external benchmark like the national average or the corresponding result for an organisation that has been identified as a role model or as having good practice policies.

Table 2.4. Logic Model for Quotas for Decision-making Bodies (Numeric Representation).

	① Resource/Input	② Activity	③ Output	④ Outcome	⑤ Impact
Intervention	A guideline/policy for the composition of decision-making bodies is formulated	The guideline is approved Staff members are informed	Staff members know and endeavour to comply with the guideline The guideline has been implemented	The composition of decision-making bodies meets the target quota At least X% (target quota) of members of a decision-making body are female	Women participate in decision-making as a matter of course Decision-making positions are equally accessible for women and men
Target	A guideline is formulated Information material is available	All staff members are informed about the guideline			
Indicator	Yes/No	Description of communication process Number of staff members who have been informed	Number of staff members who know and comply with the guideline	Share of women in decision-making bodies Share of decision-making bodies that meet the quota	Share of women in decision-making bodies vs. share of women among staff members

Source: own research.

Table 2.5. Logic Model for Anti-bias Training for Members of Decision-making Bodies (Substantive Representation).

	① Resource/Input	② Activity	③ Output	④ Outcome	⑤ Impact
Intervention	Seminar concept, target group, trainers/experts are formulated	Selection process, seminar or workshop held	Completed seminars	Participants carry out decision-making in a more gender-competent manner	Decision-making bodies behave differently
Target	Concept is developed, trainers are available, target group is invited	Seminars/workshops are held, according to schedule	Participants complete training as expected	Participants apply the training content in their everyday work	Decisions are made without an implicit gender bias
Indicator	Yes/No	Number of seminars	Number of participants by gender and other relevant criteria (e.g. target group)	Number of participants who apply the training content in their everyday work	Share of women at different stages of appointment procedures

Source: own research.

An indicator can also have limitations when it comes to the underlying construct it is intended to represent. This is the case, for example, when sex-disaggregated data is used for gender analysis. Recognising these limitations is necessary for understanding the validity of an indicator and should be explained in the analysis. Lack of data often proves to be an issue in this context. If the only data available is sex-disaggregated data that cannot be differentiated by other relevant variables, these limitations have to be considered in the interpretation. This must be done not only for the sake of clarity but also to avoid an interpretation of discrepancies between men and women as gender gaps even if they might be due to other factors (e.g. care responsibilities).

Filling existing data gaps through specific data collection or further development of administrative data sources can be formulated as an objective of a GEP. Indeed, the analysis of the monitoring may raise new questions, and changes in policy design may lead to an adaptation of the monitoring indicators. Hence, the monitoring should be interpreted as a 'living tool'. According to [Hedman et al. \(1996, p. 11\)](#) 'the production of gender statistics is a never ending process. It is a continuous process of integrating developments and improvements of gender statistics' into the monitoring system'.

Creation of Space for Reflexivity

The TARGET project assumes that the implementation of a GEP is a long-term project that requires constant reflection on the development of gender equality, the formulated objectives and targets as well as the proposed measures (Wroblewski & Eckstein, 2018). Like the process itself, objectives, targets and measures may be adapted to reflect changes in context, progress or a more in-depth understanding of the problem at hand. For example, one of the implementing institutions in the TARGET project collected information on female participation in its panel discussions for the first time. The members of its community of practice (CoP) were surprised by the significant underrepresentation of women, which led in turn to a discussion of underlying mechanisms and the formulation of a policy aiming at gender-balanced panels.

The monitoring results provide a starting point for a reflexive process that aims at increasing awareness of gender issues and building up gender competence as well as early counteraction in the event of suboptimal implementation. These two functions of monitoring should be differentiated. To initiate a gender equality discourse within the organisation, a format for discussing the monitoring results internally must be found. This requires the internal publication of monitoring results and a discursive format (e.g. a presentation or workshop) with the CoP. The discussion of monitoring results within the CoP should be seen as part of an organisational learning process ([Hallensleben, Wörlen, & Moldaschl, 2015](#); [Moldaschl, 2007](#)) and take place in an atmosphere of openness and trust. For the institutions participating in TARGET, the development and implementation of the GEP is their first attempt to pursue gender equality goals in a structured, consistent and coherent manner. It can therefore be assumed that some of the planned measures will not achieve their objectives or that the underlying assumptions behind measures will prove unrealistic. Failed attempts also provide useful

lessons learned that are of relevance for the evolution of existing measures or development of new ones. It should be clear that – even if objectives are not reached immediately – gender equality goals will remain a priority. Failure should not result in sanctions but should be turned into constructive lessons learned. This is part of the top management commitment.

Hence, the aim is not to challenge single gender equality policies or the GEP as such but to identify success and failure as starting points for their further development. Ideally, this reflection at institutional level is linked to reflexivity at individual level (Martin, 2006; Wroblewski, 2015). The discussion should aim at supporting CoP members in reflecting on their individual contribution to gender equality, detecting gender bias in their field of responsibility and developing unbiased alternative practices. Since not all members of the CoP are gender experts, the discussion within the community can contribute to raising awareness. However, gender experts should be involved in the development of alternative practices.

Spaces for reflexivity have to be specifically prepared and supported, for example, by providing a workshop moderator who is able to facilitate an open and trusting discussion, activate participants and initiate reflexivity. The gender equality discourse emerging from reflexive practices should also be used to obtain commitment for gender equality goals from all members of the organisation. This is another aspect of the top management commitment: requiring gender-competent action from all staff members within their field of responsibility (e.g. teachers in the teaching context, administrators in their administrative tasks, researchers in the context of research projects).

United Nations Economic Commission for Europe (UNECE) and World Bank Institute (2010, p. 127) recommend the use of gender indicators for communication and awareness-raising activities.

Gender statistics are valuable only if they are used to assist in understanding of gender issues. Communication is needed to encourage their use and illustrate their value to different users.

It is important in communication activities to identify the different target groups of the message and develop specific communication strategies if appropriate. One such target group is the CoP (including management) with the main aim of discussing monitoring results as part of an internal gender equality discourse. In the event that not all information obtained through the monitoring is suitable for distribution, a specific report should be developed to be distributed within the organisation and beyond. This could take the form of an annual publicly available gender report that presents the organisation as gender-sensitive and demonstrates its commitment to gender equality as well as any related progress. A gender report can also contribute to a national or regional gender equality discourse.

Conclusions

Monitoring aims at providing empirical evidence regarding developments in gender equality and GEP implementation that can be used to assess policy implementation, support policy steering and raise awareness about gender issues. As already

discussed, empirical evidence plays a crucial role for effective GEPs because a comprehensive gender analysis provides the basis for the development of GEPs and policies that address gender imbalances and the underlying mechanisms. If this stage is omitted or remains superficial, policies are at risk of becoming actionistic (Wroblewski, 2021) or being based on an inadequately formulated programme theory (Engeli & Mazur, 2018). Policy development that is not based on a sound analysis of the problem in hand risks ineffective policy implementation, wastes resources and will not contribute to change. However, even when policies are based on an empirical gender analysis, a lack of monitoring can also lead to ineffective implementation. Ideally, monitoring will reveal difficulties in correct policy implementation at an early stage (e.g. problems in reaching the target group, budgetary deviations from the plan). Hence, empirical evidence that is discussed in the CoP contributes to effective GEP development and implementation in several ways.

An evidence-based discussion in the CoP on the status quo of gender equality contributes to a shared understanding of the gender equality problem as well as a broad acceptance of the GEP and its objectives. An evidence-based approach is in line with the logic and self-image of an academic institution. Monitoring has the potential to maintain this acceptance of gender issues and the GEP. However, specific actions must be taken to support the acceptance of the monitoring, for example, by explicitly formulating and communicating the role of the monitoring to the CoP or by linking the gender monitoring to existing monitoring systems in the organisation (e.g. quality management or performance measurement). Empirical evidence contributes to creating awareness of gender inequalities and defines topics to be addressed in the context of a GEP. There is a tendency to think that only ‘what gets counted counts’ (D’Ignazio & Klein, 2020, p. 97) or that our ‘world is generated by numbers’ (Heintz, 2012) because the description of social phenomena based on statistics defines how we perceive them.

Monitoring increases transparency and thus supports reflection on an inherent gender bias in organisational processes that are generally perceived to be gender neutral and merit based. While a good database can be the starting point for equality policy, it should be just that – a starting point (Ahmed, 2012). Empirical evidence allows us to identify gendered practices and points to a need for action. If such a reflection leads to an adaption of gendered practices, it can be seen as contributing to a professionalisation of processes.

Last but not least, monitoring provides a validated starting point for a gender equality discourse within the organisation and beyond. Those involved in this gender equality discourse gain gender competence and express their commitment to gender equality. Thus, the reflection based on monitoring results should be seen as part of an organisational learning process that strengthens an organisation’s innovation potential and prepares it to meet future challenges.

Acknowledgment

The authors would like to thank Anke Lipinsky for her wonderful and constructive comments on a draft version of this chapter.

References

- Ahmed, Sara (2012). *On being included: Racism and diversity in institutional life*. Durham: Duke University Press.
- Beck, Tony (1999). *Using gender-sensitive indicators. A reference manual for governments and other stakeholders*. London: Commonwealth Secretariat.
- Butler, Judith (1990). *Gender trouble: Feminism and the subversion of identity*. New York, NY/London: Routledge.
- Childs, Sarah, & Krook, Mona L. (2008). Critical mass theory and women's political representation. *Political Studies*, 56(3), 725–726. doi:10.1111/j.1467-9248.2007.00712.x
- Criado-Perez, Caroline (2019). *Invisible women: Data bias in a world designed for men*. New York, NY: Abrams Press.
- Degele, Nina (2008). *Gender/queer studies: Eine Einführung [Gender/Queer studies: An introduction.]* Stuttgart: UTB.
- D'Ignazio, Catherine, & Klein, Lauren F. (2020). *Data feminism*. Cambridge: The MIT Press.
- Doran, George T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*, 70, 35–36. Retrieved from <https://community.mis.temple.edu/mis0855002fall2015/files/2015/10/S.M.A.R.T-Way-Management-Review.pdf>. Accessed on September 20, 2021.
- Döring, Nicola (2013). Zur Operationalisierung von Geschlecht im Fragebogen: Probleme und Lösungsansätze aus Sicht von Mess-, Umfrage-, Gender- und Queer-Theorie [On the operationalisation of gender in questionnaires: Problems and Approaches from the perspective of measurement theory, survey theory, gender theory, and queer theory.] *Gender. Zeitschrift für Geschlecht, Kultur und Gesellschaft*, 5(2), 94–113. Retrieved from <http://www.nicola-doering.de/wp-content/uploads/2014/08/D%C3%B6ring-2013-Zur-Operationalisierung-von-Geschlecht-im-Fragebogen.pdf>. Accessed on September 20, 2021.
- EC (European Commission). (2018). *Monitoring the evolution of benefits of responsible research and innovation. The evolution of responsible research and innovation: The indicators report*. Brussels: European Commission. Retrieved from <https://op.europa.eu/en/publication-detail/-/publication/2c5a0fb6-c070-11e8-9893-01aa75ed71a1/language-en/format-PDF/source-170166807>. Accessed on September 20, 2021.
- EC (European Commission). (2019a). *European research area progress report 2018*. Brussels: European Commission. Retrieved from https://ec.europa.eu/info/publications/era-progress-report-2018_en. Accessed on September 20, 2021.
- EC (European Commission). (2019b). *ERA monitoring handbook 2018*. Brussels: European Commission. Retrieved from https://ec.europa.eu/info/sites/info/files/research_and_innovation/era/era_progress_report_2018-handbook.pdf. Accessed on September 20, 2021.
- EC (European Commission) (2019c). *ERA progress report 2018. Data gathering and information for the 2018 ERA monitoring – Technical report*. Brussels: European Commission. Retrieved from https://ec.europa.eu/info/sites/info/files/research_and_innovation/era/era_progress_report_2018-technical.pdf. Accessed on September 20, 2021.
- EC (European Commission). (2019d). *She figures 2018*. Brussels: European Commission. Retrieved from <https://www.etag.ee/wp-content/uploads/2019/03/She-Figures-2018-1.pdf>. Accessed on September 20, 2021.
- EC (European Commission). (2020). *A union of equality: Gender equality strategy 2020–2025*. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM (2020) 152 final. Brussels: European Commission. Retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0152&from=EN>. Accessed on September 20, 2021.

- EC (European Commission). (2021). *Horizon Europe: Strategic plan 2021 – 2024*. Brussels: European Commission. Retrieved from <https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1>. Accessed on September 20, 2021.
- Engeli, Isabelle, & Mazur, Amy (2018). Taking implementation seriously in assessing success: The politics of gender equality policy. *European Journal of Politics and Gender*, 1(1–2), 111–129. doi:10.1332/251510818X15282097548558
- Flick, Uwe (2018). *Doing triangulation and mixed methods. The Sage Qualitative Research Kit 8*. London: SAGE Publications Ltd.
- Funnell, Sue C., & Rogers, Patricia J. (2011). *Purposeful program theory. Effective use of theories of change and logic models*. San Francisco, CA: Jossey-Bass/Wiley.
- Guldvik, Ingrid (2008). Gender quota discourses in Norwegian politics. In Eva Magnusson, Malin Rönnblom, & Harriet Silius (Eds.), *Critical studies of gender equalities. Nordic dislocations, dilemmas and contradictions* (pp. 94–111). Stockholm: Makadam Förlag.
- Hallensleben, Tobias, Wörten, Matthias, & Moldaschl, Manfred (2015). Institutional and personal reflexivity in processes of organisational learning. *International Journal of Work Innovation*, 1(2), 185–207. doi:10.1504/IJWI.2015.071192
- Hedman, Brigitta, Perucci, Francesca, & Sundström, Pehr (1996). *Engendering statistics: A tool for change*. Stockholm: Statistics Sweden. Retrieved from <https://www.scb.se/contentassets/886d78607f724c3aaf0d0a72188ff91c/engendering-statistics-a-tool-for-change.pdf>. Accessed on September 20, 2021.
- Heintz, Bettina (2012). Welterzeugung durch Zahlen: Modelle politischer Differenzierung in internationalen Statistiken, 1948–2010 [Worldmaking by numbers: Models of political differentiation in international statistics, 1948–2010.] *Soziale Systeme*, 18(1–2), 7–39. doi:10.1515/sosys-2012-1-204
- ILO (International Labour Organization). (2020). Integrating gender equality in monitoring and evaluation of projects. Retrieved from http://www.ilo.org/wcmsp5/groups/public/@ed_mas/@eval/documents/publication/wcms_165986.pdf. Accessed on October 17, 2021.
- Kanter, Rosabeth M. (1977). Men and women of the corporation. *British Journal of Sociology*, 31, 135. doi:10.1093/sf/57.1.336
- Lipinsky, Anke, & Wroblewski, Angela (2021). Re-visiting gender equality policy and the role of university top management. In P. O'Connor & K. White (Eds.), *Gender, power and higher education in a globalised world: Palgrave studies in gender and education* (pp. 163–186). London: Palgrave Macmillan.
- Markiewicz, Anne, & Patrick, Ian (2016). *Developing monitoring and evaluation frameworks*. Thousand Oaks, CA: SAGE.
- Martin, Patricia Y. (2006). Practising gender at work: Further thoughts on reflexivity. *Gender, Work and Organization*, 13(3), 254–276. doi:10.1111/j.1468-0432.2006.00307.x
- May, Judith V., & Wildavsky, Aaron B. (Eds.). (1978). *The policy cycle*. Beverly Hills, CA/London: Sage Publications.
- Meier, Petra (2008). A gender gap not closed by quotas. *International Feminist Journal of Politics*, 10(3), 329–347. doi:10.1080/14616740802185650
- Mertens, Donna M. (2017). *Mixed methods design in evaluation*. Thousand Oaks, CA: SAGE.
- Moldaschl, Manfred (2007). *Institutional reflexivity. An institutional approach to measure innovativeness of firms*. Chemnitz University of Technology: Papers and Preprints of the Department of Innovation Research and Sustainable Resource Management. Retrieved from <https://www.econstor.eu/bitstream/10419/55387/1/684991462.pdf>. Accessed on September 20, 2021.
- Rossi, Peter H., Freeman, Howard E., & Lipsey, Mark W. (1999). *Evaluation: A systematic approach*. Thousand Oaks, CA: SAGE.
- Sacchet, Teresa (2008). Beyond numbers. *International Feminist Journal of Politics*, 10(3), 369–86. doi:10.1080/14616740802185700

- Stadler, Bettina, & Wroblewski, Angela (2021). Wissen in Zahlen: Potenziale von Gender-Monitoring im gleichstellungspolitischen Prozess am Beispiel österreichischer Universitäten [Knowledge in numbers: The potential of gender monitoring in the equality policy process using the example of Austrian universities.] *Gender – Zeitschrift für Geschlecht, Kultur und Gesellschaft*, 13(2), 142–158. doi:10.3224/gender.v13i2.10
- Stefanick, Marcia L., & Schiebinger, Londa (2020). Analysing how sex and gender interact. *The Lancet*, 396(10262), 1552–1554. doi:10.1016/S0140-6736(20)32346-1
- Storvik, Aagot, & Teigen, Mari (2010). *Women on board: The Norwegian experience*. Berlin: International Policy Analysis: Friedrich Ebert Stiftung. Retrieved from <https://library.fes.de/pdf-files/id/ipa/07309.pdf>. Accessed on September 20, 2021.
- Törnqvist, Maria (2008). From threat to promise. The changing position of gender quota in the Swedish debate on women's political representation. In E. Magnusson, M. Rönnblom, & H. Silius (Eds.), *Critical studies of gender equalities. Nordic dislocations, dilemmas and contradictions* (pp. 75–93). Stockholm: Makadam Förlag.
- United Nations Economic Commission for Europe (UNECE), & World Bank Institute (2010). *Developing gender statistics: A practical tool*. Geneva: United Nations. Retrieved from https://unece.org/DAM/stats/publications/Developing_Gender_Statistics.pdf. Accessed on September 20, 2021.
- van der Haar, Marleen, & Verloo, Mieke (2013). Unpacking the Russian doll: Gendered and intersectionalized categories in European gender equality policies. *Politics, Groups, and Identities*, 1(3), 417–432. doi:10.1080/21565503.2013.816246
- Verloo, Mieke (2006). Multiple inequalities, intersectionality and the European Union. *European Journal of Women's Studies*, 13(3), 211–228. doi:10.1177/1350506806065753
- Voorspoels, Jolien, & Bleijenbergh, Inge (2019). Implementing gender quotas in academia: A practice lens. *Equality, Diversity and Inclusion*, 38(4), 447–461. doi:10.1108/EDI-12-2017-0281.
- Walby, Sylvia, Armstrong, Jo, & Strid, Sofia (2012). Intersectionality: Multiple inequalities in social theory. *Sociology*, 46(2), 224–240. doi:10.1177/0038038511416164
- Weiss, Carol H. (1998). *Evaluation. Methods for studying programs and policies*. Upper Saddle River, NJ: Prentice Hall.
- West, Candace, & Fenstermaker, Sarah (1995). Doing difference. *Gender & Society*, 9(1), 8–37. Retrieved from <http://www.csun.edu/~snk1966/Doing%20Difference.pdf>. Accessed on September 20, 2021.
- West, Candace, & Zimmermann, Don H. (1987). Doing gender. *Gender & Society*, 1(2), 125–151. doi:10.1177/0891243287001002002
- W.K. Kellogg Foundation. (2004). *Logic model development guide: Using logic models to bring together planning, evaluation, and action*. Battle Creek, MI: W.K. Kellogg Foundation. Retrieved from <https://www.aacu.org/sites/default/files/LogicModel.pdf>. Accessed on September 20, 2021.
- Wroblewski, Angela (2015). Individual and institutional reflexivity – A mutual basis for reducing gender bias in unquestioned practices. *International Journal of Work Innovation*, 1(2), 208–225. doi:10.1504/IJWI.2015.071190
- Wroblewski, Angela (2021). Monitoring of ERA priority 4 implementation – Update and final assessment. *GENDERACTION D3.3*. Retrieved from https://genderaction.eu/wp-content/uploads/2021/09/GENDERACTION_WP3_final_report.pdf. Accessed on September 20, 2021.
- Wroblewski, Angela, & Eckstein, Kristin (2018). Gender equality monitoring tool and guidelines for self-assessment. *TARGET D4.1*. Retrieved from http://www.gendertarget.eu/wp-content/uploads/2018/12/741672_TARGET_Monitoring_Tool_D4.pdf. Accessed on September 20, 2021.
- Wroblewski, Angela, Kelle, Udo, & Reith, Florian (Eds.). (2017). *Gleichstellung messbar machen: Grundlagen und Anwendungen von Gender- und Gleichstellungsindikatoren [Making gender equality measurable: Basics and applications of gender and equality indicators.]* Wiesbaden: Springer VS.