INTEGER guidelines for the self-assessment of (the implementation of) transformational gender action plans set up in higher education and research institutions

Authors:

Dr Anke Lipinsky & Maria Schäfer

GESIS-Leibniz Institute for the Social Sciences, Cologne, Germany

Contents

1 Introduction

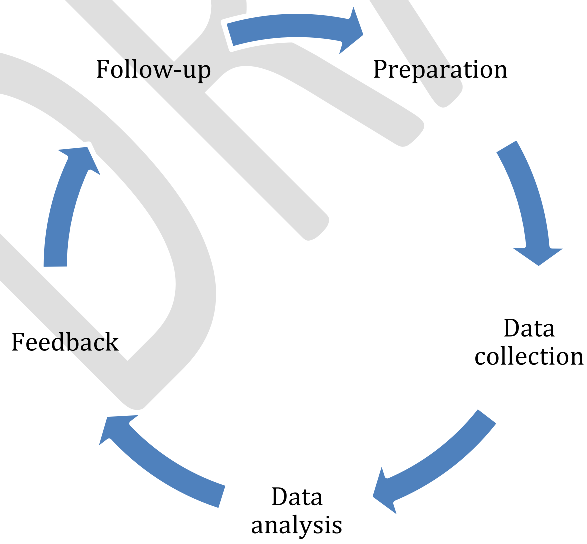
These guidelines aim to supply higher education and research institutions with tools and guidance for the assessment of their Transformational Gender Action Plans. They suggest using evaluation methodology for quality assurance of gender action plans, to support legitimacy and in-house dialogue, and to measure institutional performance of the implementation of plans to foster gender equality.

The target group of the guidelines comprises any actors in charge of or interested in conducting an assessment of (parts of a) Transformational Gender Action Plan. These may include coordinators of gender action plans, gender equality bodies and quality management entities, amongst others.

The guidelines were developed by the evaluation team at GESIS-Leibniz Institute for the Social Sciences in the framework of the project INTEGER ('INstitutional Transformation for Effecting Gender Equality in Research'), funded by the 7th Framework Programme of the European Commission.

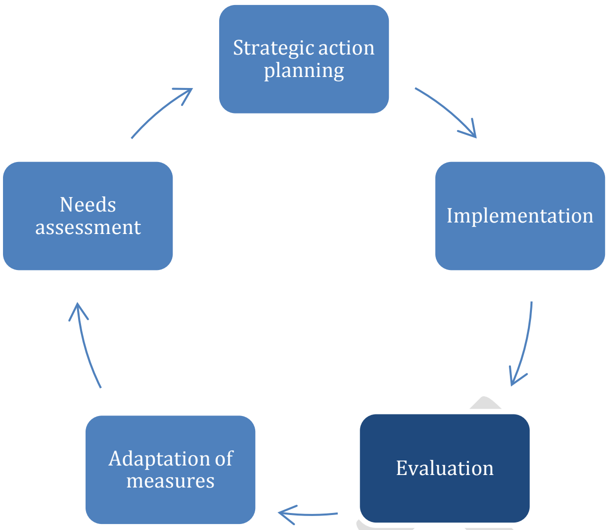
The evaluation cycle

The structure of the guidelines follows the evaluation cycle as suggested by the authors of the guidelines which is relevant for both self-assessments and external evaluations. The first step of the cycle (see graph 1) is the preparation of the self-assessment activities that is described in detail in chapter 2. Subsequently, data is to be collected, followed by the analysis of these data. The authors identified three perspectives of analysis and two phases of collecting and analysing data, and the authors consider these steps a more useful structure for the presentation of this part of the guidelines than the distinction between data collection and data analysis. Both steps of the evaluation cycle are thus addressed in chapter 3. Having analysed the data, the next step is to provide feedback on the self-assessment results. Based on this feedback, it is to be discussed how to implement recommendations resulting from the self-assessment. This follow-up is crucial in order for the evaluation to have any impact on further strategy development, structures and practices. The feedback and the follow-up steps are presented in chapter 4.



Graph 1: The evaluation cycle (source: own graph)

The whole evaluation itself can be considered to form part of a bigger cycle – the programme or project management cycle, depicted as follows:



Graph 2: The programme management cycle (source: own graph)

The self-assessment results are supposed to feed into the programme or project management cycle outlined above by informing a modification of the object and/or its implementation. In the case of a gender action plan, self-assessments can lead to an adaptation of existent measures and/or the development of complimentary measures, shifts in responsibilities or changes to the implementation process. Usually, the self-assessment feedback includes concrete recommendations on which modifications to make in order to optimize the object or the process.

In order for the self-assessment to have any impacts on developments along the programme management cycle, a clear commitment of the management to the self-assessment is essential. Such an explicit commitment should be established before the begin of the self-assessment activities.

Definition of evaluation, including self-assessment

There is a large variety of definitions of evaluation. For the purpose of these guidelines, evaluation is defined as a systematic assessment of an object (in this case: a Transformational Gender Action Plan) in view of defined goals. Simply put, a gender action plan can be considered successful if it reaches its objectives; the same holds for individual measures within the action plan.

This assessment may take place in the course of the implementation of the gender action plan and/or at the end of the validity period of the entire plan or parts of it. This guide suggests to combine formative and summative evaluation elements. Formative evaluation activities take a closer look at implementation processes and aim at contributing to an optimization of the plan and its implementation while it is being implemented by reflecting on what works and what does not work (yet) and why (not). In contrast, summative evaluations focus on the impacts of the gender action plan on the target group(s) and are usually carried out at the end or at a later stage of the implementation process. (For further information on types of evaluation please consult the section on 'Define the type of the evaluation' [link] below.)

Evaluations may be carried out by persons or entities involved in the design and/or implementation of the gender action plan, or by experts that are external to the specific setting (see below: 'Define the role of the evaluator within the institutional set-up and its potential implications for the self-assessment'). This guide specifically targets persons managing the gender action plan who are themselves about to conduct an evaluation of the latter, i.e. a self-assessment.

Limits of and to evaluations, including self-assessments

Any evaluation faces limits arousing both from the evaluation concept and from external restrictions. It is helpful to be aware of and to communicate these limits in order to create adequate expectations regarding the role of the evaluation in e.g. strategy development.

Firstly, at least as regards the type of evaluation suggested in these guidelines the scope of the evaluation is restricted to the scope and objective(s) of the gender action plan to be evaluated. Thus, the evaluation can and should take context factors into account but only to the extent that they are of expected relevance to the action plan and its implementation. Furthermore, evaluators may receive specific task descriptions by the entity or person that commissions the evaluation that may also define the focus and the objectives of the evaluation. Furthermore, evaluators may experience self-inflicted restrictions due to e.g. their position in the institutional hierarchy.

It is to be stressed that, although evaluations provide the ground for an adaptation of strategies and measures, evaluators are not responsible of the implementation of recommendations; it is decision-makers in the institutions that are in charge of this task.

Experience from the INTEGER project showed that the availability of sex-disaggregated data tends to be limited which poses restrictions to evaluation activities. In addition, the fact that internal resources, in particular money and time to be dedicated to be dedicated to the evaluation, are likely to be limited is to be taken into account at the planning stage as well.

Requirements regarding self-assessment competence, methodological skills

These guidelines explicitly target (also) "beginners" in the field of self-assessment. The guidelines are tailored for practical use and the practical examples aim to provide orientation in particular to users with limited or no previous experiences. "Beginners" (and others) are recommended to deepen their methodological knowledge following the references provided in the different sections of the guidelines.

Minimum requirements as regards methodological competencies and skills include the compilation and analysis of descriptive statistics, document analysis and qualitative research, especially the analysis of interviews. Basic skills can be obtained through (short) courses provided by a number of training centres and online platforms. It can be worthwhile to investigate on the respective offers available in the evaluator's own institution.

Furthermore, the quality of the self-assessment highly benefits from a certain gender sensitivity, a familiarity with concepts of gender as well as experience with analyses of gender inequalities.

2 The self-assessment process I: Preparation

This section aims to provide step-by-step guidance for preparing an self-assessment of a Transformational Gender Action Plan.

1. Define the type of the self-assessment

As other types of evaluations, self-assessments can be distinguished by a variety of characteristics, for example according to the subject or the goal(s) of the self-assessment, as described below. A wide-spread approach is to make a distinction between formative and summative evaluations.

Formative evaluations aim to support the optimization of the subject of the evaluation (e.g. a Transformational Gender Action Plan) and its implementation throughout the process. That is why the evaluation feedback tends to target mainly the actors who are in charge of the implementation. Formative evaluations may include a needs assessment, an analysis of the logics of the initiative and an assessment of its implementation process, amongst others. Qualitative data tend to provide particularly valuable information for this type of evaluation.[1]

Summative evaluations, in contrast, examine to what extent the subject of the evaluation has achieved its goals. They assess the outcomes and impacts of the initiative as well as unintended consequences. Summative evaluations are usually conducted during and after the end of the implementation process (of the gender action plan, for example), and are likely to resort to quantitative and qualitative data. Often, the evaluation results are provided to actors who are responsible of taking major decision on the evaluated subject, e.g. on the question which resources will be allocated to its implementation in the future.[2]

Yet, self-assessments as other types of evaluations may combine elements of both formative and summative evaluations, and it has proven fruitful to use a 'mixed methods' approach (see the section on data collection ([link]). In line with this approach, this guide is tailored to assist actors in higher education and research institutions in evaluating Transformational Gender Action Plans or selected elements of these plans by integrating elements of formative and summative evaluations.

Example from the INTEGER project

As pointed out in the introduction to this guide [link], the evaluation concept developed by GESIS-Leibniz Institute for the Social Sciences integrates elements of formative and summative evaluation. As INTEGER project partners, the evaluators have accompanied and supported the design and implementation of the TGAPs in the partner institutions from the very beginning. They contributed to the design of the baseline data collection conducted by the project partners, and collected qualitative and quantitative data both mid-way through the project and at the end of the project. Whereas the formative evaluation is considered to have provided helpful support to the implementing project partners throughout the project, and through the central evaluation report in particular, the summative evaluation only succeeded in providing a limited amount of meaningful data. The main reason for the less prominent character of the summative elements is probably the fact that the project duration was too short for the measures to produce substantial impacts, especially in quantitative terms.

References

The National Science Foundation/ Joy Frechtling (2002): The 2002 User Friendly Handbook for Project Evaluation, available at: http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf, pp. 6ff.

2. Define the subject of the self-assessment

[1] Cf. Alkin, Marvin C. (2011): Evaluation essentials from A to Z. New York: Guilford Press, pp. 11f.

[2] Cf. Ibid., p. 12.

As a basis for further steps, and in order to allow for a realistic attribution of resources (in particular: working time of dedicated staff) it is crucial to define the subject of the self-assessment. This could be, amongst others, (the implementation of) an individual measure, a programme phase, an entire programme, a gender action plan, or the overall institutional implementation of the gender equality mandate. It may encompass measures, plans or the overall set-up at the central, i.e. the organizational, level, and/or focus on (a) subordinate level(s), for example a gender equality programme in a certain faculty.

As it mostly draws on the evidence gathered within the INTEGER project, this guide focuses on the implementation of Transformational Gender Action Plans within higher education and research institutions, and can also be used for the assessment of certain elements of such an action plan (e.g. a single measure).

Example from the INTEGER project

The subject of the evaluation carried out within the INTEGER project was the implementation of Transformational Gender Action Plans (TGAPs) within the three partner institutions at different levels, namely at the level of the organization and within the INTEGER pilot institutes, faculties and laboratories, respectively. These TGAPs aimed at achieving transformational change towards greater gender equality through targeted measures in four identified areas of intervention, namely: engagement of decision-makers; organizational structure; career progression of female scientists; and work-life balance.

3. Define the role of the evaluator within the institutional set-up and its potential implications for the self-assessment

When planning a self-assessment it is, of course, to be decided who will be in charge of conducting the self-assessment. Depending on the institutional context (e.g. the competences and resources of different actors), the person or entity tasked with the self-assessment could be the person responsible of the implementation of the gender equality plan to be evaluated, (other) gender equality actors within the institution, an in-house quality management body or other actors within the institution; alternatively, external (evaluation) experts can be hired for this task.

Involving external evaluators brings the benefits of their expertise, skills and their (relatively) independent perspective, facilitating an overview of relevant features and reducing restrictions caused by intra-institutional power relations. Possibly, external experts coming from prestigious institutions may also contribute to an improved response rate, more cooperative respondents' behaviour and/or an increased interest of the target group in the evaluation results. Yet, external evaluators tend to know less about the local, institutional and, possibly, national context than actors within the institution and thus almost exclusively rely on data collected in the framework of the evaluation.

A third option may be to conduct the evaluation in-house and to resort to evaluation experts at a certain point in the process, for example when designing the evaluation concept or when conducting interviews.

Either way the quality of the evaluation benefits from a self-reflection by the evaluators on their own position and role(s) within the institutional set-up and potential implications for the assessment. This may influence the definition of (implicit) goals of the evaluation, increase awareness of limits to the evaluation and help to understand and reduce different forms of bias.

Please find below an overview of the main advantages of externally conducted evaluations, on the one hand, and self-assessments, on the other hand:

Advantages of External Evaluations	Advantages of Self-Assessments
+ Evaluation expertise & skills	+ High familiarity with internal processes, etc.

+ Rather independent view	+ Lower costs
	+ In-house capacity-building

Table 1. Advantages of external evaluations and self-assessments

Example from the INTEGER project

As pointed out in the introduction to this guide, in the INTEGER project the design and implementation of Transformational Gender Action Plans was supported by an external evaluation carried out by GESIS-Leibniz Institute for the Social Sciences. As a project partner, GESIS closely accompanied the change processes in the partner institutions from the beginning. Thus it was GESIS who, for example, provided the project partners with a basic definition of transformational change (see the section on 'Define key terms [link] below).

The evaluators at GESIS conceptualized their role as 'critical friends', supporting their project partners in the design, implementation and quality assurance of their Transformational Gender Action Plans through a critical and constructive analysis from an external perspective.[3] A basic rationale for the choice of this approach is that it provides guidance in striking "an optimal balance between the challenges of evaluation theory and practice and the fact that the utilization of evaluations depends decisively on the decision makers' interest in the results"[4].

References

Balthasar, Andreas (2011): Critical Friend Approach: Policy Evaluation between Methodological Soundness, Practical Relevance, and Transparency of the Evaluation Process, in: German Policy Studies 7 (3), pp. 187–231.

4. Define the goals of the self-assessment

In an evaluation guide targeting the Commission services, the European Commission identifies four main purposes of evaluation that can be summarized as follows: contribution to (1) the design of interventions; (2) efficient allocation of resources; (3) quality improvement of an intervention; and (4) accountability.[5] It depends on the specific context of the self-assessment which goals it is to focus on. In any case, a basic feature of evaluations is that they are supposed to feed into the optimization of the evaluated subject (and its implementation, in the case of a gender action plan) and thus constitute an essential element of the project/programme management cycle, as described above.

It is recommended to formulate concrete goals of the self-assessment and to tailor the self-assessment design to these goals. The goals defined at this stage will then guide the self-assessment and provide the ground for a quality assessment of the self-assessment itself. Communicating these goals e.g. to potential interviewees serves to explain and underline the relevance of the self-assessment.

In addition, self-assessments – as other types of evaluations – often pursue implicit goals, i.e. goals not explicitly stated neither in the self-assessment mandate nor in any further communication on the self-assessment towards e.g. decision-makers. An important goal can be the legitimation of existent gender equality structures and initiatives or an awareness-raising of gender inequalities and the need for (further) initiatives to combat these inequalities.

Example from the INTEGER project

[3] Cf. Balthasar, A., (2011), "Critical Friend Approach: Policy Evaluation between Methodological Soundness, Practical Relevance, and Transparency of the Evaluation Process", *German Policy Studies 7*, 3, pp. 191ff.
[4] Ibid., p. 191.
[5] Cf. European Commission (2004): Evaluating EU activities: A practical guide for the Commission services, available at: http://ec.europa.eu/agriculture/eval/guide/eval_activities_en.pdf, p. 9.

The evaluators identified two main goals of their evaluation of the TGAPs designed and implemented by INTEGER partner institutions: Firstly, the external evaluation should provide the local project coordinators with an independent view on the implementation process as support for programme steering and quality assurance with respect to chosen objectives, including sustainability of advancement in gender equality. Secondly, the evaluation aimed to explore outputs, outcomes and impacts of each TGAP at the organisational and subordinate levels for the purpose of proving cause-effect relations by making effects of its activities tangible. In addition, the evaluation was also supposed to feed into these guidelines and toolkit aimed to provide guidance on evaluating Transformational Gender Action Plans to other higher education and research institutions.

References

United Nations Development Programme (UNDP) (2009): Handbook on Planning, Monitoring and Evaluating for Results, available at: http://web.undp.org/evaluation/evaluations/handbook/english/documents/pme-handbook.pdf, pp. 127ff.

5. Define key terms

Ideally, the evaluated plan, measure, etc. provides clear definitions of relevant key terms. In particular, relevant documents should be scanned for definitions of what will be measured which is likely to correspond to the objective(s) of the gender action plan, the measure, etc. Still, it is often the case that not all key times are clearly defined. In these cases, definitions should be set up in accordance with the goals of the plan and its measures and, whenever necessary, in consultation with relevant actors and/or experts.

Example from the INTEGER project

As the overall goal of all TGAPs set up within the INTEGER project is to support transformational change towards greater gender equality within the partner institution, it was necessary to agree on a common understanding of transformational change between implementing partners and the evaluator. Thus, the evaluation team supplied its project partners with an operational definition of transformational change which has been adapted by each partner. The basic definition reads as follows:

"Transformational Change is a strategic mean which is steered by institutions that employ research staff. Through operating transformational change, research institutions demonstrate significant gender awareness and competency to use gender as a resource to create new knowledge and stimulate innovation by modernizing their organizational culture. The ultimate objective of the change process is to work towards a better gender relation and equal representation of both sexes in all staff categories of the institution. Operating transformational change effectively demands awareness of the statistical base, periodical examination of institutional processes (such as recruitment, promotion, retention), the willingness at the top of the institution to open up discussion and to sustain the process of self-study and change and support the achievement of organizational goals within a supportive climate."[6]

References

References to tools

The following glossaries are not at all comprehensive but may prove useful in defining key terms related to the gender equality plans to be evaluated.

[6] Cf. Lipinsky A., Schäfer M. (2014), GESIS Evaluation Concept for Transformational Gender Action Plans, INTEGER Deliverable 7.13, available at: http://www.gesis.org/cews/das-cews/cews-projekte/projekte/integer/.

- The OSCE Glossary on Gender-Related Terms of 2006, compiled by the OSCE Gender Section, is predominantly based on OSCE sources and provides definition and examples of 15 basic gender equality-related terms and concepts. The glossary is available at: http://www.osce.org/gender/26397.
- The European Commission's Glossary of Gender and Development terms that forms part of the "Toolkit on mainstreaming gender equality in EC development cooperation" (2004) focuses on gender and development cooperation but provides clear definitions of a number of gender-related concepts that can be useful for gender equality initiatives in universities and research institutions as well. It is available at: http://ec.europa.eu/europeaid/sites/devco/files/toolkit-mainstreaming-gender-section-3_en.pdf.
- The Glossary of Gender-related Terms and Concepts developed by the United Nations International Research and Training Institute for the Advancement of Women (UN-INSTRAW) contains around 20 brief definitions of basic concepts relevant to gender equality-related work, including definitions of different degrees of integration of a gender perspective in projects. The glossary can be accessed via the following link: http://unamid.unmissions.org/Default.aspx?tabid=11012&language=en-US.
- The European Institute for Gender Equality (EIGE) is envisaging to establish a Europe-wide gender equality thesaurus and glossary in consultation with stakeholders which can be expected to become a useful tool in the future.
- For the definition of key terms related to the self-assessment itself the Glossary of Key Terms in Evaluation and Results Based Management (trilingual version) is recommended. It is available online via the following link: http://www.oecd.org/derec/dacnetwork/35336188.pdf.

6. Define the time-frame

Two time-frames are to be defined before conducting the self-assessment: the time-frame of the self-assessment activities and the time-frame related to the subject of the self-assessment. The first time-frame refers to the planned dates of the beginning and the completion of the self-assessment activities, and the work phases during this period, namely data collection, data analysis and presentation of results. As the second relevant time-frame it is to be defined which period of time will be assessed in the self-assessment, such as the implementation period of a (phase of a) gender action plan since its entry into force.

In both regards it is recommended to adapt the time-frame to certain internal and/or external cycles or caesura, where possible, in order to seize a certain momentum. Such cycles or caesura may be political (e.g. related to the university elections, or a relevant national law) or administrative (e.g. a redefinition of staff categories).

Example from the INTEGER project

The INTEGER project did not leave much scope for defining time-frames of the evaluation and thus did not allow for taking into account institutional or local caesura. The period of time to be assessed corresponded to the time-frame of the evaluation itself, starting with the design of the Transformational Gender Action Plans.

However, the evaluators could define the points in time of the different phases of the evaluation. Thus, they opted for scheduling the central evaluation round for around 1 ½ years after the official beginning of the implementation of the Transformational Gender Action Plans, and the final evaluation round for one year later. The experiences from the INTEGER project show that a time-span of four years between the baseline data collection and the (final) evaluation is rather short, at least if impacts are to be measured.

7. Identify corpus, sources of information and material to be collected and analysed

In order to provide for a sound self-assessment, the analysis carried out needs to be based on a clearly defined corpus, i.e. the material to be analysed. The definition of the corpus should be guided by the objectives of the self-assessment and the methods of analyses (see section on 'Implementation' [link] below), and need to take into account practical considerations like the time-frame and the resources available for the self-assessment. It is recommended to consider both quantitative and qualitative data.

Sources of information to be identified can include, amongst others, administrative (HR) data, further compilations of sex-disaggregated (staff) data at national, organizational and/or subordinate levels, official institutional documents, policy and legal documents at regional, national and European levels, the gender action plan etc. itself, interviews, group discussions and surveys [link to tool: 'List of potential sources of information']. In order to select methods of data collection, it is first to be assessed which data is already available within the institution and beyond. Data gaps are to be identified and it is to be decided which data needs to and can be harmonised. Subsequently, it can be determined which data is to be produced in order to fill existing data gaps.

Example from the INTEGER project

In a detailed evaluation concept shared with the project partners the evaluators within the INTEGER project defined the corpus of material, comprising: each institution's Transformational Gender Action Plan and supplementary documents; a completed template for collecting HR statistics; the INTEGER baseline data report that features, amongst others, key results of the staff surveys conducted at the implementing partner institutions; each project partner's self-report; individual and group interviews as well as group discussions with staff, decision-makers and gender equality actors; site-visits to the pilot entities within the partner institutions; and policy and legal documents at national, organisational and local level collected through additional desk research [each with links to the respective sections].

8. Select data collection methods

The definition of the corpus and the assessment of available data (see above [link]) give information about which data is available and needs to be collected and which data is to be produced in the course of the self-assessment.

It is recommended to use a mixed methods approach, i.e. to use both qualitative and quantitative methods. That is because mixed methods tend to produce more comprehensive insights than one of the methods alone. The use of several methods for the assessment of the same object (a technique called triangulation) also serves to check the validity of the results, and thus supports solid, sound and credible research. It can also be argued that in any way both types of data are closely related (as Trochim (2000) points out: "[a]ll quantitative data is based upon qualitative judgments; and all qualitative data can be described and manipulated numerically"[7]).

Selected data collection methods that are considered relevant for evaluating Transformational Gender Action Plans are briefly outlined in the section 'Overview of relevant data collection/production methods' (link) below. For further information and practical guidance please consult to the entire section on the implementation [link] of the self-assessment and the recommended literature.

Example from the INTEGER project

In the framework of the INTEGER project, the evaluators opted for a mixed-methods approach, combining all data collection methods mentioned above. Broadly speaking, desk research was particular important at the beginning of the project, whereas the central evaluation round resorted to interviews, group discussions, site visits, and self-reports provided by the project partners. As regards the final evaluation web surveys were used. Quantitative staff data was collected at regular intervals throughout the project, facilitating a sound data monitoring.

In the selection of data collection methods and in their application the evaluation team benefitted from the fact that the team is composed of researchers with different methodological expertise and skills which facilitated fruitful exchanges.

[7] Cf. Trochim, William M. (2000): The Research Methods Knowledge Base, 2nd Edition, online version available at http://www.socialresearchmethods.net/kb/.

3 The self-assessment process II: Implementation

The preparatory work described in the previous chapter paves the ground for the actual implementation of the self-assessment.

This guide suggests the self-assessment to integrate three perspectives on the gender action plan to be evaluated, namely to examine (1) the framework conditions for creating and implementing the TGAP; (2) the implementation process of the TGAP and (3) the impact created by the TGAP and its measures on site.[8]

Following this approach, two phases of data collection and analysis can be distinguished: The first phase covers an assessment of the framework, implementation process, input and output of the gender action plan (i.e. the first and second perspective), and the second phase comprises an analysis of its outcome and impact (i.e. the third perspective).

Within each phase, relevant data is to be both collected or produced and analysed. The decision on which data is to be collected and which data needs to be produced is rooted in the mapping of available data and the identification of data gaps carried out in the preparatory phase.

Overview of relevant data collection/production methods

Relevant qualitative data collection methods include document analyses, interviews, group discussions, surveys and guided field/site visits.

An analysis of key documents at European, national, institutional and/or local level [link: 'Framework analysis'] is crucial for adequately taking the contextual conditions of the design and implementation process into account.

Interviews with actors in charge of the implementation of the gender action plan and possibly further gender equality actors are particularly instructive for an analysis of implementation processes and outcomes of gender equality measures. Furthermore, expert interviews can guide the document analysis and/or complement the insights gained through that method in order to assess the framework conditions [link: 'Framework analysis']. Interviews with the target groups of measures – e.g. staff or decision-makers –, in turn, can shed light on impacts of these measures [link: 'Assessment of impacts']. For that purpose, group interviews [link: …] are particularly recommended.

Group discussions [link: 'Assessment of impacts' > 'Group discussions'] and surveys can be another way to assess target group reach-out and impacts on target groups. Whereas group interviews and discussions may allow for deeper insights into the impacts of certain measures on a small group of people, a survey provides the opportunity to measure reach-out and impacts of measures on a larger group of staff, and can possibly even be deemed representative.

Field/site visits [link: 'Framework analysis' > 'Site visit(s)'] in which the evaluators are guided by staff are a useful method of obtaining insights into specific day-to-day framework conditions like, amongst others, the work environment and the work climate.

Quantitative data can contribute to an overall picture of existing gender inequalities and imbalances in the institution [link 'Framework analysis' > 'Statistical data'] and can provide information on impacts of gender equality measures [link 'Assessment of impacts' > 'Collection of quantitative data']. Depending on the goals of the gender action plan and the self-assessment, the evaluators are recommended to compile sex-disaggregated HR data on e.g. the employment status and pay as well as further sex-disaggregated staff data on e.g. participation in committees. In order to detect developments over time that may eventually be causally attributed to the gender action plan, these data are to be collected repeatedly. Furthermore, relevant quantitative data may be produced through a survey. Climate surveys facilitate an assessment of the starting

[8] Cf. Lipinsky A., Schäfer M. (2014), GESIS Evaluation Concept for Transformational Gender Action Plans, INTEGER Deliverable 7.13, available at: http://www.gesis.org/cews/das-cews/cews-projekte/projekte/integer/ (direct link: http://www.gesis.org/cews/fileadmin/cews/www/download/D7_13-Detailed-Evaluation-Concept.pdf).

point as regards the work climate, and possibly a comparison between the situations at different points in time. Surveys may also be used to measure different kinds of impacts in quantitative terms.

References

The National Science Foundation/ Joy Frechtling (2002): The 2002 User Friendly Handbook for Project Evaluation, available at: http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf, pp. 43ff.

United Nations Evaluation Group (UNEG) (2013): Handbook for Conducting Evaluations of Normative Work in the UN System, available at: http://www.uneval.org/document/detail/1484, pp. 48ff.

Krueger, Richard A. (2002): Designing and Conducting Focus Group Interviews, available at: http://www.eiu.edu/~ihec/Krueger-FocusGroupInterviews.pdf.

3.1 Phase I: Assessment of the framework, implementation process, input and output

For the first phase, i.e. the assessment of the framework, implementation process, input and output of the Transformational Gender Action Plan, it is recommended to focus mainly on qualitative data. Relevant data can be collected or produced through document analysis, interviews and self-reports. Yet, quantitative data can provide very relevant information regarding the national, institutional and local context as well.

Framework analysis

Purpose of the framework analysis is to understand the contextual conditions and possible constraints at organisational and subordinate levels in relation to (elements of) the Transformational Gender Action Plan. Organisational structures significantly determine the implementation process of gender equality activities.[9]

Background information on higher education legislation and research governance policy, employment policy in public research, gender equality policy applicable to higher education and research institutions, etc. as well as relevant characteristics of the specific institution and entity – i.e. the status quo regarding the participation of women in decision-making, career paths of female and male scientists, financial resources dedicated to gender equality measures, the dominant gender equality discourse, persistence of gender stereotypes, etc. – serve as background for weighting more specific assessments of processes, outcomes and impacts, e.g. the potential of the institution to demonstrate change within a specific area of the gender action plan.

It is recommended to sum up the contextual conditions in a 'Framework report' at the beginning of the self-assessment, and to revisit it at a later stage when assessing the outcomes and impacts. Depending on the role of the evaluator(s) in the institution and in the implementation process – and thus on their knowledge – the evaluator(s) may decide to write the report themselves or to delegate (parts of) this task to another actor in the institution. Such a 'Framework report' may contain the following sections: (1) summary of the national, institutional and local baseline situation; (2) structure and organization of gender equality in the institution; and (3) objectives and overall concept of the gender action plan. The toolkit [link] contains a detailed template for such a 'Framework report'.

Data collection/production

Document analysis

Relevant information on the legal and policy context at national, institutional and local levels can be obtained by means of document analysis [link 'References']. It is recommended to include the following (types of)

[9] Cf. Löther A., Vollmer L., (2014), "Erfolge durch Strukturen? Hochschulische Gleichstellungsarbeit im Wandel". In: Andrea Löther und Lina Vollmer (Eds.): Gleichstellungsarbeit an Hochschulen. Neue Strukturen – neue Kompetenzen. Leverkusen: Barbara Budrich (cews.Beiträge Frauen in Wissenschaft und Forschung), pp. 15–53.

documents in this analysis: Relevant national (gender equality, research) policies and legal provisions, etc. (if existent); institutional gender equality policy, action plans, etc. (if existent); further institutional policies and other relevant institutional documents at different levels within the institution; and documents related to the Transformational Gender Action Plan to be evaluated.

Interviews

Further insights can be gained from interviews with representatives of the management, gender equality actors and research staff. For further information on conducting and analysing interviews and group discussions please see the section on 'Process analysis' [link]. A self-report (see 'Process analysis' [link]) can also be useful for collecting data on framework conditions.

Statistical data

In addition, quantitative data can provide a broader understanding of gender (im)balances in decision-making positions in management and research and career paths of female scientists and other staff (e.g. drop-outs), including for example career breaks. Such data may be available from national statistical offices or the administration of the higher education and research institution, for example, or could be collected. For further guidance on the collection and analysis of quantitative data please consult the section on 'Assessment of impacts' [link].

Field visit(s)

Field visits within the institution can provide insights through observations and are likely to facilitate informal discussions with staff. They are a fruitful method to investigate on the work environment (e.g. the work place facilities), the work climate (e.g. the overall atmosphere, collegial cohesion and social activities), the visibility of gender equality-related initiatives, and the existence and quality of child-care facilities, amongst others.

<u>Example from the INTEGER project</u>

At the beginning of the evaluation activities, the evaluators extensively analysed political and legal documents at national level as well as key policy documents at organisational level in order to gain an understanding of the context in which the Transformational Gender Action Plans were drafted and implemented. Yet, in particular as regards the evaluation of Šiauliai University's Transformational Gender Action Plan, the corpus of documents that could be analysed was restricted due to language barriers.

The evaluators asked the local project coordinators and TGAP managers to provide a self-report on the contextual conditions and the implementation of the TGAPs. For this purpose the evaluators provided a self-report template that served as the basis for designing the 'Framework report' and 'Process report' templates provided in the toolkit. As the evaluators were external to the national and institutional contexts as well as to the implementation processes, the insights gained from the self-report – although the framing of certain information can be expected to have been influenced to some extent by the authors' own interest – were of high value and fed into the preparation of the interviews and the site visit.

Furthermore, the evaluators carried out site/field visits at the pilot faculties/institutes during which they were guided by local staff. The field visits offered informative insights into the work environment, the work climate and further aspects of interest.

<u>References</u>

The National Science Foundation/ Joy Frechtling (2002): The 2002 User Friendly Handbook for Project Evaluation, available at: http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf, pp. 49ff.

Process analysis

In order to assess the operationalization of the gender action plan, it is suggested to apply a process analysis. By looking at key actors involved in the implementation of the action plan as well as the institutional behaviour (actors not directly involved but potentially affected), strengths and weaknesses of the institution managing the change process can be revealed.

Relevant elements of the implementation process to be analysed include: resources (e.g. financial resources and time dedicated to the implementation of the gender action plan and the access to relevant expertise/competence); implementation structures (e.g. types of actors involved, their responsibilities and their level of ownership of the gender action plan, and the role and position of the coordinator of the plan); further collaboration within and beyond the institution; internal communication on gender equality measures; public relations related to the action plan; the work flow regarding the selection of activities and their implementation; drivers as well as resistances and conflicts; and prospects for the sustainability of the processes.

Data collection/production

Promising data collection/production methods for an assessment of implementation processes include interviews, 'Process reports' and field visits.

Process report

In order to establish a comprehensive data base for the process analysis it has proven useful to issue a 'Process report'. Such a report serves to compile relevant data in a structured manner and can constitute a starting point for deeper investigation e.g. during interviews. The report should be written by persons actively involved in the implementation of the Transformational Gender Action Plan – i.e. by the evaluator(s) themselves and/or by other actors –, and the work may also be divided between different actors according to their experience. In any way the evaluators should provide a clear template for the report that could include the following topics: (1) level of implementation of the gender action plan; (2) strengths and weaknesses of implementation structures; (3) strengths and weaknesses of communication processes; (4) internal quality control; (5) cooperation on gender equality within and beyond the institution; and (6) sustainability mechanisms. A template for such a 'Process report' is provided in the toolkit [link].

Interviews

Depending on the specific focus of the interview session with each respondent, it may be more fruitful to conduct an individual interview (probably the most common form) or a group interview. Whereas individual interviews allow for questions specifically tailored to the individual respondent and tend to facilitate openness due to the more confidential situation, group interviews can offer valuable insights into roles of actors as well as communication and work processes, and may even evoke a more relaxed and open atmosphere. Group interviews may be particularly useful if they bring together actors involved in the implementation of gender equality measures who find themselves at a similar hierarchical level, as they can complement each other's narrations and do not feel troubled by the presence of superiors that may increase self-censorship.

The first step in the run-up to an interview is to decide who is going to conduct the interviews. Assigning this task to the evaluators themselves brings the benefit of their deep understanding of the self-assessment purpose, the subject of the self-assessment and the questions to be asked. In some cases, in particular if the person(s) in charge of the self-assessment is/are deeply involved in the organisational structures and hierarchy, it may make sense to have another person – e.g. a professional evaluator – carry out the interviews. In this case, comprehensive briefings by the person(s) in charge of the self-assessment are crucial. Usually, these considerations are made already when deciding whom to task with the whole self-assessment.

Subsequently, the evaluators need to identify which interviewees may be the most suitable ones for which kind of information to be collected. Key actors involved in the design and implementation of the

Transformational Gender Action Plan as well as (other) key decision-makers in the institution should be interviewed in any case. In addition, it may be helpful to interview e.g. administration staff as well.

As regards the identification of interview partners it is to be taken into account that the same interview participant may provide information about various factors relevant to the implementation of the gender action plan, e.g. the framework conditions, the process and the outputs.

In order to increase the participation rate it is important to address the potential participants adequately, explaining the purpose of the self-assessment and pointing out the compliance with data protection standards. It may be helpful to have a look at the sample invitation letter provided in the toolkit [link]. Furthermore, it is important to provide for favourable external conditions, e.g. an adequate room with a calm, pleasant atmosphere. Depending on culturally specific practices, available facilities, possibly their position in the institution's hierarchy and personal preferences, interviewees may prefer to be interviewed in their own office. This may provide the interviewer with interesting insights into the interviewee's workplace and may offer stimuli for further questions.

Furthermore, it is to be verified in advance if an interpreter (spoken or sign language) is needed for the interview, and, if so, it is favourable to involve an interpreter with an understanding of gender inequalities and to brief the interpreter about key features of the Transformational Gender Action Plan and the context.

It is recommended to record the interviews in order to allow for a thorough analysis. This requires the informed consent of each participant; and high standards of data protection must be guaranteed (see the toolkit for a sample consent form [link]). A smooth run of the technical equipment is to be guaranteed. If a participant does not agree to the recording of the interview it is recommended to task an additional person (if possible a member of the self-assessment team) with taking notes. Notes-taking can also be helpful to provide guidance in the analysis of interviews, and to directly retain insights to be used in subsequent interviews.

For conducting the interview a semi-structured form is recommended, as it reduces the risk of a little fruitful interview and still leaves room for adapting to the flow of the conversation. During the preparation of such a semi-structured interview, the information to be gathered is to be defined, and subsequently suitable questions are to be developed. Whereas some information can be easily extracted, other issues may require more indirect questions. A draft structure can guide the interviewer through the interview. It has proven useful to develop questions and potential follow-up questions, including references to practical examples or academic literature, for example. The toolkit contains an example of such a draft structure (link: Template for the preparation of interviews/group discussions).

In addition to a thorough preparation as regards content, it is recommended to acquire basic interview techniques in the run-up to the interviews. Trainings allow the interviewer to already practice certain skills, but if trainings are not an option it can be useful to consult literature on interview techniques (see the section 'References' below [link]).

<u>Example from the INTEGER project</u>

As mentioned above (link: Framework analysis > Example from the INTEGER project), the evaluators asked the local project coordinators and TGAP managers to provide a self-report that featured several sections on the implementation process.

In addition, the evaluators conducted a total of 41 interviews and group discussions with representatives of each institution's senior management, members of governing bodies, representatives of the central and de-central administration, members of the TGAP implementation teams, and research staff and senior academics.

All three evaluators attended all interviews and discussions. In the preparation phase they had assigned specific tasks to each of them for all the interviews and discussions and had these tasks rotate among each other. Thus, the lead interviewer was always supported by a second interviewer, intervening only at certain points, and by a person taking notes and keeping the time. This set-up has proven instrumental in all interviews. Whenever necessary, the evaluators were accompanied by interpreters (English-French-English and English-Lithuanian-English) to guarantee the 'freedom of expression' of each interviewee and to reduce misunderstandings and subsequent misinterpretation of the empirical data collected.

References

Baranowski, T., Stables, G. (2000): "Process Evaluation of the 5-a-Day Projects", in: Health Education and Behavior, 2000, 27 (2), pp. 157–166.

Cohen D., Crabtree B. (2006): "Semi-Structured Interviews", Qualitative Research Guidelines Project, available at: http://www.qualres.org/HomeSemi-3629.html.

Analysis of inputs

These self-assessment guidelines do not focus on input-output-efficiency. Yet, it is informative to analyse the inputs, amongst others because this can provide an understanding of the possibilities and constraints of gender equality measures as well as of the political will to implement the Transformational Gender Action Plan.

Resources may encompass, for example, financial means, (paid and unpaid) time dedicated to the implementation of the gender action plan, and expertise in terms of gender concepts, gender equality work and self-assessment.

Data collection/production

It is recommended to collect data on inputs in the 'Process report' (see above [link]). By means of a table template the person(s) tasked with the 'Process report' can provide information on (amongst others) input and output for each gender equality measure. The toolkit [link] provides a table template as used in the INTEGER project.

Example from the INTEGER project

The evaluators within the INTEGER project collected information on inputs by means of a table to be filled in by the project coordinators at the partner institutions. This table lists all measures featured in the institution's Transformational Gender Action Plan and is to be completed with data on the budget and time attributed to each measure, amongst others. In addition, the evaluators included a question on sources of expertise the TGAP managers resorted to in the self-report to be delivered by the local project coordinators.

Analysis of outputs

Outputs can be defined as the direct effects of the evaluated subject. In the context of Transformational Gender Action Plans, the implemented gender equality measures themselves (for example a completed briefing, a communications campaign or the establishment of a mentoring programme) can be considered to constitute the outputs of the gender action plan. Following the logics of the logic chart model [link], measures are successful if these outputs generate the desired outcomes and impacts.

Data collection/production

As briefly described above [link: Analysis of inputs > Data collection/production], a structured way to collect data on outputs is a 'Process report'. Please consult the toolkit [link] for a respective table template.

Example from the INTEGER project

As suggested above, the evaluators identified the implemented TGAP measures as the outputs of the TGAP. In order to obtain comprehensive information on the status quo of the implementation of each measure, they asked the TGAP managers at the partner institutions to complete a table template in which the partners could tick a box if a measure had already been (partly) implemented, and provide additional information.

3.2 Phase II: Assessment of outcomes and impacts

The recommended methodological approach to be followed in the analysis of the outcomes and impacts of the Transformational Gender Action Plan is the logic chart model[10]. It aims at shedding light on causal relationships between outputs, outcomes and impacts of measures. The use of the logic chart template [link] is highly recommended as it provides clear and comprehensive guidance in this analysis.

As described above, gender equality measures implemented, such as products delivered, constitute the output. Outcome, in turn, refers to specific changes directly resulting from the output, for example, increased awareness of gender inequalities among participants in a certain training. Impact is defined as the wider effects on the target group(s) of the measures that can be causally attributed to the implementation of the Transformational Gender Action Plan, for example the removal of barriers to the career progression of female scientists.

A measure can be considered successful if it reaches its objective. To the extent that this is possible, the analysis should differentiate between outcome and impact at different levels within the institution, especially if the gender action plan targets different levels specifically.

Assessment of outcomes

In order to assess the outcomes of gender equality measures it is necessary to depart from each output [link] and to analyse which changes can be directly attributed to specific outputs.

<u>Data collection/production</u>

Interviews

The most promising method to collect data on outcomes is to conduct interviews with staff who directly benefitted from certain measures, for example the participants in a mentoring programme. Both individual and group interviews can be fruitful in this regard (see the section on 'Process analysis' [link: 'Interviews'] for (dis)advantaged of each of the two methods).

If the expected outcome is e.g. an increased awareness of the existence of gender inequalities or of gender equality initiatives, it is recommended to conduct group interviews or group discussions with randomly selected staff members, or a group of staff members that is considered representative of the institution. Group discussions can be considered particularly useful if the focus is put on attitudes including awareness, stereotypes, etc. For further information on the benefits and challenges related to group discussions please consult the section 'Analysis of impacts' [link: 'Group discussions'].

It can also be informative to ask actors who are closely involved in the implementation of the Transformational Gender Action Plan to provide their assessment of the outcomes. This information may be obtained in interviews or in a 'Process report' (see 'Process report' [link]).

Field visit(s)

[10] Cf. Balthasar A., (2011), "Critical Friend Approach: Policy Evaluation between Methodological Soundness, Practical Relevance, and Transparency of the Evaluation Process", *German Policy Studies* 7, 3, pp. 187–231; Knowlton Wyatt L., Phillips C. C., (2009), *The Logic Model Guidebook. Better Strategies for Great Results*, Los Angeles, California: Sage Publications.

Field visits do not only provide information on the setting in which the Transformational Gender Action Plan is being implemented but can also offer practical insights into the implementation of certain measures, for example of measures aiming at improving intra-department communication or childcare. For further information on field visits please see the section on 'Field visits' [link] in the chapter on framework analysis.

Example from the INTEGER project

In the evaluation carried out in the INTEGER project, the analysis of outcomes was guided by the logic chart model [link] departing from the (partly) implemented measures as outputs.

Data was mostly collected through individual interviews with the project partners, group interviews with the TGAP implementation teams and group discussions with staff. This information was complemented by insights obtained from the 'self-report' provided by the project partners and, to some extent, from insights gained during site visits, e.g. as regards the establishment of child care facilities. Although relying heavily on actors involved in the implementation of measures to be evaluated carries the risk of a significant bias, the data collected from the actors involved in the INTEGER project can be considered quite comprehensive, also because of the fact that these actors are not a homogenous group as regards their perspectives and motivations.

Assessment of impacts

The assessment of impacts indicates to what extent a measure has reached its objective, and thus sheds light on the success of the Transformational Gender Action Plan as a whole. The task of the self-assessment in this regard is to assess effects on the target group(s) (e.g. female researchers) that can be attributed to measures introduced through the gender action plan. For this analysis of causal relations these effects have to be assessed in the lights of outputs and outcomes. Yet, whereas outcomes can usually be observed in the short term, impacts may only manifest themselves in the medium or long term. Thus, the fact that a certain desired impact cannot yet be observed does not always imply that it is not to come.

Data collection/production

Collection of quantitative data (data monitoring)

As staff data is often collected and stored by various entities at central and de-central levels, a time-consuming, yet important task is to compile the available sex-disaggregated data and harmonise it – if possible. The HR department is probably most likely to be able to provide useful staff data, especially on forms of employment and pay. However, it may be necessary to contact each faculty, institute, department, laboratory and/or governing body to fill remaining gaps. Compiling data from different sources also poses the challenge of definitions of staff categories that may vary between institutes, etc.

The objective should be to establish a consistent database with sex-disaggregated data on e.g. research and management positions as well as participation in decision-making bodies (depending on the objectives of the Transformational Gender Action Plan) that allows tracking developments over time. For this purpose it is crucial to collect these data at regular intervals which would benefit future self-assessments as well.

The first step is to assign the responsibilities for data collection, taking its (most probably) time-consuming character into account. It can be useful to delegate the data collection at de-central levels to staff who is familiar with the local setting; yet, the coordination should be placed in the hands of one person or entity who stays on top of all activities and is responsible of the harmonisation.

Subsequently, categories of staff etc. are to be defined, guided by the objective(s) of the self-assessment. This may require the harmonisation of different co-existing definitions. It is recommended to set up a data monitoring template in which to insert the data. The toolkit [link] contains a data monitoring template as used in the INTEGER project. As mentioned before, the actual data collection is then to be carried out at regular intervals, e.g. once a year, and should preferably be continued after the end of the self-assessment.

Survey(s)

Surveys can be considered a very informative method to collect data on impacts of the Transformational Gender Action Plan among a large group of people. Depending on the target group(s) of the gender action plan they may address e.g. research, technical and/or administrative staff in the institution or in a certain section, students and/or members of certain decision-making bodies. The most feasible and accessible type of survey for this purpose is a web survey.

On the down side, surveys tend to be quite time-consuming if thoroughly prepared, and require (access to) respective methodological and technical competence. Furthermore, any data protection issues need to be clarified in advance. Most importantly, surveys need to be based on valid indicators. The development of these indicators is to be rooted in existing knowledge acquired in previous research (e.g. desk research, interviews).

An extensive amount of literature on survey methodology, including questionnaire design provides guidance in constructing a survey (see 'References' [link]).

Interviews

Insights into impacts may as well be obtained through interviews with members of the target group(s) of the measures, for example female researchers or decision-makers in the management. As most probably only a small number of people can be interviewed, this method cannot be expected to produce representative results. Interviews usually serve a rather explorative purpose, and may even feed into the design of a survey at a later stage.

Please consult the section on 'Process analysis' [link] above for further information on individual and group interviews.

Group discussions

In addition, it may prove useful to conduct group discussions among members of the target group(s) of the Transformational Gender Action Plan. Group discussions aim to provide insights into opinions and experiences of the selected group, and are thus a particularly suitable method to investigate on understandings of gender, gender stereotypes, the awareness of gender inequalities, and attitudes towards gender equality measures. Yet, conducting a group discussion is a challenging task requiring particularly thorough preparation. That is because moderating techniques need to be carefully selected in order to facilitate an open discussion without inflicting too much bias on the participants. Indirect questions are key to this. It is recommended to pay particular attention to the design of the opening question that is supposed to stimulate a first exchange of thoughts. It may resort to e.g. recent events, quotations or pictures.

The toolkit [link] contains an example of a plan for a group discussion among academic and research staff.

Field visit(s)

As pointed out above [link: 'Framework analysis' > 'Field visit(s)'], field visits to e.g. institutes and laboratories – possibly those specifically targeted by the Transformational Gender Action Plan if the plan does not address all sections of the institution alike – facilitate insights into the work environment, the work-climate and the visibility of e.g. childcare-facilities, amongst others. Thus, it may also be indicative for an ad-hoc impression of the impact of certain measures targeting exactly these features. These impressions can complement respective information shared by e.g. interviewees.

<u>Example from the INTEGER project</u>

The evaluators in the INTEGER project conducted analysed the impacts of the Transformational Gender Action Plans at two points in time. The first impact analysis was conducted during the central evaluation round when it was guided by the logic chart model, departing from specific measures. Interviews and group discussions with members of the target group of measures provided insights into the (limited) impacts that had manifested themselves at this stage. This impact analysis was of rather exploratory nature, as no impact indicators could be defined at this point.

The second impact analysis was carried out ca. one year later in the framework of the final evaluation. This analysis examined to what extent the overall objective of the Transformational Gender Action Plans and the project as a whole have been reached. For this purpose, the analysis departed from the concept of 'transformational change' as formulated in the INTEGER project, and the evaluators defined three indicators and seven sub-indicators of 'transformational change' in this context. These indicators fed into a web survey that was conducted among the staff of the pilot entities and the control groups at the INTEGER partner institutions.

In addition, quantitative personnel data was collected during both evaluation rounds. Whereas the qualitative analysis provided at least some interesting results regarding the impacts, no impacts could be measures in quantitative terms yet. The short timeframe of the EU project INTEGER in view of the slow pace at which transformation change tends to manifest itself are considered to play a significant role in this regard.

Providing for the accessibility of interviews and/or surveys and facilitating 'freedom of expression'

In order to reduce bias, efforts should be made to make the interviews, group discussions and/or surveys as accessible as possible, and to guarantee – to the extent possible – 'freedom of expression' during these activities. For this purpose, it is to be made sure that all potential participants receive the invitation (which may not be the case if certain restricted mailing lists are used), that the time to take part in the interview or survey does not exceed what is considered bearable (which differs from institution to institution and across disciplines), and that interviews can be scheduled flexibly regarding both time and place. Anonymity needs to be guaranteed, and this is to be communicated to potential participants. It is also recommended to use language that is generally understandable, and to resort to practical examples where possible. Furthermore, it should be considered to involve interpreters in interviews or group discussions with individuals who are not native speakers in the language in which the self-assessment activities are carried out and/or who communicate in sign language.

Example from the INTEGER project

In the INTEGER project, a particular challenge regarding the accessibility of the interviews and group discussions were the language barriers the evaluators – all of whom are German native speakers – faced in all partner institutions to a different degree.

During their visits to CNRS and Šiauliai University the evaluators were accompanied by interpreters (English-French-English and English-Lithuanian-English, respectively). The position and role of the interpreters as regards the organisational setting differed: Whereas in one case the professional interpreters were hired externally, in the second case the interpreter was a member of the institution. Both cases carry specific risks: Whereas external interpreters tend to be less familiar with the organizational structure and current developments in the institution, which may pose difficulties to their understanding of some discussions, interpreters who belong to the same institution as the interviewees may be tempted to influence the discussions. In both scenarios, a thorough briefing of the interpreter(s) is of high importance.

For the web survey, the evaluators had the questionnaires translated into French, for the data collection at the CNRS, and into Lithuanian, for the data collection at Šiauliai University. The limited resources available for translations led to the evaluators' decision against open-ended questions as the answers would have needed to be translated at least in the case of the questionnaire in Lithuanian.

The project coordinators in the partner institutions supported the evaluators in adequately taking national, institutional and/or local specificities into account when scheduling and conducting interviews. Thus, the evaluators learned that in one institution the order of the interviews had to be guided strictly by the professional hierarchy (with the highest-ranking interviewee to be met before all others), and in which settings it was common to conduct interviews in the interviewees' offices and in which ones it was considered more appropriate to schedule the interviews in meeting rooms.

<u>References</u>

Squires, Allison (2009): Methodological challenges in cross-language qualitative research: A research review, International Journal of Nursing Studies, 46:2, pp. 277-287.

4 The self-assessment process III: Feedback and follow-up

Feedback and follow-up activities constitute the final phase of the self-assessment process. Providing feedback is understood as presenting the self-assessment results to target groups within the institution. The follow-up process, in turn, is the next step following the feedback. It implies moving along the management cycle (see above [link 'Introduction']. The aim of follow-up procedures is to figure out how results and recommendations from the self-assessment feed into adjustments of policies and structures. Evaluators can initiate follow-up procedures but are not in charge of guaranteeing for the implementation of recommendations from the self-assessment process.

4.1 Feedback

Rationale for feedback

The rationale for providing feedback is rooted in the goals of the self-assessment. The feedback procedure aims at providing the ground for an adjustment of the Transformational Gender Action Plan and/or its implementation process with the aim of optimizing it. Obviously, the presentation of the self-assessment results is a precondition for follow-up procedures to integrate suggested changes. In addition, the feedback procedure can pursue less explicit goals. It can benefit the legitimation of existent gender equality structures and initiatives, and it can fulfil the goal of raising awareness of gender inequalities and the need for (further) initiatives to combat these inequalities more broadly.

Target groups of feedback

The identification of the target groups of the feedback should also correspond to the goal(s) of the self-assessment.

It is to be considered that the selection of the target group influences the self-assessment process. That is because, generally speaking, the bigger the target group of the feedback the less likely it is that respondents/interviewees answer openly, given that the respondents are informed about the feedback and follow-up procedures in advance, which is highly recommended due to transparency reasons. It can be expected that, even if anonymity is guaranteed, some respondents may adjust their answers if they know that the self-assessment results will be presented to e.g. a broader public or funding authorities. In this case, respondents may be less critical of internal processes etc. out of a sense of responsibility of their own institution, for example.

A distinction can be made between direct and indirect target groups of feedback. Direct target groups can be defined as all actors who are in charge of the design, implementation and/or modification of the evaluated Transformational Gender Action Plan. This group is likely to include senior management at institutional and/or subordinate level(s) and gender equality actors within the institution. In any case, feedback is to be provided to the person or entity that commissioned the self-assessment.

In contrast to these direct target groups that should receive comprehensive feedback on the self-assessment results, indirect target groups may be included as long as this serves the goal(s) of the self-assessment. These indirect target groups can include administration staff (e.g. the HR officer), certain other groups of staff or even all staff, students and funding authorities, amongst others.

The feedback methods and the type of feedback can differ between the two types of target groups. To give an example, it may prove useful to provide feedback to funding authorities after having taken into account reactions to feedback within the institution, and to put emphasis on aspects of particular relevance to these authorities.

<u>Example from the INTEGER project</u>

The direct target group of the feedback identified by the evaluators are the TGAP managers and the implementation teams, on the one hand, and the top-level decision-makers within the partner institutions, on the other hand. Both groups received oral feedback on the preliminary evaluation results; this feedback took place during a seminar that brought together the implementation teams of all three implementing partner institutions, and a meeting targeting the leaders of each institution.

Printed versions of the evaluation reports were provided to the local project coordinators and TGAP managers as well as the official 'owner' of the TGAP, i.e. the person who is ultimately responsible of the project and the TGAP (e.g. the president of the institution or the dean of the pilot faculty). In addition, the evaluators sent pdf versions of the reports to the local project coordinators so that they could circulate it for example among the implementation teams and participants of the 'review seminars' [link]. Thus, the project coordinators are free to share the report with any actor they wish.

Feedback methods

Feedback can be provided at different points in time and by means of written and/or oral methods.

A first feedback can be provided via a de-briefing directly after the collection of qualitative data on site. Such a brief presentation of general impressions is best targeted at the person(s) responsible of (the implementation of) the evaluated Transformational Gender Action Plan. It should be clearly communicated that this feedback is based on first impressions and that these may be subject to adjustments in the course of the structured analysis. The main aim of this early feedback is to satisfy curiosity and tension related to the self-assessment results that are usually keenly awaited.

The most prominent, very useful feedback method is to provide a report pointing out the self-assessment concept, the self-assessment methods and the results of the analysis, including recommendations for adjustments of the gender equality plan and its implementation (see the Toolkit for a Self-Assessment Report template [link]). It is recommended to attach an executive summary to the report in order to put an emphasis on the most relevant results and to make it more likely that e.g. key decision-makers with limited availabilities who are not very committed to the topic of gender equality are informed about the main issues.

The quality of the report benefits from including a feedback loop for the correction of mistakes. A person responsible of (the implementation of) the evaluated gender action plan or possibly another actor with relevant expertise is to be asked to point out evident mistakes that can then be corrected by the evaluators. When communicating this request it is to be underlined that any comments that go beyond the correction of mistakes, e.g. requests to modify the recommendations, are not taken into account.

The final self-assessment report can be circulated among the target groups electronically or in printed form. It is recommended to make an electronic version available anyway in order to allow for further circulation of the results, if desired.

In order to provide the person(s) responsible of (the implementation of) the gender action plan with the possibility to react to the self-assessment results, it has proven useful to invite them to issue a statement on the final report and to attach it when distributing the report in the follow-up phase.

An oral presentation of the self-assessment results can be very fruitful as it increases the visibility of the results among key target groups and creates space for direct questions and first discussions. The results can be presented to all relevant target groups in a joint meeting or workshop, or presentations can be adjusted to the specific interests of certain target groups and are thus to be held to different target groups separately. It may be useful to tailor the presentations to key decision-makers (providing a summary of the results with an emphasis on the needs for action), on the one hand, and persons in charge of the implementation of the gender action plan (presenting a more detailed version), on the other hand.

The oral presentation(s) can take place after or before the written report is circulated among the target group, or may replace a written report altogether. However, the latter is not recommended, as only a written report provides space for a detailed and comprehensive presentation of methods, results and recommendations, and is more likely to reach the entire target group.

The presentation of the results may take place in the framework of a workshop that includes a first discussion on the follow-up, i.e. on how to modify the gender action plan and/or the implementation process based on the self-assessment results.

Based on the experiences gathered within the INTEGER project it is recommended to combine written and oral feedback for the benefits of both forms mentioned above, and to include a de-briefing directly after the site visit. The latter has proven to contribute to strengthening the role of the evaluator as a 'critical friend' (cf. Introduction [link]).

Example from the INTEGER project

Within the INTEGER project, the lead evaluator conducted a short de-briefing to the local programme coordinators directly after the evaluation visit without opening up a space for discussion.

Before finalizing the evaluation reports, preliminary evaluation results were presented orally towards all project partners and the teams involved in the implementation of the Transformational Gender Action Plans during a session of about 90 minutes length. Additionally, the lead evaluator presented a less detailed overview of evaluation results to key decision-makers and official TGAP owners in the framework of the *INTEGER partnership group meeting.*

Both occasions provided the opportunity for questions and remarks, yet, due to time restrictions only to a limited extent. Anyway, the fact that the audience had not read the reports made it difficult for them to provide a comprehensive feedback on the results.

An evaluation report for each institution was provided in English, and a brief executive summary of the report was attached in English (Trinity College Dublin), French (CNRS) and Lithuanian (Šiauliai University). Having received the finalized reports the local project coordinators were invited to issue a statement on the report which they did. The evaluation reports and these statements provided the basis for discussions on the follow-up among the local project team, decision-makers and external experts.

References

Arlen Gullickson & Daniel Stufflebeam (2001): Feedback Workshop Checklist, available at: http://www.wmich.edu/evalctr/archive_checklists/feedbackworkshop.pdf

European Commission (2003): Guidelines for Dissemination and Feedback of Evaluations, available at: http://ec.europa.eu/europeaid/evaluation/methodology/examples/guidelines_feedback_en.pdf. *(This document – although targeting EuropeAid's Evaluation Unit – may serve as an example of how to formulate the provisions on feedback in advance.)*

Reactions to feedback

Reactions to presentations of the self-assessment results can differ between actors and can range from constructive reactions over low interest to defensive demeanor. It is to be underlined that the objective of the self-assessment and the feedback is not to please the target groups. Techniques to influence the reactions to the feedback do not aim at putting the results in a more positive light but to reach the target groups adequately and thus to increase the chances that the feedback is taken into account. In the following, some techniques that may impact reactions to feedback are presented.

In general, involving the target group, i.e. key actors, in the self-assessment process from an early stage tends to increase the acceptance of the self-assessment as a whole and the self-assessment results in particular. Grounding the self-assessment in a sound conceptual and methodological framework and explain this framework to the target audience further increases the credibility of the evaluators and possibly the acceptance of the results. In written and oral presentations of self-assessment results it is important to communicate both positive and negative feedback; it is recommended to start with positive aspects.

Furthermore, it is useful to consider some other communication techniques like figuring out how to speak the audience's language.

Example from the INTEGER project

It required a thorough communication of the objectives of the evaluation to the project partners in advance and to the implementation teams during the evaluation visit to convince them of the fact that the evaluation aims at supporting their activities, and that pointing out certain room for improvement does not automatically imply criticism of their work.

The evaluators were satisfied with the great interest in the evaluation results and recommendations shown not only by the project partners but also by a variety of actors involved in the implementation of the Transformational Gender Action Plans.

4.2 Follow-up

As explained above, the follow-up to the self-assessment constitutes the next step on the management cycle [link], initiating the modification of strategies, structures and practices on the basis of the self-assessment results. A recommendable method is to hold a workshop or various workshops convening key decision-makers and gender equality practitioners within the institution and to initiate discussions on how to modify the Transformational Gender Action Plan and/or the implementation process based on the self-assessment results. These workshops or the follow-up process in general may be supported by external experts (see section on 'quality management' [link]). Furthermore, it may be worthwhile to disseminate the self-assessment results among indirect target groups (e.g. funding authorities) as well in order to influence political processes.

The next step to follow is the actual implementation of recommendations from the self-assessment or, in a broader sense, adjustments drawing on self-assessment results. It is important for the evaluators to see and communicate clearly that this process is not covered by the self-assessment mandate. It is the decision-makers within the institution who are responsible for strategy development.

The follow-up may also include a retrospective quality assessment of the self-assessment in view of implications for future self-assessments. A final reflection on the self-assessment process is recommended for quality management reasons, and it is highly recommended to document this assessment in order to be able to resort to these lessons learnt in the run-up to future self-assessments.

Example from the INTEGER project

Following the presentation of the preliminary evaluation results to the project partners and the members of the implementation teams from the three implementing institutions, these members of the target group and the evaluators split into working groups for discussions on further steps to be taken in the implementation of the Transformational Gender Action Plans. Yet, as this session brought together members of all partner institutions, specific recommendations could not be addressed in detail. However, the participants benefitted from this exchange of experience.

The implementing partners were in charge of organising specific follow-up seminars ('review seminars') during which the evaluation reports were discussed in detail among the actors involved in the implementation of the Transformational Gender Action Plans, key decision-makers in the institution and external experts.

The evaluation team conducted a retrospective reflection on the evaluation itself directly after each evaluation round, and more comprehensively in the process of preparing these guidelines.

References

World Health Organization (2013): WHO Evaluation Practice Handbook, available at: http://apps.who.int/iris/bitstream/10665/96311/1/9789241548687_eng.pdf; pp. 66ff.

5 Quality assurance of self-assessments

Just as the self-assessment plays an important role in the quality assurance of the evaluated Transformational Gender Action Plan and its implementation,[11] attention is to be paid to the quality assurance of the self-assessment itself.

Systematic quality assessment assures, amongst others, the accuracy, validity, soundness and utility of self-assessment results. That is why quality assurance tends to increase the credibility of the self-assessment among key decision-makers and other actors,[12] and thus potentially the impact of the self-assessment results on strategy development and the management of the gender equality plan.

It is worthwhile to take the time at the beginning of the self-assessment to reflect upon how to assure the quality of the self-assessment, to examine different methods and instruments and to establish a quality management plan pointing out standards and instruments most suitable for this self-assessment. A comprehensive reflection on the self-assessment can take place in the course of the process and/or retrospectively.

Overview of approaches to quality assurance of self-assessments

The quality of self-assessments and their results can be assured drawing on in-house quality management instruments and/or by resorting to external experts. It is recommended to use basic quality management instruments like documentation and reflection on the self-assessment process in any case, and to consider additional options. Yet, the choice of options to some extent depends on the financial resources available, as it may be costly to hire external experts, in particular.

A manifest form of in-house quality management is checking the self-assessment performance against well-established evaluation standards (see section 'references' [link] below). To give an example, the general assembly of the German Evaluation Society (DeGEval) suggests assessing the quality of evaluations by their utility, feasibility, propriety and accuracy. It may also make sense to adapt a certain standard to the specific circumstances of the evaluation, e.g. to specify what it implies in the specific context.

Evaluation guidelines can provide valuable guidance throughout the implementation process and can thus help to assure the quality of the self-assessment. The guidelines developed within the INTEGER project aim to fulfil this purpose. Please consult the section 'references' [link] below for a list of other potentially useful guidelines and checklists. The available guidelines range from guidelines covering the whole evaluation process – like the INTEGER guidelines – to guidelines specifically targeting one phase or element of the evaluation process, like the evaluation report.

A thorough documentation of the self-assessment also contributes to quality assurance and facilitates quality control. According to (social) science standards, the conceptual framework of the self-assessment, the methods that are/were applied, sources of information and information on collected data are to be documented. It is recommended to make this information available (at least) to the direct target group of the self-assessment feedback. As regards the publication of (specific information on) collected primary data, data protection provisions must be taken into account.

The quality assurance of the self-assessment may also be assisted by external experts. It can be considered most useful to involve either evaluation experts or gender equality practitioners (or other actors in higher education institutions) with experience in the evaluation of gender equality plans in this peer counselling process. The experts can lead or support a joint reflection on the self-assessment process, and may offer their advice either for free or against remuneration. Apart from the expert's knowledge and experience in the field of evaluation, the process also benefits from their independence as external actors tend to be more likely to

[11] Cf. e.g. Conole, Grainne (2004): The Role of Evaluation in the Quality Assurance of eLearning, in: Learning & Teaching in Action, 3 (2), p. 17, available at: http://www.celt.mmu.ac.uk/ltia/issue8/conole.pdf.

[12] Cf. e.g. Office of Energy Efficiency & Renewable Energy of the U.S. Department of Energy (2006): EERE Quality Assurance Guidelines for General Program Evaluation Studies, available at: http://www1.eere.energy.gov/analysis/pdfs/eval_guide_appendix_7.pdf.

identify specific characteristics of the local context and may formulate criticism more openly. On the down side of this option, external actors possibly lack an overview of the institutional situation and insights into specific contexts, and hiring external experts may be cost-intensive.

Striving to implement a self-assessment of highest quality the evaluators will anyway be confronted with a number of constraints, often caused by the limited time and budget as well as data gaps. It is crucial to be aware of these constraints and subsequent gaps, and to highlight them in the final self-assessment report.

Example from the INTEGER project

Throughout the evaluation process, the evaluators reflected on their performance against evaluation standards as lined out in the Standards developed by the German Evaluation Society (DeGEval) [link].

Comprehensive information on the conceptual framework as well as methods of data collection and analysis was provided by the evaluators in the detailed evaluation concept that was shared with the project partners and is available on-line [link: http://www.gesis.org/cews/en/about-cews/cews-project-work-en/eu-projects/integer/]. Both the three evaluation reports and the oral presentations (see section on 'feedback' [link]) did not only present the evaluation results but also documented the approach rooted in evaluation research and the methods applied.

The evaluators at GESIS considered to hire external experts to accompany them on the three evaluation visits. For this task the evaluators selected experts who had given advice to the local project coordinators at CNRS, Šiauliai University and Trinity College Dublin, respectively, in the course of the project. This choice is rooted in the fact that the evaluators themselves were external to the institution implementing the gender equality plan. Thus, these academics were expected to assure the quality of the evaluation results by double-checking the assessments made by the evaluators in the light of the insights into the realities and initiatives in the INTEGER partner institutions they had gained when being consulted before. In the end, this option had to be dismissed out of financial reasons.

References

United Nations Development Programme (UNDP) (2009): Handbook on Planning, Monitoring and Evaluating for Results, available at: http://web.undp.org/evaluation/evaluations/handbook/english/documents/pme-handbook.pdf, pp. 163ff.

World Bank (2006): "Conducting quality impact evaluations under budget, time and data constraints", Independent Evaluation Group (IEG) working paper series, Washington, D.C., available at: http://documents.worldbank.org/curated/en/2006/06/7287644/conducting-quality-impact-evaluations-under-budget-time-data-constraints.

Standards

DAC Network on Development Evaluation (2007): DAC Evaluation Quality Standards, available at: http://www.oecd.org/development/evaluation/dcdndep/38686953.pdf.

DeGEval – Gesellschaft für Evaluation e.V. (2001): Evaluation Standards (DeGEval-Standards), available at: http://www.degeval.de/images/stories/Publikationen/DeGEval_-_Standards_short.pdf.

[Available only in German: DeGEval – Gesellschaft für Evaluation e.V. (2004): Empfehlungen zur Anwendung der Evaluationsstandards im Handlungsfeld der Selbstevaluation, available at: http://www.degeval.de/fileadmin/Publikationen/DeGEval_-_Empfehlungen_Selbstevaluation.pdf.]

United Nations Evaluation Group (UNEG) (2005): Standards for Evaluation in the UN System, available at: http://www.uneval.org/document/detail/22

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011): The program evaluation standards: A guide for evaluators and evaluation users, 3rd Edition, Thousand Oaks: Sage.

Guidelines for evaluations:

(These guidelines do not specifically address the evaluation of Transformational Gender Action Plans in higher education and research institutions but may nevertheless provide some useful information.)

Austrian Development Agency (2009): Guidelines for Project and Programme Evaluations, available at: http://www.entwicklung.at/uploads/media/Guidelines_for_Project_and_Progamme_Evaluations_Juli2009_02.PDF.

Canadian International Development Agency (2004): CIDA Evaluation Guide, available at: http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUImages/Performancereview5/$file/english-e-guide.pdf.

European Communities (2006): Evaluation Methods for the European Union's External Assistance, available at: http://ec.europa.eu/europeaid/evaluation/methodology/examples/guide1_en.pdf.

Molund, Stefan & Schill, Göran (2007): Looking Back, Moving Forward. Sida Evaluation Manual, 2nd revised version, available at: http://www.sida.se/globalassets/publications/import/pdf/en/looking-back-moving-forward_2561.pdf

Scriven, Michael (2013): Key evaluation checklist, available at: http://www.michaelscriven.info/images/KEC_3.22.2013.pdf (draft).

World Food Programme Office for Evaluation (2013): Evaluation Quality Assurance System for Impact Evaluations – Guidance Materials, available at: http://documents.wfp.org/stellent/groups/public/documents/reports/wfp231096.pdf.

Checklists for/quality criteria of evaluations

The Evaluation Center at Western Michigan University provides a collection of checklists for evaluations available at: http://www.wmich.edu/evalctr/checklists/.

Recommended checklists featured in this collection include:

Stufflebeam, Daniel L. (2007): CIPP Evaluation Model Checklist, available at: http://www.wmich.edu/evalctr/archive_checklists/cippchecklist_mar07.pdf.

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

Stufflebeam, Daniel L. (1999): Program Evaluation Metaevaluation Checklist, available at: http://www.wmich.edu/evalctr/archive_checklists/program_metaeval_10point.pdf (long version) or at http://www.wmich.edu/evalctr/archive_checklists/program_metaeval.pdf (short version).

Checklists for/quality criteria of evaluation reports

Canadian International Development Agency (2002): How to Perform Evaluations No. 7 – Evaluation Reports, available at: http://www.oecd.org/derec/canada/35138852.pdf.

European Commission (1999): Quality assessment grid for assessing the quality of evaluation reports, http://ec.europa.eu/europeaid/evaluation/methodology/guidelines/gui_qal_flr_en.htm.

United Nations Evaluation Group (UNEG) (2010): UNEG Quality Checklist for Evaluation Reports, available at: https://www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/about_iom/eva_techref/UNEG_Eval_Report.pdf.