



**INSTITUTIONEN FÖR SOCIOLOGI  
OCH ARBETSVETENSKAP**

# **MONITORING HANDBOOK**

Methods and tools for monitoring developed in the *GenderTime* project

Helen Peterson & Jennifer Dahmen

Samling "Gothenburg Studies in Work Science"  
Nr 1, 2018

ISBN 978-91-87876-19-6

# CONTENT

- INTRODUCTION .....3**
  - THE GENDERTIME PROJECT ..... 3
  - THE MONITORING HANDBOOK ..... 3
- DEVELOPING A MONITORING PLAN .....4**
  - EVALUATION AND MONITORING IN THE GENDERTIME PROJECT .....4
  - THE RELATION BETWEEN MONITORING AND EVALUATION .....6
  - TIMING AND FREQUENCY OF MONITORING ACTIVITIES .....7
  - MONITORING METHODS ..... 8
  - THE ADVANTAGES OF MONITORING.....9
  - A SYSTEMATIC APPROACH ..... 11
- THEORETICAL CONTEXTUALISATION.....12**
  - EVALUATION SCIENCE ..... 12
  - THE INTERPRETATIVE, DIALOGUE- AND CUSTOMER-ORIENTED APPROACH ..... 13
  - THE THEORY-DRIVEN APPROACH..... 13
  - THE REALISTIC APPROACH ..... 15
  - THE FEMINIST APPROACH..... 16
- EMPIRICAL CONTEXTUALISATION .....19**
  - MONITORING AS KNOWLEDGE EXCHANGE ..... 19
  - MONITORING AGENTS, MULTIPLE PROJECT ROLES AND MONITORING BUDGET ..... 19
  - QUALITATIVE AND QUANTITATIVE MONITORING TOOLS.....20
  - DISTRIBUTION OF THE MONITORING TOOLS DURING PROJECT PHASES .....21
- PRESENTATION OF THE GENDERTIME MONITORING TOOLS.....23**
  - 1. CULTURAL STAFF SURVEY .....23
  - 2. NATIONAL SURVEY REPORT .....26
  - 3. EXCHANGE WORKSHOP .....28
  - 4. OPERATIONAL PROCESS MONITORING TOOL .....30
  - 5. SELF-ASSESSMENT OF CHANGE AGENT ROLE .....33
  - 6. INTERIM FEEDBACK REPORT .....36
  - 7. PEER CONSULTATION REFLECTION SESSION.....38
  - 8. INCREMENTAL TRANSFORMATION MONITORING TOOL.....41
  - 9. MOST SIGNIFICANT CHANGE TECHNIQUE .....45
  - 10. FINAL FEEDBACK REPORT .....48
- RECOMMENDATIONS FOR MONITORING.....51**
- REFERENCES.....53**

## **Introduction**

### **The GenderTime project**

GenderTime was an action, support and intervention programme funded by the European Commission between January 2013 and December 2016. The purpose of GenderTime was to promote organizational structural change and increase the participation and career advancement of women researchers in seven research and higher education institutions in seven European countries.<sup>1</sup>

The method for generating change that was adopted in the GenderTime project utilized Gender Equality Plans (GEPs), which were tailor-made to fit each of the seven participating institutions. The implementation of the GEPs was managed and organized by seven national GenderTime teams, whose members constituted the GenderTime consortium.

The seven GEPs together organized almost 200 actions and change interventions that were implemented during four years. The actions targeted organizational and managerial processes and procedures and aimed at for example creating gender sensitive recruitment, retention and promotion policies; supporting and improving work-life balance; establishing a more inclusive work culture, and; increasing gender awareness throughout the organizations.

The implementation of the GEPs in the participating institutions was carefully monitored throughout the project. The name GenderTime indicates the prominent place monitoring had in the project as the acronym TIME stands for: Transferring, Implementing, Monitoring Equality.

### **The Monitoring Handbook**

This Monitoring Handbook offers guidelines and recommendations concerning monitoring which were developed and applied during four years of monitoring activities in the GenderTime project. It also includes detailed information about the specific monitoring tools that were designed and implemented within the GenderTime project. As a Handbook, it thus presents the reader with “guidance on ‘how to’ and practical tools” (UNDP 2009: 2) that can be used to strengthen monitoring efforts in change programmes. It is meant to provide inspiration about

---

<sup>1</sup> GenderTime received funding in the 7<sup>th</sup> Framework Programme of the European Commission. Consortium members were: Egalité des Chances dans les Etudes et la Profession d’Ingénieur en Europe, France (coordination); Interdisziplinäres Forschungszentrum für Technik, Arbeit und Kultur, Austria, Università degli Studi di Padova, Italy; Gothenburg University, Sweden; Université Paris Est Créteil, France; Mihailo Pupin Institute, Serbia; Bergische Universität Wuppertal, Germany; Loughborough University, UK; Fundacion TECNALIA Research & Innovation, Spain; Donau-Universität Krems, Austria.

how to think creatively about different options available by presenting a range of tested methods and tools.

We believe that the handbook, by providing this type of practical information, can fill a gap not only in the existing literature on evaluation and monitoring but also be a contribution to the literature on change management and action research. We hope that the handbook can be useful to others involved in gender equality change projects. The handbook can, hopefully, also be relevant for those involved in change projects more generally, not exclusively targeting gender equality. The overall purpose of GenderTime is common for many change programmes, i.e. to generate “changes in patterns of behaviour, events or conditions by bringing fresh inputs to that system in the hope of disturbing and re-balancing it” (Pawson & Tilley 2004:3). Many of the practical guidelines and recommendations are wide-ranging and of a general character, meaning that even if they do not constitute a standard set of tools and methods that can be adopted to any situation, they are not specifically or necessarily linked to intervention programmes dealing with gender equality.<sup>2</sup>

## **Developing a monitoring plan**

### **Evaluation and monitoring in the GenderTime project**

In the GenderTime project, monitoring and evaluation were distinct project activities designed to complement each other. An external team, not involved in the implementation of the (GEPs), performed the evaluation.

The external evaluation team developed and used different evaluation tools from month 1 of the project, covering the cooperation and collaboration in the GenderTime consortium as well as the long-term impact of the GenderTime project. Two types of evaluation were thus used in the GenderTime project: impact evaluation and progress evaluation. Impact evaluation focused efficiency and sustainability in the outcomes and results in comparison to the stated objectives and goals, and external accountability. Progress evaluation focused the internal process of learning in the project, internal accountability, cooperation and collaboration between the project members (Siebenhandl & Mayr 2015).

Monitoring, as used in the GenderTime project, refers to two different types of activities. The first type of activity was the monitoring of certain indicators of existing gender inequalities and gender segregation in each GenderTime institution. This type of monitoring was included in several of the objectives of the seven GEPs that the GenderTime teams implemented. These objectives dealt with

---

<sup>2</sup> A previous, shorter, version of this Handbook was presented on April 5<sup>th</sup> 2017 at the Working Life and Welfare Seminar, the Department of Sociology and Work Science, Gothenburg University, Sweden. Insightful and helpful comments and suggestions from the opponent Gunnar Gillberg, and from other seminar participants, are gratefully acknowledged by the authors.

reviewing and monitoring the recruitment process, salaries, career development and career opportunities, and funding and fellowships (cf. Barnard et al. 2013a). Establishing monitoring systems such as these was also an inherent part of the action research approach that characterized the GenderTime project (Ahnberg et al. 2010).

The second type of monitoring, and the type that this Handbook deals with, was organized within the monitoring work package. The monitoring activities developed and organized in this work package encompassed the overall implementation process of the GEPs in the GenderTime institutions. This type of monitoring involved an on-going collection of project data in order to assess whether the project was going in the right direction and complied with the pace and stages set (cf. Kriszan & Lombardo 2013). In the GenderTime programme this kind of monitoring was defined as an activity separate from the evaluation system that focused the outcomes of the project in comparison to the stated objectives and goals. The purpose of monitoring was thus to systematically and continuously document key aspects of the programme performance that assessed whether the programme was operating as intended (Rossi et al. 2004). Monitoring, as used in the GenderTime programme, therefore had a focus on the implementation process and can be referred to as process monitoring combined with a form of performance monitoring.

In contrast to evaluation, which was performed by an external team, monitoring in the GenderTime project was internal, meaning that the monitoring team regularly reported back to the implementing national GenderTime teams the results of the monitoring activities (cf. OECD 2002). Monitoring in the GenderTime project was a team effort and a collaborative undertaking (Rossi et al. 2004) which involved self-assessment (Lipinsky & Schäfer 2016) and participatory monitoring tools (Haylock & Miller 2016) used by all seven implementing national GenderTime teams.

The internal aspect also concerned the role of the two leaders of the monitoring work package, i.e. the monitoring agents who are also the authors of this handbook. Both were members of the GenderTime consortium's core management team and one of the two leaders of the monitoring work package was also the team leader of one of the national GenderTime teams, and therefore also played an active role in the implementation of the GEPs (Chen & Rossi 1989).

The development of a plan for monitoring of the overall implementation process in the GenderTime programme was guided by questions such as: "What are the issues, risks and challenges that we are facing that need to be taken into account to ensure the achievements of results? What decisions need to be made concerning changes to the already planned work in the subsequent stages? What are we learning?" (cf. UNDP 2009: 82). Process monitoring also means focusing on (intermediate) outputs like gradual learning, knowledge, skill or attitude changes in

people involved in the intervention programme rather than (long-term, major) impacts (Scriven 2013).

### **The relation between monitoring and evaluation**

Evaluation is generally defined as a structured way of strategically examining and assessing the results of different intervention programmes in order to learn from successes and challenges how to continue and move forward (Vedung 2010). The definition is, however, somewhat less straightforward when taking into consideration the many different types of evaluation that exist. While summative evaluation (or impact evaluation) is conducted at the end of an intervention to provide information about the extent to which anticipated outcomes were produced, formative evaluation (or implementation evaluation) is conducted during the implementation phase of an intervention in order to improve the performance (OECD 2002). The formative evaluation is therefore also often a type of process evaluation that focuses on the internal dynamics of the implementing organization (Rossi et al. 2004).

Monitoring, on the other hand, is sometimes narrowly understood as a method for data collection to find out what is happening in an on-going project or programme. To find out *why* this is happening you need to evaluate the collected data (Kusek & Rist 2004). Contrary to evaluation, this definition of monitoring does not imply that you do anything special with the data (Funnell & Rogers 2011). As a rule, careful monitoring, i.e. the collection of data, is understood as facilitating evaluation and as necessary for evaluation, although not sufficient for evaluation.

Another way of defining monitoring is as: “the ongoing process by which stakeholders obtain regular feedback on the progress being made towards achieving their goals and objectives” (UNDP 2009:8). With a more limited definition this means reviewing the progress made in implementing activities. In this Handbook, we refer to monitoring in a broader sense, meaning that monitoring also involves reviewing progress against achieving goals (UNDP 2009). This broader definition means that evaluation and monitoring are closely related concepts and as such they often seem to refer to overlapping activities.

Previous research on evaluation sometimes includes what we here define as monitoring in the definition of evaluation, for example when referring to process-oriented or on-going evaluation (closely related to what we call monitoring) as a specific part of evaluation, separate from impact-based or outcomes-based evaluation (what we call evaluation) (Upton et al. 2014). On-going evaluation is, however, typically also associated with a top-down-perspective (cf. Ahnberg et al. 2010), which distinguishes it from monitoring as defined in this Handbook. The distinction between monitoring and evaluation has been depicted as constituted by exactly this feature: that evaluations are done independently to provide decision-makers with an objective assessment of whether the intervention is on track. In addition to this, evaluations have been described as “more rigorous in their

procedures, design and methodology”, and involving a “more extensive analysis” than monitoring (UNDP 2009: 8).

More often, however, the distinction between evaluation and monitoring is blurry or at least utterly complex. Williams and Hummelbrunner (2010) developed the principles for what they call Process Monitoring of Impacts, which entails finding out whether intervention programmes “contribute to intended effects and what actions should be taken during implementation to improve effectiveness” (Williams & Hummelbrunner 2010: 94). This means recognizing and classifying the intended and expected effects, outcomes and objectives but not focusing on causal relations and attributing effects to a specific intervention. Instead, the main task is to understand the less linear manner in which the interventions influence actors and to which extent they actually lead to the impacts intended (Williams & Hummelbrunner 2010).

Rossi et al. (2004) define what they call programme process monitoring as a type of process evaluation (or implementation assessment), that is, as an on-going function in an intervention programme that involves repeated measurements and recordings of information about the programme’s operations over time (cf. Kusek & Rist 2004). They also emphasize the overlapping elements regarding the data collected and the data collection procedures used in evaluation and monitoring (cf. Kusek & Rist 2004). In accordance with this definition of (process) monitoring proposed by Rossi et al. (2004) the definition of monitoring adopted in this Handbook is somewhat less narrow and also includes analysing the data collected through monitoring activities.

### **Timing and frequency of monitoring activities**

Some of the most apparent distinctions between evaluation and monitoring activities regards timing and frequency of the activities. While evaluation can take place after a project starts and often continues even long after it ended, monitoring is often initiated already in the planning phase of a project and is continuously performed during a project in progress (Holbrook & Frodeman 2011). Usually then, monitoring is performed before evaluation and provides the basis for later evaluation activities. Evaluation results can, however, also be fed back to the continuous planning of monitoring activities (UNDP 2009). There are, however, also variations when it comes to these dimensions. Funnell and Rogers (2011), for example, distinguish between different types of monitoring: routine, regular and frequent monitoring. And besides the more common ex-post evaluation that takes place after the intervention has been completed, there is also something called ex-ante evaluation that is performed before the implementation, focusing on appraisal and quality at entry (OECD 2002). Nonetheless, evaluation is usually undertaken after the completion of interventions with the intention to assess the sustainability of results and impacts and to identify the factors of success or failure (OECD 2002).

Monitoring needs to be performed on regular basis in a project. Sometimes it is even described to be an activity that should be executed in a routine-wise manner (Funnell & Rogers 2011). The character and design of the monitoring activities should also be adapted to the specific phase of the implementation process and whether the project is in the launching phase with early implementation interventions, in the main implementation phase or in the final reflection phase where the interventions are being completed. Monitoring plays different roles in the different phases (Lipinsky & Schäfer 2016; UNDP 2009).

The role of monitoring during the launching phase is to gather information and data to establish relevant baselines that can be used for time-related comparisons of before and after the implementation started (Kusek & Rist 2004). This launching phase can benefit from quantitative monitoring tools. The reflection phase, in comparison, requires qualitative monitoring tools (Rossi et al. 2004). The intensity of the monitoring activities should also be adapted to the specific phase. The implementation phase, for example, might call for a concentration of more monitoring tools than the other phases. It should, however, be noted that monitoring takes time away from the implementation. To find a balance is therefore important, between monitoring regularly and allowing enough time for implementation of the interventions for the project members and stakeholders.

### **Monitoring methods**

Planning monitoring activities in more detail entails making important decisions regarding which data to collect. Such decisions should be informed by considerations about which data that can be considered as relevant and reliable measures of achievements, performance and progress. The methods used to gather data usually differs between evaluation and monitoring (Kusek & Rist 2004). Monitoring has been associated with ad-hoc or periodic reviews and rapid and light assessments of the performance and operational issues of an intervention, performed by those internal to the intervention programme or the organization where it is implemented (Rossi et al. 2004). Evaluation, on the other hand, often involves more thorough assessments and rigorous methodology, conducted by independent evaluators (UNDP 2009).

Evaluation also often uses a quantitative approach with clearly defined performance indicators to measure expected outcomes (Badaloni & Perini 2016). Monitoring, on the other hand, can use a qualitative and more inductive approach, which builds on continuous observations of the activities in the change project (Kusek & Rist 2004). Funnell and Rogers (2011: 418), however, suggest that more qualitative aspects are difficult to routinely monitor, and instead need to be evaluated through occasional surveys.

In addition, already Lincoln and Guba (1985), emphasized the importance of an ethnographic approach and of using qualitative methods for monitoring. Using qualitative methods for monitoring produces rich data (Chen 1990) and qualitative



methods have also been promoted within evaluation (Shaw 2002). As a rule, it is advisable to gather a variety of data, using sophisticated monitoring tools of both quantitative and qualitative character: questionnaires, surveys, interview guidelines, checklists, focus groups, participant observations, statistics, templates and workshop concepts (cf. Lipinsky & Schäfer 2016; Pawson & Tilley 2001; Scriven 2013). Adopting a mixed-methods approach will increase the likelihood that the information about the effectiveness of the interventions, the achievements and progress is as reliable as possible (Haylock & Miller 2016). This includes making informed decisions about *whom* to monitor based on considerations of which actors possess important information regarding the interventions and the implementation (Rossi et al. 2004).

Monitoring in the GenderTime programme was characterised by a mixed-method approach and both qualitative and quantitative monitoring data was collected. Qualitative data, however, was given precedence as it was decided it could provide better evidence of improved gender integration and gender equality. Monitoring in GenderTime was also dedicated to systematize knowledge about the implementation process. As suggested by Funnell and Rogers (2011: 434), comparison was an inherent aspect of monitoring in GenderTime, especially so comparison between actual and planned implementation, and between before and after implementation started. Comparison of different sites was also included among the tasks addressed, as necessary when analysing monitoring data from a complex, multi-sited intervention programme as GenderTime (cf. Rossi et al. 2004).

### **The advantages of monitoring**

It is essential to include a robust system of evaluation and monitoring activities in programmes and projects that set out to produce some kind of change in behaviours, events, cultures or conditions in society, sectors or in organizations (Kusek & Rist 2004). The general purpose of adopting such systems is to find out whether the programme is working as it is expected to and to inform the development of future policies and practices (Pawson & Tilley 2004).

Despite problems with distinguishing between monitoring and evaluation, there are important advantages of the two types of project activities, which makes it relevant to include them both in intervention programmes.

Generally, the major purpose of evaluation is to provide information for policy making (Chen 1990). In addition, if used strategically and systematically the many benefits of monitoring can be essential for the success of a change project (UNDP 2009). One of these benefits is that monitoring contributes to systematizing change interventions and the implementation process (Rossi et al. 2004). Implementation processes in change projects are often complex and multi-layered and can at times also be confusing for those involved (Chen & Rossi 1989). Monitoring is useful as it can produce detailed and structured information about what is happening in the

change project and how the interventions are going. It tracks progress and reports on achievements at different times in the project while also identifying problems (Funnell & Rogers 2011). Monitoring information, if fed back to the practitioners involved in the change process, can thus contribute to their increased motivation and improved organizational and individual learning (both single-loop and double loop learning) (Patton 2011; UNDP 2009). In so doing, monitoring can systematize individual and shared reflection and can be used to provide a framework to facilitate knowledge sharing between participants and stakeholders in a project. These reflections can be used to improve the interventions and the implementation of the change plans, but also to develop plans for how interventions can be adapted to other circumstances and transferred to other settings (Kusek & Rist 2004).

Collecting monitoring information data, while comparing it to the baseline and to the expected outcomes, makes it possible to identify where there is room for improvements. Monitoring identifies whether or not the desired results are achieved and can be used to develop corrective actions to optimize future achievements (Kusek & Rist 2004). This systematic knowledge can be further used to re-adjust objectives and goals and keep them realistic (cf. Kotter 1995). Monitoring can therefore facilitate incremental corrections and improvements (Kriszan & Lombardo 2013).

Moreover, monitoring also provides clues to whether or not an evaluation programme is needed because an understanding of the implementation process can help in the development of strategies appropriate to its measurement (Chen & Rossi 1989). Monitoring can contribute to systematizing the evaluation process as it can identify the need for evaluation and let new important indicators and measures emerge inductively (Funnell & Rogers 2011). Monitoring thus facilitates and lays the foundation for evaluation (UNDP 2009).

Finally, although evaluation also contributes to keeping a change programme on track, monitoring has an advantage when it comes to resources. Impact assessments can be both expensive and time consuming (Chen & Rossi 1989). Funnell and Rogers (2011: 418) describe monitoring as “less expensive and easier to do than evaluation”. Considerations concerning financial resources are vital to consider in the development of a realistic, suitable and useful plan for monitoring, and Lipinsky and Schäfer (2016) emphasize that data collection can be time-consuming. Time and resources therefore always have to be taken into consideration when planning for monitoring activities and they set limitations to these activities (Rossi et al. 2004). Inadequate resources can seriously hamper the quality of monitoring (UNDP 2009). Nonetheless, some methods are less resource intensive than others. Previous research has for instance highlighted that one of the advantages of using self-assessments in evaluation is that it requires minimal resources (D’Eon et al. 2008). In addition, how well an intervention programme is using its resources is important to monitor (Rossi et al. 2004; Scriven 2013).

Efficient, systematic and well-planned monitoring reduces the risk of cost overruns and time delays during an intervention programme (UNDP 2009).

### **A systematic approach**

In order to develop a plan for the best possible monitoring strategy a structured and systematic approach is recommended (Rossi et al. 2004). Such a systematic approach entails; an understanding of non-linear interrelationships and how different components in a system (key variables, context, actors, behaviours, actions, assumptions, results etc.) affect, relate and link to each other; a commitment to multiple stakeholder perspectives, and; an awareness of boundaries concerning what is relevant for an intervention (Williams & Hummelbrunner 2010). This means making strategic and well-informed decisions during each phase of the monitoring process – most urgently about *what*, *who* and *when* to monitor – based on theories about change and about how intervention programmes work (Pawson & Tilley 2004). These decisions can be facilitated by a combined inductive and deductive approach, i.e. by adopting an abductive and pragmatic approach (Patton 2011).

Using a *deductive approach* here means making the decisions based on research literature. Research-based theories about organizational change can provide information that is important in order to understand change processes (activities, contexts, inputs, outputs), problems and effective practices, and thus maximize the relevance and applicability of the monitoring activities (Pawson & Tilley 1997). This is information that also can be obtained from interviews with experienced evaluators or other experts, practitioners and change agents within the relevant field in question (Brouselle & Champagne 2011; Sridharan & Nakaima 2011). This type of information was collected in the GenderTime programme, using expert interviews during the initial phase of developing the plan for monitoring. Previous research on gender change projects was also consulted in order to identify what type of information that can demonstrate a positive change and therefore needs to be collected using monitoring tools (cf. UNDP 2009).

With an *inductive approach* the decisions about *what*, *who* and *when* to monitor are based on information gained from observations in the monitored project. This is an approach akin to the utilization-focused approach to evaluation that requires considerable experimentation and adaptation of available tools and methods (Haylock & Miller 2016; Patton 2011). The inductive approach to monitoring thus entails less emphasis on the process of formulating indicators beforehand (UNDP 2009). The value and usefulness of indicators for the overall monitoring strategy in the GenderTime programme was deemed limited for steering and influencing behaviours of actors involved (e.g. target groups, beneficiaries, implementing teams) and they were therefore downplayed in the plan for monitoring that was developed (cf. Williams & Hummelbrunner 2010). Instead, the inductive approach allowed for tailor making the monitoring tools to fit the needs of the specific project at hand. In GenderTime, the inductive approach to monitoring allowed for

each national team to use the results of the monitoring tools to formulate in detail the specific, quantitative and qualitative, indicators that were relevant to their organization.

As a result, an *abductive approach* to monitoring was adopted in the GenderTime programme. Such an abductive approach, i.e. the combination of a deductive approach with an inductive approach, constituted a sound and solid basis for the development of the monitoring tools where reasoning and dialogue was emphasized as a way of solving problems and promoting inference to the best explanation (cf. Patton 2011).

## **Theoretical contextualisation**

### **Evaluation science**

The overall contribution of this handbook is primarily empirical, with recommendations based on practical and hands-on work performed in the action-research project GenderTime. In order for it to be a Handbook, however, it needs to provide guidance for those attempting monitoring for the first time. This entails knowledge of fundamental principles concerning purposes, norms, and standards for monitoring (cf. UNDP 2009). We here therefore also outline the theoretical context for these empirical observations and accounts. This theoretical context is primarily and above all constituted by concepts, methods and techniques developed by different traditions within evaluation science.

The attempts above, to point out some of the distinguishing features of monitoring and evaluation, can be summarised by emphasising the lack of agreement on clear definitions of monitoring and evaluation in previous research. Another observation that can be made based on an overview of existing literature is that monitoring has attracted much less attention than evaluation. While different evaluation approaches have been developed and argued for, how these affect the systems for monitoring has been left more or less completely unexplored. This Handbook will not further elaborate on the problem with distinguishing between two types of activities that are clearly overlapping. Instead, it will illustrate how the development of a plan for monitoring can draw on evaluation science, thereby benefitting from the advantages of a rich, complex and expanding theoretical field.

As already illustrated, there are numerous perspectives and different evaluation traditions that define evaluation, and monitoring, in slightly different manners (Equality Challenge Unit 2014). Developing a coherent and efficient plan for monitoring presupposes familiarity with some of these different traditions and definitions. Evaluation science, however, as a distinct academic field, is still relatively new and has previously been described as “contested” (Scriven 2002). Further, evaluation is a notoriously multi-faceted concept, referred to by Vedung (2010: 264) as a “nebulous concept”. These circumstances present challenges for

anyone attempting to theoretically contextualise monitoring practices within evaluation science.

Here, we will therefore only briefly refer to the specific theoretical aspects that guided the development of a plan for monitoring in the GenderTime programme. For those who are looking for a more thorough presentation of evaluation we refer to other, more comprehensive, sources (e.g. Funell & Rogerts 2011; Rossi et al 2004).

### **The interpretative, dialogue- and customer-oriented approach**

When trying to learn about the evaluation field, Vedung (2010) suggests that it is helpful to think about how different traditions have developed during the last 50 years as four *evaluation waves*; the science-driven wave, the dialogue-oriented wave, the neo-liberal wave, and the evidence wave. Evaluation practices developed within these four waves have distinct elements, for example regarding how evidence is ranked in the evidence hierarchy (Lindgren 2011). The expression “the evidence hierarchy” refers to different evaluative designs, or monitoring methods, where the science-driven wave gives precedence to experiments and randomized trials (cf. Tilley 2000). The other waves acknowledge the problems with experimentations and therefore emphasize the limited contribution of experiments to our understanding of how programmes work (Chen & Rossi 1983). Instead, these other waves include elements of process evaluation, action research, ethnographic research, and focus on case studies, good practices, and expert opinions as well as user opinion (Vedung 2010).

As will become clear in this Handbook, the monitoring practices developed in GenderTime, were primarily inspired by the interpretative, dialogue-oriented wave and the customer-oriented, neo-liberal wave.

### **The theory-driven approach**

GenderTime is sometimes referred to as a project in this Handbook, which reflects standard academic language and the language of the funding body: the European Commission. Using evaluation language, however, the more correct definition is to refer to it as a programme (Scriven 2013). GenderTime fits this definition due to the complexity of the different sets of interventions, involving multiple activities that cut across several different geographic areas (OECD 2002). GenderTime was thus a large, complex multi-site, action research intervention programme which meant that it had a common set of aims, but that specific questions were addressed at the different sites (Mullen 2002). Within this broader *programme*, the specific national GenderTime *projects* were implemented in each country. Such complex intervention programmes are characterized by being, among other things, uncertain, dynamic, emerging, non-linear and adaptive. These are characteristics which necessarily also should be taken into account when developing a plan for monitoring (Patton 2011).

Change programmes most often have an underlying logic or theory of change, i.e. programme theory. Programme theories define the nature of the problems that need to be corrected, they describe the actions and interventions needed to create change and outline what a successful outcome would look like (Funnell & Rogers 2011; Lindgren 2012). Programme theories thus define how the programme is supposed to work and can be articulated in a series of “if... then” propositions that describe the hypothesized processes by which an intervention programme can bring about change (Sridharan & Nakaima 2011). Programme theories are therefore not always research based theories but can be theories in a looser view of theory (although they can be more or less sparse or rich) that sets out to identify the processes linking program treatments to desired outcomes (Chen & Rossi 1989). Ideally, programme theories are constituted by two components: first, a theory of change that explains the drivers of change; and second, a theory of action that explains how intervention programmes are constructed in order to activate such changes (Funnell & Rogers 2011).

According to Funnell and Rogers (2011: 431) it can be difficult to determine exactly which factors to monitor, especially so in complex programmes, which therefore calls for a monitoring process that is “sufficiently open-ended” in order to identify these factors (cf. also Sridharan & Nakaima 2011). A plan for monitoring that is grounded in programme theory has an advantage when it comes to identifying these factors.

That the overall and main contribution of this handbook lies in the empirically based recommendations does not mean that the monitoring approach adopted was solely inductive and not involved theories. Instead, the approach was influenced by the theory-driven approach to evaluation. Although the plan for monitoring of the implementation process in the GenderTime programme was not developed based on a programme theory as strictly defined by Funnell and Rogers (2011) per se, the system did rely on theories and on specific visions about how the interventions were supposed to produce change. These theories and these visions were also influenced by feminist theories that attempt to explain gender relations and what gives rise to unwanted behaviour, discriminatory events and inequalities of social conditions (cf. Pawson & Tilley 2004). These feminist theories in combination with programme theories formed the basis of the GenderTime programme and also informed the plan for monitoring. As in most large and complex intervention programmes expert knowledge about these theories was necessary in order to develop efficient and valid plans for evaluation and monitoring (cf. Patton 2011).

Feminist theory provided guidelines for determining which methods were most relevant to use when monitoring the activities in the project. In that way, the monitoring in GenderTime was theory-driven, or, perhaps more correctly; theory-oriented and guided by a more or less explicit programme theory about how the intervention programme was supposed to work. This (deductive) approach, however, was combined with a focus on empirical research methods and the

importance of them as useful tools for obtaining empirical knowledge about the implementation (in line with the abductive approach described above) (Chen 1990).

Drawing on the theoretical framework concerning evaluation developed by Chen and Rossi (1989), gender theory can be referred to as a superordinate theory for the GenderTime programme. This means that gender theory contributed to an understanding of the different domains of importance for evaluation, for example the implementation environment domain that concerns the environment under which the program is implemented.

The theoretical basis and how it shaped the monitoring activities more in detail will be further outlined below in the Handbook (p. 19 onwards), specifically in relation to each monitoring tool.

### **The realistic approach**

The premises of the programme theory adopted in the change programme GenderTime were in line with a so-called realist or realistic approach to evaluation (Pawson & Tilley 1997). Using a realistic approach to evaluation entails focusing on what it is about a program that makes it work (Sridharan & Nakaima 2011). It is characterised by an emphasis on the context and the conditions under which interventions are being introduced and thereby drawing on theories characterized by realist components (Pawson & Tilley 2004). More specifically, the realistic approach can be illustrated by two different types of questions asked by evaluators. Instead of asking: “Does this intervention programme work?” the realistic evaluator asks the more complex and multi-faceted question: “What works for whom under which circumstances?” (Tilley 2000). The realist evaluation thus seeks to discover what it is about intervention programmes that works, for whom, in what circumstances, in what respects, and why (Pawson & Tilley 2004).

Adopting this more multi-faceted, realistic and sensitive approach to evaluation (and monitoring) produces nuanced information about how intervention programmes will work in local settings depending on the specific conditions and about how certain contexts will be supportive to the interventions and others will not (Pawson & Tilley 1997). The underlying assumption here is that the complexity of the intervention programme and the context in which it is being implemented entails low predictability, low control and high uncertainty which calls for a monitoring approach that is emergent, dynamic and changing rather than fixed and predetermined (Patton 2011).

The advantage of the realistic approach is the possibility to better understand the varying impacts interventions can have depending on the circumstances and the intervention context, including unwanted outcomes (outcomes defined as changes in the behaviour targeted). The realistic approach understands intervention programmes as embedded and as parts of open systems, which means that they are

never implemented in the same ways (Pawson & Tilley 1997). A basic assumption in the realistic approach is thus that any programmes will have mixed so-called 'outcome patterns'. Outcome patterns involve "the intended and unintended consequences of programmes, resulting from the activation of different mechanisms in different contexts" (Pawson & Tilley 2004: 8). The consequence of this for the planning of monitoring activities is the inclusion of a range of different measures and indicators to assess the programme.

The purpose of the realistic approach is to refine programmes by testing the core underlying theories, and the visions of change, that they are based upon. This is done by investigating whether the planned and the implemented interventions are sound, plausible, valid and practical (Pawson & Tilley 2004). In order to investigate this, one more basic concepts, besides 'context' and 'outcome patterns', is central: 'mechanism' (Tilley 2000). Mechanisms describe what it is in an intervention programme that brings about specific outcomes. Pawson and Tilley (2004) explain in the following way how they use the concept 'mechanism' in realistic evaluation:

*This realist concept tries to break the lazy linguistic habit of basing evaluation on the question of whether 'programmes work'. In fact, it is not the programmes that work but the resources they offer to enable their subjects to make them work. This process of how subjects interpret and act upon the intervention stratagem is known as the programme 'mechanism' and it is the pivot around which realist research revolves. (Pawson & Tilley 2004: 6)*

The monitoring plan that was developed in the GenderTime intervention programme adopted the main principles of the realistic approach and focused specifically on producing contextual knowledge about the varying intervention contexts where the interventions were implemented. The monitoring tools were also developed in order to particularly be able to capture the most relevant mechanisms in the implementation process and thereby contribute to the explanation and understanding of the varying outcome patterns in the different national contexts. Furthermore, in accordance with the realistic approach, monitoring in the GenderTime programme included data collection that was both qualitative and quantitative (Pawson & Tilley 2004). The monitoring plan was thus developed in order to support the implementation process by exploring possibilities, generate learning and ideas and trying them out, rather than trying to control the design of the implementation (Patton 2011).

### **The feminist approach**

The overall goals of the GenderTime intervention programme were related to women's rights and gender equality and the activities were designed in order to be consistent with feminist principles, so also the monitoring activities. This meant for example that feminist considerations guided the design of the monitoring activities and they were developed in line with appropriate sensitivity to women's authentic



experiences while being cautious and modest about goals and outcomes (Reid et al. 2006).

When developing a plan for monitoring in such an intervention programme it is helpful and important to take feminist evaluation principles into account. Feminist evaluation attempts to challenge and transform gender inequalities and power relations and expose discrimination by creating empowering and collaborative processes and participatory tools that facilitate reflection, awareness-raising and sense-making among stakeholders (Haylock & Miller 2016). Adopting this feminist approach to evaluation (and monitoring) means that the process of evaluation (and monitoring) is just as important as the findings and results of the evaluation (and monitoring) (cf. Scriven 2013). Characteristic of feminist evaluation is also that evaluation is considered a political activity that includes power analysis, focusing on gender power dynamics, power imbalances and power relationships (Frisby et al. 2009).

Kriszan and Lombardo (2013) emphasize the importance of considering both the process dimension and the content dimension simultaneously when evaluating the quality of gender equality change policies. According to them the process dimension includes two criteria of relevance when evaluating the quality of gender equality change policies. One of these criteria concerns empowerment through the inclusion of women in policy-making processes. The other criterion concerns incremental transformation of gender relations with reference to institutional and contextual legacies, meaning that transformation towards increased gender equality must be seen in relation to status quo (Kriszan & Lombardo 2013).

We have already previously argued the importance of shifting attention from (only) the outcomes of change projects to the process of change. The complexity of the GenderTime programme and the GEPs implemented in the different varying national contexts made the development of common indicators difficult, if not impossible. The actions in the seven GEPs implemented shared neither a common set of beneficiaries nor a single set of mechanisms for reaching them. This complexity was inherent to the integral and horizontal nature of GEPs. The GEPs were constructed in order to solve complicated, multi-dimensional and complex problems concerning gender inequality (Bustelo 2003).

However, it was also important to monitor the content of the gender equality change activities as this dimension is closely related to the process dimension. We for example expected the different success factors and hindering factors to be related to the specific character and content of the different actions. Thus, even if the impact of the GEPs was not in the focus for the monitoring activities, the goals and the visions of gender equality that underpinned the actions in the GEPs were of utmost importance to take into consideration in the monitoring activities (cf. Kriszan & Lombardo 2013). The content dimension was also heavily influenced by

feminist theories, because, as already explained, the plan for monitoring was developed in a theory-driven manner (cf. Funnell & Rogers 2011).

The following chapters contain further information about the monitoring approach in the GenderTime project and the empirical contextualisation. Details about the different monitoring tools that were developed and tested within the project will also be presented.

## **Empirical contextualisation**

Monitoring was an integral part of the GenderTime project and encompassed numerous activities. The most fundamental achievement was to develop a systematic and purposeful monitoring plan for the duration of the 4-year project. At the core of this plan was the 10 monitoring tools. Before each individual tool will be described in detail, the empirical context of these 10 monitoring tools will be outlined.

## **Monitoring as knowledge exchange**

Monitoring in the GenderTime project started with the realization that it would be an impossible task to develop monitoring tools for each of the almost 200 actions and interventions in the seven GEPs (Barnard et al 2013). Instead the main task of monitoring became to support the national teams with reflection tools for identifying facilitators and barriers for structural change within their organizations.

The monitoring work in the GenderTime project focussed more on accompanying the implementation processes within the institutions rather than on evaluating the outcomes. One particular emphasis laid on stimulating knowledge exchange and sharing experiences of the involved change agents. Therefore, the monitoring tools included in this handbook often compromised arranging arenas for discussions and interactive settings like workshops.

For us as monitoring agents it was important to offer room for such exchanges for the national GenderTime teams, especially since this is an aspect which easily gets neglected in cross-cultural projects. Working in this kind of collaborative settings brings up several challenges. Time for face-to-face exchange and discussions is rare since the bi-annual project meetings, that are standard in these type of projects, mostly have a duration of only two full days, where all current tasks for each work package need to be considered. The overall aim of this type of monitoring, focussing on knowledge exchange, was, to gain a better understanding of how the implementation processes of the GEPs worked, by identifying challenges and success factors on both individual and organisational levels. The collected data retrieved out of the monitoring tools was fed back to the participants on a regular basis to support the individual as well as the group reflections.

## **Monitoring agents, multiple project roles and monitoring budget**

As mentioned in the beginning of this Handbook, one of the two leaders of the monitoring work package, was also the leader of a national GenderTime team that implemented a GEP. In the role as monitoring agents we both, however, primarily defined ourselves as objective, but supportive, observers of the implementation processes. The success of the GenderTime project and implementation of GEPs did not depend on the monitoring outcomes or us, and the results of the monitoring

activities were not connected to potential budget cuts in the case of not meeting previous set goals. These are two important framework conditions, which ensured a certain neutrality of the project members responsible for monitoring, i.e. the monitoring agents, even though one of them had a double-role through her direct involvement in the implementation process.

In fact, most of the members of the seven national GenderTime teams, to a certain extent, took on the role as internal monitoring agents during limited periods of the project duration. The reason for this was the necessity to involve them in producing the monitoring data. The monitoring tasks that each national team performed involved; participation in workshops arranged as part of the monitoring activities; collecting and reporting monitoring data, and; documenting and sharing own experiences and knowledge (see the descriptions of the monitoring tools in this handbook). Each national team was allocated at least 8 person months (PMs) during a course of four years to perform these monitoring tasks.<sup>3</sup>

The two leaders of the work package “Monitoring” devoted 16 PMs each to monitoring tasks during the length of the 48-months project. Their tasks comprised developing the monitoring plan; constructing the monitoring tools; implementing a majority of the monitoring tools; initiating and overseeing the implementation of the remainder of the tools; documenting the implementation of the monitoring tools, collecting monitoring data; analysing the data and the results; and finally, feeding back the results to the rest of the GenderTime consortium.

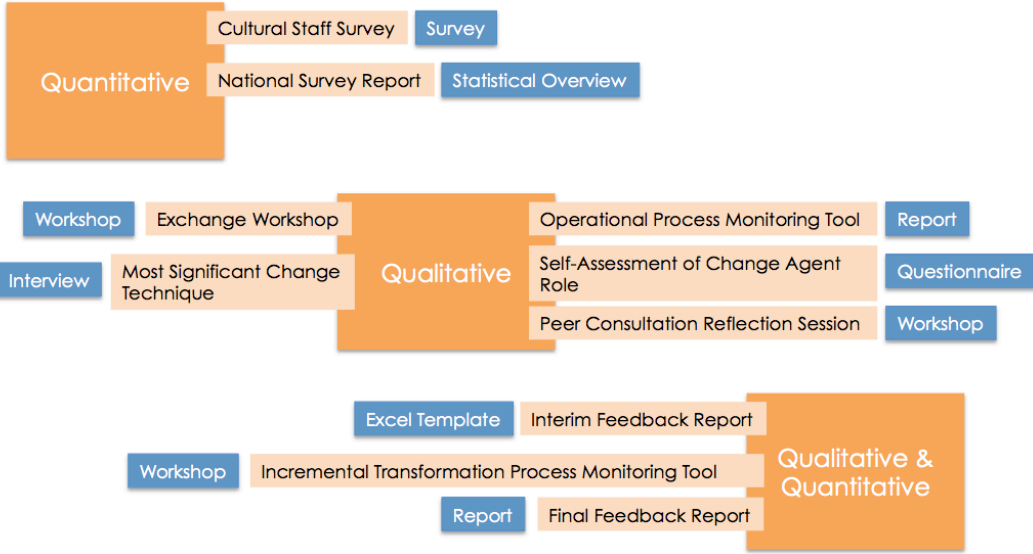
### **Qualitative and quantitative monitoring tools**

Figure 1, below, illustrates that the monitoring tools were fairly balanced along the quantitative/qualitative spectrum although there was a predominance of qualitative tools. The decision to give precedence to qualitative monitoring tools was based on concerns about avoiding a simplistic view of gender equality as only dealing with numbers. Drawing on previous studies, feminist theories and expert interviews, gender equality was operationalized in qualitative terms as involving culture, structures and attitudes and therefore difficult to quantify. However, quantitative measures and indicators are an important complement to qualitative tools and some of the monitoring tools offered the opportunity to combine the qualitative methods with quantitative ones, adopting a multi-method approach.

---

<sup>3</sup> Each national GenderTime team consisted of between 2 and 5 individual members.

Figure 1: Qualitative and quantitative monitoring tools



**Distribution of the monitoring tools during project phases**

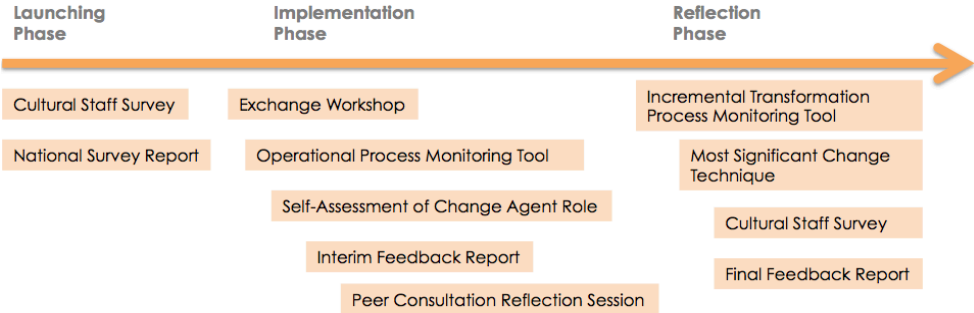
Another important characteristic of the monitoring programme in the GenderTime project concerned the timing of the different monitoring tools, i.e. how they were distributed during the different phases of the project. Figure 2 below illustrates this property of the programme and shows that monitoring activities played an important role already in the launching phase of the project, i.e. month 1 to 6. During this phase, baselines were established and the aim of the monitoring tools was thus to collect data about the status quo in each of the seven institutions where the GEPs were implemented.

Two to three monitoring tools were used each year during the 4-year project with the exception of year 3 in the project when 4 different monitoring tools were implemented. The reason for this was that during year 3 all interventions were in progress, had been initiated or had already been terminated, which called for an intensification of the monitoring activities.

Decisions regarding the distribution of the tools over the 4 years also took into account that not too many monitoring tools could be implemented during each phase, since the staff resources allocated to monitoring activities, although generous, was limited to around 8 PMs for each national GenderTime team. Each monitoring tool required the involvement of the implementing GenderTime members in some ways and it was acknowledged that their workload was already high.

The wish to create a plan for systematic and purposeful monitoring also entailed a distribution of different types of tools during the different phases – hence for example alternating quantitative and qualitative tools. This facilitated the collection of rich data that could produce a “thick description” of the implementation process.

Figure 2: Timing of the different monitoring tools



The tools are presented in detail below in the order they were implemented in the GenderTime project. The presentation thereby illustrates a form of progression in the monitoring process, where the subsequent monitoring tools built on the previous ones.

## **Presentation of the GenderTime monitoring tools**

### **1. Cultural Staff Survey**

#### ***Aim of the tool***

*The Cultural Staff Survey* was the first monitoring tool developed and applied in the project. The aim of this monitoring tool was to provide the seven national GenderTime teams with a quantitative indicator they could use to measure the level of gender equality at the participating institutions at the start of the project, i.e. to establish a baseline for the implementation of the GEPs.

#### ***Background of the tool***

The Cultural Staff Survey was based on a Culture Analysis Tool<sup>4</sup> developed by UKRC-WISE, as part of the HEFCE funded national HE STEM programme with the aim of sharing good practice in gender equality in higher education. The aim of this questionnaire is to help university departments understand how male and female staff experience their working environment and what, if any, improvements may be needed to ensure equality of opportunity. Content wise it focuses on four areas:

1. Participation and promotion practices
2. Workplace culture
3. Leadership and management commitment
4. Institutional reputation and social responsibility

#### ***Structure of the tool***

The monitoring agents immediately began developing this monitoring tool during the project's launching phase. Six months into the project the national teams were involved in tailor making the survey to fit their specific national and cultural context but also to fit the intended target groups of the survey (e.g only faculty members or including also administrative staff). The survey was tailor made in this way, as the questions were adapted to meet the specific context of each institution also with regard to already existing gender equality activities of the organization.

---

<sup>4</sup> See: <https://www.wisecampaign.org.uk/resources/2010/06/staff-culture-analysis-survey> for more information. [22.2.2018]

This inclusion of specific institutional questions allowed for a broader efficiency assessment of the implemented GEPs.

After translation into national languages the seven different versions of the survey were set up on the project webpage as an online version, which allowed a better dissemination of the questionnaire within the institutions/organizations (Barnard et al. 2014).

Figure 3 on the next page shows a screenshot of the first page of the staff survey implemented at Loughborough University. It illustrates for example that answering categories of the closed questions were organized along a 5-point-Likert-Scale plus the category 'Not applicable/I don't know'.

The questionnaire further contained sections for remarks where participants were free to add information if necessary.

### ***Results of the tool***

The national teams were responsible for organizing the data collection at their own institution. The strategies used for promoting the participation in the survey were similar: emails to all targeted employees, announcements in internal newsletters, or e.g. via the company internal web platform. Partly the commitment from senior management was used to motivate staff members to participate and in some cases 'survey ambassadors' were recruited, who informed colleagues about the aim and importance of the study. Some of the national GenderTime teams reported that they had 'neutralized' the survey announcement due to strategic reasons. Deleting words like 'gender equality' was believed to increase the participation also of male staff members and support their inclusion. The reason for this was that 'gender equality' was thought to be interpreted, in some national and cultural contexts, as a subject only concerning women.

A special web application was developed, which allowed all national GenderTime teams to retrieve the absolute gender-segregated data of their institutional survey plus the overall cross-national data. The latter one also was made available to allow for women and men only comparison between the countries. Preliminary results were presented and briefly discussed by each national team during a project meeting. For a better understanding of the results more information of the national context was necessary, which led to the development of a national report template (see description of the next tool).

The implementation and analysis of this survey symbolized a benchmarking for the national teams about the work culture in their institutions. It is a useful tool for getting information on staff member's perceptions of working conditions,



management structures, communication flows and organisational environment not only with respect to gender equality matters. For achieving gender equal workplace conditions, it is important to focus on the prevalent organisational structure itself (Castaño et al. 2010). The results further showed that the implementation of gender equality measures or GEPs can be even more difficult if actors in an organisation are not well informed about the legal, national and organisational context (Achterberg & Dahmen 2017).

A tool like the Cultural Staff Survey offers the possibility for an on-going reflection of the implementation process on a structural and individual level, which is important since organisational culture is not a static concept, it is fluent, and therefore it is necessary to explore the beliefs and behaviours within institutions (Peterson 2010). For these reasons, it can be recommended to implement staff surveys periodically or at least at the beginning and the end of an intervention for comparing the effects and results.<sup>5</sup>

*Figure 3: Screenshot of the first survey page at Loughborough University*

Loughborough University School of Civil and Building Engineering Staff Survey

View Results

**PARTICIPATION AND PROMOTION PRACTICES**

*This section assesses whether your Department has policies and practices that encourage the participation and promotion of women and men at all levels. Please indicate your level of agreement or disagreement with each statement.*

**1. In my School, staff are treated on their merits irrespective of their gender (e.g. both women and men are actively encouraged to apply for promotion and take up training opportunities).**

Strongly disagree

Disagree

neither agree nor disagree

Agree

Strongly agree

Not applicable or don't know

**2. In my School, work is allocated on a clear and fair basis irrespective of gender**

Strongly disagree

Disagree

Neither agree nor disagree

Agree

Strongly agree

Not applicable or don't know

---

<sup>5</sup> The second implementation of the Cultural Staff Survey in GenderTime happened at the very end of the project and was an optional monitoring tool. The analysis of these outcomes was not reported back to the monitoring agents and is therefore not included in this handbook.

## **2. National Survey Report**

### ***Aim of the tool***

The aim of the monitoring tool called the *National Survey Report* was to produce contextual data that would support and facilitate a cross-national analysis of the results of the *Cultural Staff Survey*.

### ***Background of the tool***

This tool filled an important function in order to develop a shared understanding of the national, organisational, cultural and social context in which each of the seven GEPs were implemented. Without this context, it would not only be difficult to interpret the results of the survey, but it would be difficult to monitor the implementation of the GEPs. This context was also essential for the dialogue-based knowledge exchange between the national teams, which was a central feature of the GenderTime project. Without it misunderstandings and miscommunication could easily create problems.

### ***Structure of the tool***

The National Survey Report was designed as a Word-template and consisted of two parts. Each part included questions and text boxes where the requested data could be inserted by the national teams. They were encouraged to write as much as they wanted and give as much detail as they could as a reply to each question. The space where they were supposed to insert text as a reply could easily be extended, without word limit. This layout of the tool was coherent with the other templates that were used for monitoring within GenderTime. This coherence was considered beneficial because it would evoke recognition value for the national teams.

The first part of the National Survey Report focussed general information about the Cultural Staff Survey (see figure 4 below for an example). The national teams were here invited to add information about the national and organisational context and about the implementation of the survey at their institution; e.g. if they had received any feedback from staff members.

In the second part of the National Survey Report the national teams were asked to add their own interpretation of the data and the results of the Cultural Staff Survey regarding some specific questions, which were particularly difficult to interpret for persons who might not have the organizational background knowledge (in this case this meant the monitoring agents and other national teams). In addition, if applicable, any interesting responses in open answer categories or remarks sections for single questions in the National Staff Survey were supposed to be translated and included in the National Survey Report. This was done by using a survey ID

for the respondents, which included their gender, so that it would be possible to refer to them/cite them.<sup>6</sup>

**Results of the tool**

The National Survey Report provided the possibility to compare the survey outcomes cross-nationally. By doing so, cross-national similarities and differences of the totally 1053 respondents were revealed, with partly surprising findings. For instance, male staff in Austria and Germany were more likely to agree that it was clear for them what their department/faculty policies looked like in relation to gender equality compared to their British and Spanish counterparts. And generally, both men and women of the total sample agreed with around 50% that they understood why the institution/faculty needed to take up actions on gender equality. Results like these offered indications for the further implementation as well as monitoring process.

Figure 4: First page of the National Survey Report template

General information	
Please insert the short name of your institution here:	
How did you spread the information within your organization about the survey?	Please insert text here
Please specify the target population of the staff survey in your country:	<input type="checkbox"/> All staff members (academic and non-academic) <input type="checkbox"/> Academic staff members <input type="checkbox"/> Non-academic staff members
	<input type="checkbox"/> Other, please specify: <input type="checkbox"/> Whole university / institute <input type="checkbox"/> Single faculty/faculties (please specify): Please insert text here
Total sample number:	Please insert text here
Response rate in %	Please insert text here
Did you receive any feedback (positive/negative) or remarks from the participants concerning the survey?	Please insert text here

---

<sup>6</sup> The reporting of the results of the survey was based on gender-segregated data. Initially the idea was to use a more inclusive approach by not using the binary gender categories ‘male’/‘men’ and ‘female’/‘woman’ in the survey. The option ‘other’ would leave room for those who do not identify with one of the traditional categories. It was, however, not possible to reach agreement between all national teams about this.

### **3. Exchange Workshop**

#### ***Aim of the tool***

The aim of the *Exchange Workshop Monitoring Tool* was to allow all seven national GenderTime teams to work together on the implementation of the GEPs. This collaborative effort was achieved through facilitating a collective reflection and an exchange of experiences between the national GenderTime teams focusing on a discussion of successes and challenges.

#### ***Background of the tool***

Just as a majority of the monitoring tools, the Exchange Workshop was based on research-based theories about organizational change and feminist theories about challenges and resistance related to gender equality interventions. This deductive framework emphasized the need to understand change processes, problems and effective practices. The need for this was also emphasized in the interviews with experienced evaluators and other experts, practitioners and change agents within the relevant field in question that were performed during the initial phase of developing the plan for monitoring in GenderTime.

#### ***Structure of the tool***

The Exchange Workshop Monitoring Tool was structured as an internal workshop concept and implemented during a GenderTime project meeting. The workshop was part of the in-process consultation model of the implementation of the GEPs (cf. Arrizabalaga et al. 2014). Prior to the workshop participants had to prepare a list of three challenges and three successes related to the process of implementing GEPs at their organizations. These lists were the entry points for the discussions at the tables during the workshop.

The first part of the workshop was run in World Café format. World Café methodology is a simple, effective, and flexible format for hosting large group dialogue (DG 2011). This format gave the national teams the chance to share their experiences of implementing GEPs at their organizations. In the second part of the workshop two national teams were paired to work on a practical 6-month action plan of how they could support each other. The lists of challenges and success constituted the starting point for this phase. The action plans consisted of a prioritized list of challenges to be addressed by the national teams together with a strategy for how the other national team, that was assigned as a collaborative partner, could assist in addressing them.

### ***Results of the tool***

The results of the Exchange Workshop showed that persons involved in change processes tend to focus on “challenges” rather than on “success” factors (see second photo below). This is especially crucial since theories about organizational change have emphasized the importance of acknowledging also short-term wins right from the beginning of the implementation process (Kotter 1995). It is not only essential for keeping up the enthusiasm and positivity of the change agents but also for communicating success in the organization and give legitimacy. A monitoring tool was therefore later developed in order to acknowledge short-term wins. The challenges identified by this tool were related to “communication” (how to explain understanding of project), “precarity” (recruitment, promotion mechanism), “balance” (work life balance, gender balance), “data” (lack of disaggregated data), “environment” (scepticism, competitive culture) and “institutionalize” (decision-making issues, commitment of management).

*Figure 5: Photos from the Exchange Workshop. First photo: “World Café” setting. Second photo: Post-Its documenting successes and challenges*



#### **4. Operational Process Monitoring Tool**

##### ***Aim of the tool***

The *Operational Process Monitoring Tool* provided a detailed overview of the progress towards the objectives in the implementation of GEPs, during the early implementation phase. The purpose of the tool was to identify important success factors and hindering factors, related to the specific objectives in each institution.

##### ***Background of the tool***

This monitoring tool, with the focus on success factors and hindering factors, was developed primarily using a deductive approach, drawing on previous action research on feminist interventionist change projects. Obviously, due to the purpose of the interventions in the change project, feminist theories became important for the abductive approach used for developing the monitoring program in the GenderTime project.

Awareness of organizational power structures is necessary in projects aiming to change organizations (Haylock & Miller 2016). Gender equality projects can undermine power structures and reduce management control, which most probably will appear as undesirable by powerful organizational actors. It is for example important to identify the positions in which actors have the power to achieve the changes that are sought in the project, often line positions rather than staff positions (Coleman & Ripping 2000). Based on this previous research the tool for example included a question about collaborations within the participating institutions, with the aim to identify actors and internal partners important to negotiate with and win the trust from in order to succeed with the implementation of GEPs.

Special attention was also paid to gathering information about whether some actions in the GEPs had been cancelled, and why that was so. This was viewed as important information to be provided in relation to hindering factors and barriers.

##### ***Structure of the tool***

This tool was designed as a template to collect information from all the national GenderTime teams. Compared to several of the other monitoring tools, aimed at facilitating reflection, filling out this template was intended to be an internal team effort, completed by each GenderTime national team, not the individual team member. The Operational Process Monitoring Tool thus provided the national GenderTime teams time to discuss and reflect together as a team on different experiences, perceptions and opinions regarding the implementation process.

The template requested information about the implementation process that was structured around the objectives stated in the GEPs, rather than the specific actions

in each institution. The tables in the templates were therefore linked to every objective of the tailor-made GEPs and the monitoring tool was thus similarly tailor-made for each national team (for examples of the objectives in GenderTime see figure 6 below).

This also meant that the tool provided limited information about the specific actions in the GEPs. The template, however, allowed for the teams to include information about specific actions - in progress or completed. The teams could decide to write generally about the objective or specifically about only one or two actions. The responses show that both ways of reporting were used. As with the other monitoring tools, the national teams were encouraged to write as much as possible and share as many experiences and as much detail as possible – without word limit in the template. The templates were distributed via email during month 22 of the project.

Figure 6 below illustrates the design of the template for this monitoring tool. It displays one of the pages in the template document where the teams were supposed to fill out objective-related monitoring information.

### ***Results of the tool***

The seven national GenderTime teams reported on a wide range of success factors that they identified had supported the pursuit of the objectives, for example good collaborations inside the institutions and management commitment. Similarly, they reported on a wide range of hindering factors, for example existing informal arrangements and lack of gender awareness. The tool also provided examples of how challenges in the implementation process, related to lack of time, staff and funding, had been overcome by the use of different strategies. It also showed how some actions in the GEPs had to be abandoned, due to challenges, and instead replaced with other ones.

The results from the tool proved to be highly relevant as it allowed for a deeper understanding of the progress related to each objective – which was important at this early stage of the implementation process. However, when developing the next monitoring tool the focus was again on specific interventions in order to retrieve more detailed information on the activities in each institution.

Figure 6: One page of the template for the Operational Process Monitoring Tool

Objective 1		
1) Please specify your institutional objective(s):	<input type="checkbox"/> Improve employment opportunities for all at institutional level	<input type="checkbox"/> Review and monitor salaries
	<input type="checkbox"/> Review and monitor recruitment processes	<input type="checkbox"/> Create gender balanced decision bodies
	<input type="checkbox"/> Diagnosis on the current situation regarding gender	<input type="checkbox"/> Work-life-balance
		<input type="checkbox"/> Not applicable for my institution
2) Please identify the main actors for the implementation	<input type="checkbox"/> GenderTime members	<input type="checkbox"/> Other, please specify:
	<input type="checkbox"/> Transfer Agents	
3) How important is this objective for your institution compared to the other objectives?	<input type="checkbox"/> Very high	<input type="checkbox"/> Medium
	<input type="checkbox"/> High	<input type="checkbox"/> Low
		<input type="checkbox"/> Very low
		<input type="checkbox"/> Can't decide
4) Please describe the institutional relevance of this objective (e.g. aims, hopes, background, reasoning, outcome) with regard to achieving structural change:		
Insert text here		
5) Please describe, which factors support you in achieving this objective (e.g. managerial, individual, institutional etc.):		
Insert text here		
6) Please describe, which factors hinder you in achieving this objective (e.g. managerial, individual, institutional etc.):		
Insert text here		
6 a) If you already have met barriers in achieving this objective, how have you handled these? What were the consequences (e.g. change of measure(s), aims, persons involved etc.)? Please describe this process:		
Insert text here		
6 b) If you already have met barriers in achieving this objective but did not overcome these yet, what could help you? Please describe below:		
Insert text here		
7) How would you assess the success in achieving this objective until now?	<input type="checkbox"/> 1 - excellent	<input type="checkbox"/> 3 - average
	<input type="checkbox"/> 2 - good	<input type="checkbox"/> 4 - satisfactory
		<input type="checkbox"/> 5 - poor
		<input type="checkbox"/> Can't decide
8) Any further remarks you want to add related to this objective?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
	If yes, please specify: Insert text here	
9) Please insert names of team members involved in finalizing this template here:		
Insert text here		



## **5. Self-assessment of Change Agent Role**

### ***Aim of the tool***

The aim of the *Self-Assessment of Change Agent Role Monitoring Tool* was to document the factors that motivate academics and practitioners to engage in activities aimed at transforming research institutions into gender equal organizations.

### ***Background of the tool***

The background of the Self-Assessment of Change Agent Role Monitoring Tool was the considerations expressed in the expert interviews that were performed during the initial phase of developing the plan for monitoring in the GenderTime project. These gender equality experts emphasized the importance of also monitoring the people who actively implement GEPs. This was an important piece of information about how to support organizational change processes that seemed to be missing from the research literature on monitoring and evaluation.

The research literature, however, provided two other important pieces of information that were used as starting points when developing this monitoring tool. One such piece of information was found in the evaluation literature that emphasized the importance of allowing for reflection during the implementation phase (Coleman & Rippling 2000). Another piece of information was found in previous literature on the importance of change agents for changing organizations towards increased gender equality (Parsons & Priola 2013).

This tool, like many others, also drew on feminist organization theories. These theories provided a critical perspective on how organizational processes, practices and structures are gendered. Gendered organizational structures and practices are notoriously rigid and resistant to change (Linstead et al. 2005). Change agents for gender equality therefore often face resistance. Feminist theories provided us with a framework for understanding the role of active and passive resistance to gender equality and to feminist activism.

The Self-Assessment of Change Agent Role Monitoring Tool was thus an example of a tool developed using a deductive approach, meaning that it was inspired by research-based theories about organizational change and gender relations (cf. Pawson & Tilley 1997) and by interviews with experienced evaluators, experts, practitioners and change agents (cf. Sridharan & Nakaima 2011).

### ***Structure of the tool***

The Self-Assessment of Change Agent Role was a monitoring tool with a qualitative character that involved individual self-reflection by every member of all

seven national GenderTime team. These team members, who were active in the implementation of the GEPs, were asked to describe their personal experiences as Change Agents for gender equality. The tool was implemented during month 24 of the project, i.e. exactly at “half-time” of the project, which is a critical phase when it is necessary to identify obstacles and provide support for corrections.

The tool provided a personal and individual perspective on success factors and challenges, especially resistance. In this way, this tool produced data that complemented the information collected with the help of the previous monitoring tool – the Operational Process Monitoring Tool, which collected information on an institutional and organizational level.

The tool collected individual memories of situations in which the Change Agents had experienced negative and/or positive reactions to their attempts to improve gender equality in an institutional context. It was a tool that identified patterns of privilege and disadvantage and thus recognized the existence of a gender order, meaning that the notions and actions of individuals both produced and were the result of gendered power relations.

The tool was constructed as a questionnaire and as part of the framework of different, but coherently designed, templates developed for monitoring within the GenderTime project. The template consisted of six questions (see figure 7 below), considered to be the most relevant for the purpose of the tool – i.e. to collect data about the experiences of change agents for gender equality. Each team member was encouraged to write down as many reflections as possible. Most of the members of the national GenderTime teams filled out the questionnaire template and submitted their personal reflections as part of a monitoring session during a project meeting. In order to include also the reflections of team members who could not participate in the project meeting, the template was also distributed via email. Some of the team members accordingly submitted their reflections electronically by returning the filled-out questionnaire via email.

### ***Results of the tool***

The Self-Assessment of Change Agent Role Monitoring Tool collected information about different kinds of resistance and how such resistance can be overcome. The team members described a wide range of supportive factors that strengthened them in their role as change agents as well as challenges that hindered them. These are personal stories about possessing or lacking different types of resources that are important for a change agent to be efficient.

All change agents referred to their role within the organization as an essential resource for the outcomes of the implementation process. Different variations of

formal as well as informal roles were described as important resources. Being a change agent can often be an informal role within an organization and that role can be strengthened by for example a formal role such as being a professor. Having a formal role within the human resource department was also described as an advantage as it provides a platform to build a network within the organization. A position such as that can also open doors and lead to increased access to information. On the contrary, having a precarious position in the organization was deemed as utterly detrimental for a change agent.

The problem with lack of resources such as time and money was also mentioned, as well as gaining legitimacy of the gender change project and support for it. In addition, the tool produced important information about the motivation of the change agents (sense of fairness, awareness of problems, personal experiences of discrimination etc.) and about requests for supporting factors (more extended networks, less precarious positions, increased resources etc.). Finally, the change agents also submitted useful and important recommendations for efficient change agency, based on their own experiences: search for allies; build a team; know the organization; be persistent; rely on facts; piggyback on existing initiatives etc.

*Figure 7: Questions of the Self-Assessment of Change Role tool*

- 1) Please describe factors that have strengthened and/or supported your role as a Change Agent for gender equality. It can be individual (position, knowledge, networks etc.) or structural (management commitment, policies etc.) factors.
- 2) Please share some factors that have hindered and/or challenged you in our role as a Change Agent for gender equality. It can be individual (position, knowledge, networks etc.) or structural (management commitment, policies etc.) factors.
- 3) What motivates you as Change Agent for gender equality? Please share your personal motivation.
- 4) What can we learn from other Change Agents in our institutions/other institutions? What makes them efficient and successful? It can be individual (position, knowledge, networks etc.) or structural (management commitment, policies etc.) factors.
- 5) What could strengthen you in your roles as Change Agents in your institutions? It can be individual (position, knowledge, networks etc.) or structural (management commitment, policies etc.) factors.
- 6) Recommendations for efficient Change Agency: What do efficient Change Agents need and/or do? It can be individual (position, knowledge, networks etc.) or structural (management commitment, policies etc.) factors.

## 6. Interim Feedback Report

### *Aim of the tool*

The overall aim of the *Interim Feedback Report* was to let the national GenderTime teams report back on success factors and challenges, linked to each specific intervention activity in the GEPs. The purpose was to receive provisional information that would constitute an important base for the decisions regarding which of these actions to monitor more closely.

### *Background of the tool*

A similar tool, documenting the progress of the implementation of every activity in the seven GEPs, was used at the beginning of the project. That version of the tool, however valuable it was as it systematized the collection of data, only comprised implementation related information (and is not included in this handbook). The development of the Interim Feedback Report as a monitoring tool meant an extended and more comprehensive reporting on the activities. The report therefore not only included quantitative information but also qualitative data.

### *Structure of the tool*

The Interim Feedback Report was developed as an Excel document, which has limitations when it comes to the amount of information that is possible to gather. The information provided by each national GenderTime team was thus necessarily very brief. However, as this was an interim feedback report, only provisional information was requested and an Excel sheet was therefore appropriate.

The report consisted of ten columns where the national GenderTime teams would update already previously provided information about each and every action in the GEPs. The data to be inserted was divided into implementation related (six columns) and monitoring related information (four columns) (see figure 8).

*Figure 8: Information collected in the Interim Feedback Report*

Implementation related information	Monitoring related information
<ul style="list-style-type: none"><li>• Action</li><li>• Objective</li><li>• Stage indicated in previous report</li><li>• Qualitative information provided in previous report</li><li>• Update on qualitative information</li><li>• Stage (completed/in progress/cancelled)</li></ul>	<ul style="list-style-type: none"><li>• Deviations in relation to the initial GEP</li><li>• Success/support factors to report?</li><li>• Challenges/resistance to report?</li><li>• Short description of success factors/ challenges</li></ul>

### ***Results of the tool***

The Interim Feedback Report provided evidence of the progress of the implementation processes at the involved institutions by depicting in a comprehensible way the advancement of the GEPs. The tables helped to identify delays, postponements, resistance, challenges and hindering factors. It was also important for identifying support and success stories.

For actions that the national teams indicated was still progressing according to the plan, they were asked if they would categorise this as an example of a success story or not. The actions, described as successes, would be studied more in detail in order to investigate the supporting factors and understand why these actions had become success stories.

The actions that according to the teams were not proceeding like planned would be further investigated in order to possibly identify and understand challenges and hindering factors.

The information, gathered by this tool, was used for the development of the Peer Consultation Reflection Session workshop concept, which is described below.

## **7. Peer Consultation Reflection Session**

### ***Aim of the tool***

The aim of the *Peer Consultation Reflection Session Monitoring Tool* was to create an arena where participants in the project could listen, reflect, share ideas and solve problems together and learn from each other's experiences of the implementation progress across the national teams.

### ***Background of the tool***

This tool was, just as the Self-Assessment of Change Agent Role Monitoring Tool, based on considerations about reflection as an important monitoring methodology (Coleman & Ripping 2000). However, instead of personal reflections this tool introduced the concept of peer consultation.

The development of the tool was inspired by previous literature on peer consultation reflection exercises (Brown et al. 1999). Peer consultation reflection exercise is an innovative strategy for personal and professional development that can be used in many different settings - creating a valuable opportunity to interact and learn from colleagues. It is a method that encourages and helps people to see their situation from a different perspective and to understand problems in a different way. When peer consultation reflection sessions are working well they promote equality among participants and generate new ideas and solutions for both presenters and participants.

The tool therefore had an inductive quality, developed as a result of continuous observations of the activities in GenderTime (cf. UNDP 2009). These types of arenas for knowledge sharing was something that had been requested within the project and something that had previously been appreciated.

### ***Structure of the tool***

The Peer Consultation Reflection Session Monitoring tool was developed as one of several workshop concepts within the GenderTime monitoring plan. The tool was implemented during the project meeting in month 30 – more than halfway through the project – when most interventions had commenced and there was a need for support and sharing of knowledge and experiences. The participants in the project meeting were divided into four different peer reflection groups to discuss four different themes. Each peer reflection group consisted of representatives from different national GenderTime teams. After 30 minutes discussion in these four groups, four new groups were formed, discussing four new themes.

Besides the discussion themes, the participants were also provided with instructions for *how* to discuss the themes. The participants were expected to

switch between taking on the role as a Presenter, i.e. a person that describes and reflects over a situation, action, challenge or good practice, and the role as Reflection facilitators. Reflection facilitators were expected to pose questions to the Presenter to further clarify the challenges or to further increase the understanding of the success factors. The participants in the groups thus formed a so-called peer reflection team that not only facilitated the reflections of the Presenter but also contributed with their own reflections on how to, for example, overcome challenges or how to best take advantage of success factors.

The Peer Consultation Reflection Session Monitoring Tool was tailor-made because the eight themes that were discussed in the peer reflection groups were based on the information that the teams had provided in the Interim Feedback Report (see previous tool description in this handbook).

In order to ensure that the reflections would be shared in a helpful manner the themes were constructed as so-called “how to”- themes. This was also believed to stimulate the sharing of good advice and solutions to challenges. The reflection experiences were supposed to focus on successes and challenges within the specific “How to”-theme.

The first round of reflection sessions was organized around the following four themes:

- How to assess equality in complex organisations (employment/salary)?
- How to understand the gendered career through interviews and focus groups?
- How to raise awareness within the organization?
- How to support women through mentoring?

The second round of reflection session was organized around the following four themes:

- How to monitor through gender sensitive indicators?
- How to identify career obstacles through exit interviews?
- How to disseminate good practices outside the organization?
- How to challenge male dominance through women’s networks?

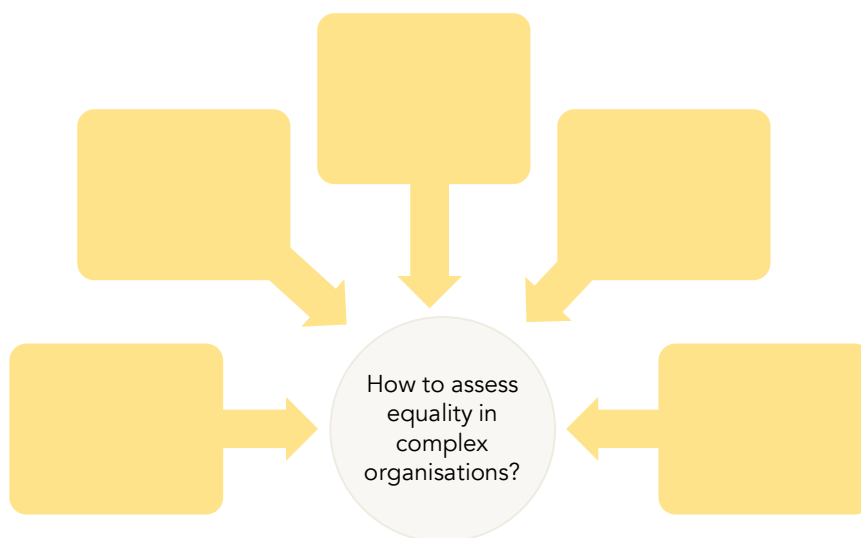
After 60 minutes, everybody had participated in two Peer Consultation Reflection Sessions and a short “check-out” phase commenced, where one rapporteur in each of the eight teams reported back their main key insights that had surfaced during the reflection sessions. This “check-out” round revealed that the workshop concept had been successful as to providing an arena for sharing and reflection over a wide range of relevant and important themes. During the sessions, there had also been a note taker in each of the eight teams and the notes were made on so called “reflection sheets” (see figure 9 below for an example of such reflection sheet).

### ***Results of the tool***

This tool provided interesting results for all eight different themes. The challenges that were discussed concerned for example problems with developing indicators to make inequalities visible and lack of resources to collect data. The solutions involved suggestions about how already existing databases can be used in collaboration with HR departments.

Another important discussion concerned the role of women's network. Here, some of the challenges identified concerned problems with involving women in networking activities and how to make women's networks powerful. Experiences of how men resisted women's networks were also shared as well as suggestions about how to create empowering face-to-face mentoring situations.

*Figure 9: Example of a "How to"-theme reflection sheet*





## **8. Incremental Transformation Monitoring Tool**

### ***Aim of the tool***

The *Incremental Transformation Monitoring Tool* aimed at defining impact relatively and contextually in relation to the starting point in each participating institution. It was the first tool to be implemented during the final and third phase, the reflection phase, and it was therefore designed to allow for reflection on the two previous phases. As such it focused on facilitating an internal discussion within the GenderTime consortium about setting and achieving intermediate goals in a change project.

### ***Background of the tool***

The Incremental Transformation Monitoring Tool was developed using an abductive approach, i.e. by combining induction with deduction. The deductive approach, drawing on a comprehensive theoretical framework, was initially inspired by recommendations within evaluation theory, feminist theory and policy research, emphasizing the importance of including process criteria in evaluations of gender equality policies. According to for example Krizsan and Lombardo (2013) evaluation involves linking such criteria to “incremental change in relation to institutional and contextual legacies” (Krizsan & Lombardo 2013: 86). Focusing incremental change in specific contexts means among other things to acknowledge that the success of gender equality actions is dependent on the status quo compared to which these actions are adopted.

The Incremental Transformation Monitoring Tool was also influenced by a theoretical model, developed by John P. Kotter (1995). Kotter’s model describes an 8-step process of successful organizational change. This model constituted the foundation for the development of the tool. The eight different steps of the model are as follows: 1) Creating a sense of urgency about the changes needed. 2) Building a coalition within the organization. 3) Forming a strategic vision and initiatives about the changes. 4) Enlisting volunteers that are committed to the change. 5) Enabling action by removing barriers. 6) Generating short term wins. 7) Sustaining acceleration. 8) Instituting change.

These 8 steps are divided into three phases where steps 1 to 3 concern creating a climate for change. Steps 4 to 6 concern engaging and enabling the organization for change. Finally, steps 7 to 8 concern “making the change stick”.

The abductive approach used when developing this monitoring tool proved helpful when it became clear that all phases and several of the steps described in Kotter’s model had already been in focus on several of the previous monitoring tools. However, the analyses also identified that some of these steps had not received

enough attention. These steps were deemed highly important and therefore selected for the tool.

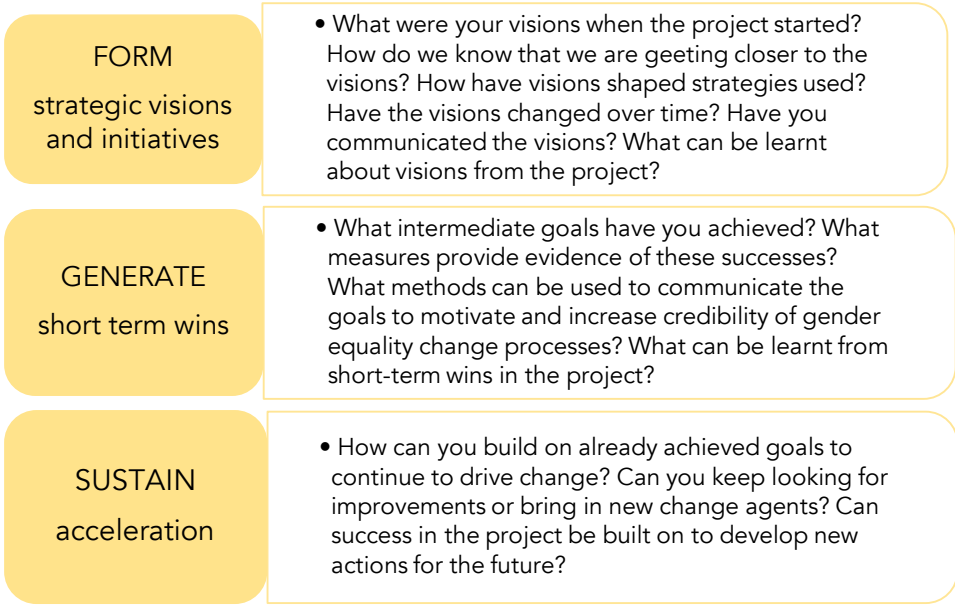
**Structure of the tool**

The Incremental Transformation Process Monitoring Tool was, similar to several other tools, developed as a workshop concept, and implemented during a project meeting. The workshop started with a theoretical introduction of the tool and of the aim of using it. After this introduction, the 15 workshop participants, i.e. members of the national GenderTime teams who implemented GEPs, were divided into three groups for 45 minutes intense discussion and knowledge sharing. The discussion was organized in the three following different topics:

- 1) How to form a strategic vision
- 2) How to generate short term wins
- 3) How to sustain acceleration of change processes

Every topic was elaborated into a number of questions that would facilitate the discussion on the topic (see figure 10 below).

*Figure 10: Topic-related questions of the Incremental Transformation Process Monitoring Tool*



Each group documented their discussion in specific tailor-made templates (see figures 11-13 below), which were collected afterwards by the workshop leaders. The outcomes of the small group discussions were reported and discussed within the whole team.

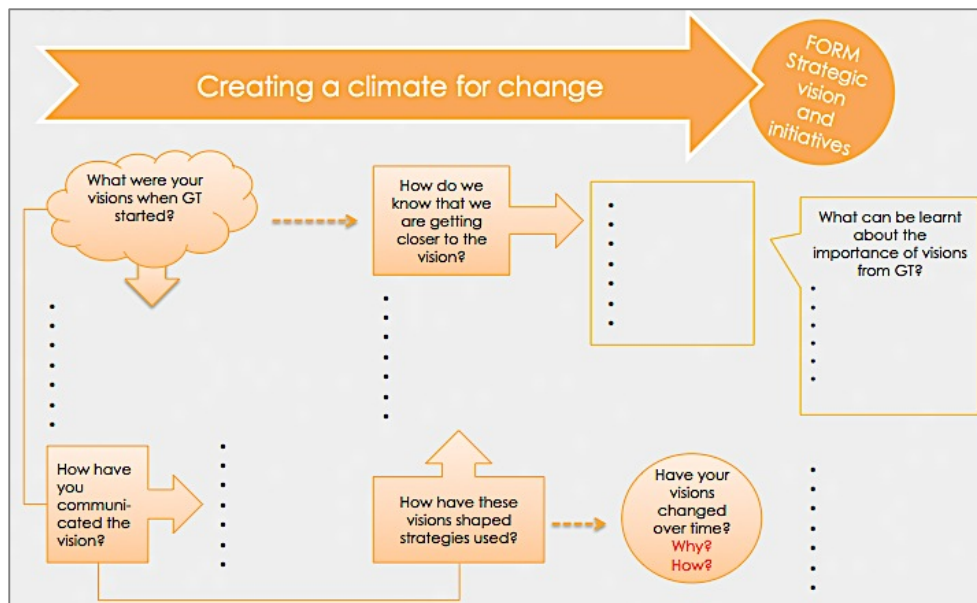
### ***Results of the tool***

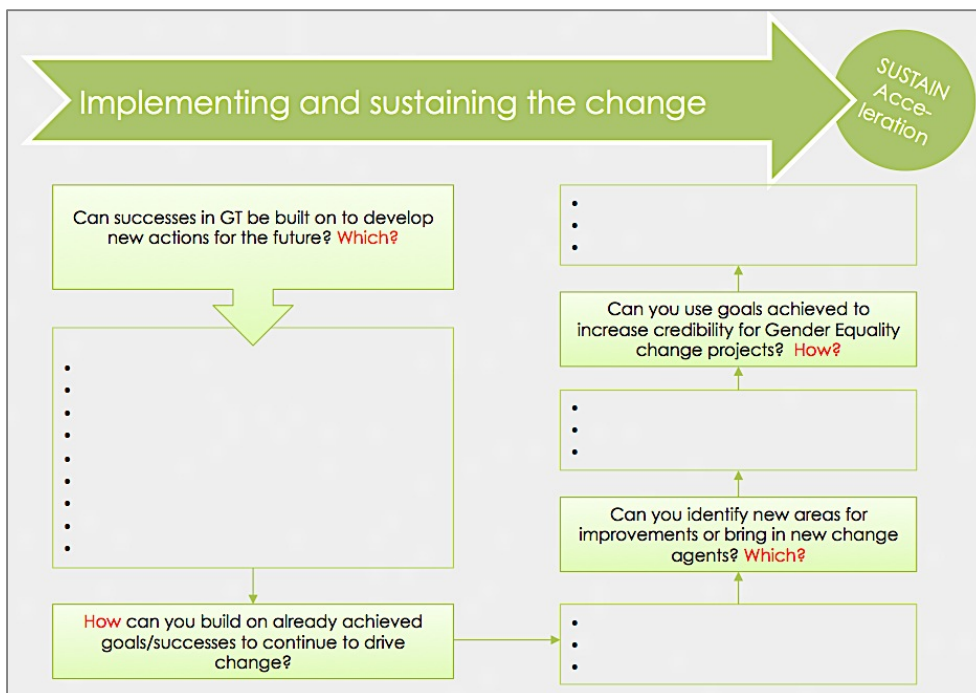
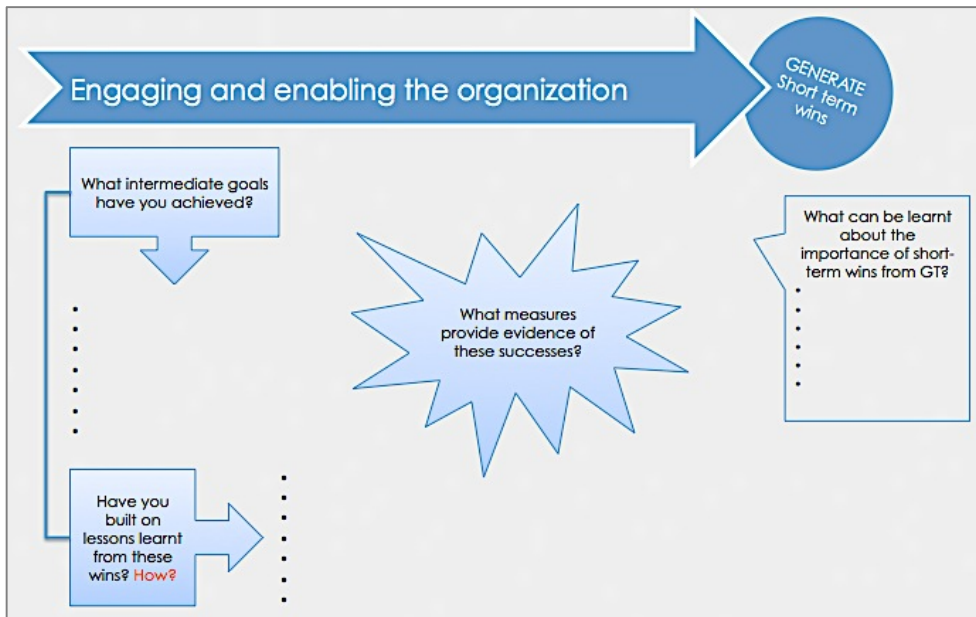
The workshop resulted in interesting discussions, reflections and recommendations. The discussions about the first topic showed how visions had changed over time in the national GenderTime teams. While some had started from minimum requirements, the visions had become more detailed, long-term and sometimes more linked to considerations about sustainability and beyond the project lifetime. The importance of having realistic visions and flexible goals was emphasized as well as communicating the visions clearly in order to engage others in the change interventions.

The discussions about the second topic, about generating short term wins, resulted in recommendations about not starting to implement interventions where the resistance is the greatest. It also provided examples of how small successes could encourage and motivate change agents. Another recommendation concerned addressing personal interest and highlight the personal benefits.

The discussion about the third topic gave rise to recommendations about building a “super team” and bringing in new change agents in the project as part of succession planning. Networking was emphasized, both inside and outside of the organization.

*Figures 11-13: Reflection templates for the Incremental Transformation Workshop*





## **9. Most Significant Change Technique**

### ***Aim of the tool***

The aims of the *Most Significant Change Technique*, used as a monitoring tool in the GenderTime project, were:

- to collect evidence of intangible changes, like behavioural and attitudinal changes.
- to collect information on positive and unexpected changes.
- to collect information about changes of both individual character and organizational character (but from an individual perspective).

### ***Background of the tool***

Rick Davies and his colleagues developed the most significant change methodology already in the 1990s. It is a qualitative form of participatory monitoring and evaluation method that is based on stories (Dart & Davies 2003). It can provide evidence base for project improvement and impact evaluations. When used at the end of a change project it can gather evidence that can be used to introduce changes to future program indicators and contribute to the sustainability of the project.

The most significant change technique gathers rich descriptions that make intangible changes visible. It captures change stories as personal experiences and observations and requires the person that writes down the story to explain why a change was significant for them and provide both descriptive and interpretive information.

The Most Significant Change Technique was a monitoring tool used to collect participatory stories about change in order to monitor and evaluate the impact of the GenderTime project. It focused on learning rather than accountability and identified the effects of the implementation of the GEPs on people's lives and on their working environment. The tool facilitated the gathering of information about outcomes and results from the change project that might have remained hidden if only quantitative monitoring tools had been used. The tool enabled the members of the national GenderTime teams to come to a better understanding of cultural changes and of their own role in the change process.

### ***Structure of the tool***

The Most Significant Change Technique was a monitoring tool that involved the beneficiaries and target groups of interventions, i.e. persons who participated in gender equality actions arranged within the GenderTime project. The tool was designed as a questionnaire with only two, but complex, questions (see figure 14 below). The purpose of these questions was to collect participatory stories by way

of asking informants to reflect over, and share, the most significant change of the GenderTime project from a personal and an organizational perspective, respectively.

The questionnaire template was distributed, together with a separate document with user guidelines, to the seven national GenderTime teams, who identified potential informants, sent the questionnaire to them, collected them again approximately a week later and sent the answers received back to the monitoring team who analysed the responses. The national teams were also given the choice to use the questionnaire as an interview guide and ask the questions in person while taking notes or recording the persons' answers.

In addition, this tool was used as a workshop concept, implemented during the very last project meeting of GenderTime. The tool then allowed the participants in the consortium to share their views on the most significant change of the project. Again, the template was distributed to each national team and during a 30 minutes discussion they reflected on this, taking down notes that were collected for analysis.

*Figure 14: Questions of the Most Significant Change Technique tool*

1) Describe the project related activity/measure that you took part in. Reflect over the most significant change that you personally experienced in relation to participating in the activity.

It can be a direct change (you learnt something) or an indirect change (you made a change due to the information that you learnt).

Please explain why this change is important to you.

*Examples: change in career possibilities and opportunities; change in awareness about gender and gender equality; changes in networks and contacts*

2) In your opinion, what has been the most significant change that has occurred in your institution as a result of the project?

Please describe why this change is/was important to you.

*Examples: change in attitudes, climate and culture; change in administrative routines; change in leadership and management; change in the physical environment etc.*

### ***Results of the tool***

The change stories about the implementation of the GEPs that were produced by this tool were relatively short, but still highly multifaceted and informative. Not including the results from the internal workshop, 23 individual stories about 9 different types of personal changes were submitted. These stories could be categorized into three different types of changes: changes in knowledge/awareness, changes in behaviour and changes in daily lives.

In addition, not including the results from the internal workshop, 31 stories about 14 different types of institutional changes were submitted. Also, these change stories could be categorized into three different types of changes: changes in culture, changes in practices/policies and changes in structures/management. Remaining challenges were also identified: “Increased awareness and discussions did not lead to change in practices” and “No clear change in my institution”.

Based on the stories submitted a theoretical framework could be developed – demonstrating a typology of dimensions of change along the lines of individual/collective, informal/formal and intangible/tangible. This framework and these dimensions were used when implementing the monitoring tool as a workshop concept. The workshop participants were then asked to give examples of changes that could fit into these different categories.

Examples of individual and formal changes were described as involving greater importance in decision-making processes: “Management board includes me in many more issues because I could help improve certain topics with my gender knowledge, which is valued”. The examples of tangible and collective results involved the realisation of a new, gender equal, salary system, gender budgeting, improved gender balance in boards and committees, gender networks and improved communications.

## **10. Final Feedback Report**

### ***Aim of the tool***

The *Final Feedback Report* was the final monitoring tool to be implemented in the GenderTime project. It was, together with the Most Significant Change Technique, implemented during the very last months of the project. The aim of this monitoring tool was to offer all team members the possibility to personally reflect one last time over the implementation process and progress during the last four years.

More specifically, the Final Feedback Report had the following purposes:

- to collect information about lessons learned during the participation in the project
- to sustain the experiences made and make them available for future change agents
- to receive a final reflection on supporting and hindering factors during the implementation processes

### ***Background of the tool***

The Final Feedback Report was developed primarily using an inductive background. The tool was designed taking previous feedback reports into account. The first of these reports, implemented already during the launching phase of the project, was not included in this monitoring handbook because it was more or less just a checklist over the different interventions in the seven different institutions and an update of the GEPs. The second feedback report was the Interim Feedback Report, previously described.

The Interim Feedback Report contributed with important and relevant information during the implementation phase of the project. The information gathered, however, was brief and primarily provided only an overview of the progress, suited for a quantitative analysis. The Interim Feedback Report was therefore considered limited. Although the same excel sheet that was used for the Interim Feedback Report could have been used also for the Final Feedback Report a different approach was considered more relevant and valuable for this final reflection phase.

The Final Feedback Report, in contrast, needed to collect more extensive information about the implementation process. Hence, a more comprehensive, qualitative approach was adopted.

### ***Structure of the tool***

The Final Feedback Report was designed in accordance with the other templates developed in the GenderTime project. As such it was structured as a questionnaire, with three open-ended questions leaving considerable space for the participants to answer by writing quite extensive answers (see figure 15 below). The Final



Feedback Report was distributed the same way, using the same distribution channel as the Most Significant Change Technique, i.e. via mail. Instructions about how to fill out the questionnaire were also distributed together with the template.

The Final Feedback Report collected lessons learned that could be shared with people, who want to take up the challenge to make their institution more gender equal for all. The Final Feedback Report, as a monitoring tool, was thus developed as part of the sustainability of the project.

*Figure 15: Questions of the Final Feedback Report*

1) What kind of recommendation(s) or advice(s) would you give to a person, who wants to start implementing gender equality plans in her/his organization, based on what you have learned during the project?

Please explain why you think this is an important recommendation/advice.

2) If the GenderTime project would start all over again next year – what would you do/organize differently?

Please explain in your answers also why you come to this conclusion.

Remark: Your answers should not be related to consortium related collaboration and work, we are interested in answers about the actual implementation process.

3) And finally, what would you do/organize exact the same way if the GenderTime project would start all over again next year?

Please explain in your answers also why you come to this conclusion.

Remark: Your answers should not be related to consortium related collaboration and work, we are interested in answers about the actual implementation process.

The tool offered all members of the national GenderTime teams the possibility to personally and individually reflect one last time on the implementation process and progress in the project during the four years. Instead of focussing on collecting quantitative data, as previous feedback reports, the tool covered the broader picture taking into account the experiences and knowledge gained by the team members directly involved in the work as change agents in their institutions.

***Results of the tool***

The replies to the first questions included a diverse set of recommendations. Commitment from the top management as well as from middle management was emphasized as important. Here it was stated: “The champion for change is key to success!”. Others however also suggested that it was important to include all staff

in the change efforts and engage them in a participatory way, e.g. by emphasizing the benefits of it for all. To be able to gain such commitment and involvement it was further recommended to “gather evidence!”. This meant undertaking a survey or similar to confirm the relevance of a change project. This would allow you to rely on facts instead of making assumptions about the situation. Awareness raising, defining and setting-up monitoring frameworks, and identifying allies and supports were also suggested as important recommendations.

The replies to the second question, about what they would do different, included improving the timing of the work plan and making modification to the GEPs in order to make them more tangible and manageable. Engaging people to take ownership for the actions planned and build a “task force” in the project was also considered as a way of improving the effectiveness and the results achieved. Defining and setting goals and a common understanding for the organization about what gender equality means was another task that could have been improved. Not all GEPs included interventions for administrative and technical staff in the institutions and this was something that was considered problematic and therefore something that should be changed. Finally, it was suggested that they would not again underestimate the necessity of providing educational and information material on gender equality and gender issues for staff members and this was something that could have been improved.

Finally, the Final Feedback Report provided insights into what had been particularly successful aspects of the GenderTime project. Although most of the participants that submitted their responses to the questions took the opportunity to reflect over possible improvements, several measures and actions in the GEPs were underlined as highly influential, effective and successful. The specific knowledge transfer activities and the inclusion of transfer agents were also mentioned as important features to include in a gender equality change project such as this (Thaler 2016).

## Recommendations for Monitoring

Based on the experiences from monitoring in GenderTime the following recommendations can be formulated.

### 1. Make monitoring an integrated part of the project

This is a comprehensive recommendation that has several implications for the overall **planning** of monitoring. To integrate monitoring in the project means that it is considered a core part of the project and a continuous process throughout the different phases of the project. The implication of this is that monitoring activities should be planned parallel with other project activities.

### 2. Make monitoring a collaborative effort

This recommendation concerns the question about **participation** and who should be involved in monitoring activities. Making monitoring a collaborative effort means involving all project members. This involvement concerns not only the collection of information but also the assessments and analyses of the data gathered. Feeding back the results from the monitoring activities to the project members creates a dialogue about future improvements of the project, also of monitoring. This dialogue is essential for adopting an abductive approach.

### 3. Combine a deductive approach with an inductive

This recommendation concerns the overall strategy for the process of **developing** the monitoring strategy. Combining a deductive approach with an inductive means to draw on research-based literature about organizational change but to also let the monitoring process be flexible and guided by empirical observations of the change process being monitored.

### 4. Adapt to the project phases

This recommendation concerns the **timing** of monitoring. The intensity and frequency of monitoring should take into consideration the specific project phase. The implementation phase e.g. might call for more intense monitoring than the launching phase. The monitoring tools should also have different design depending on the phase. During the final phase, they should e.g. allow for reflections of the complete implementation process (and thus might overlap with certain evaluation activities).

### **5. Combine qualitative and quantitative tools**

This recommendation concerns how to **design** the monitoring tools. A diverse and unique set of monitoring tools should be adopted, with a wide range of methods and performance indicators; questionnaires, checklists, interview guides and workshop concepts for team discussions. Decisions about when to implement which tools should be based on considerations of the project phase. While the launching phase of a project may benefit from quantitative tools, the reflection phase may require qualitative tools.

### **6. Tailor-make monitoring tools**

This recommendation concerns the **character** of the monitoring tools. Tailor-making monitoring tools means to design them while taking into account the specific, local, context that they will be implemented in and allow for them to be adapted to this context. This tailor-making aspect cannot be achieved without monitoring being an integrated part of the project or without it being a collaborative effort.

### **7. Collect diverse and varied data**

This recommendation concerns the **data** collected. Monitoring tools should collect data and information with a focus on diversity, variety and inclusion. The implication of this is that monitoring tools should include the direct involvement of both project members and so-called target groups or beneficiaries of project activities.

### **8. Allocate sufficient resources to monitoring activities**

The final recommendation concerns the **budget** and **person months** dedicated to monitoring activities. Monitoring should not be regarded as a side-product to the actual implementation work. Rather it is an important accompanying measure, which helps to understand underlying processes for enhancing the effectiveness of the practical work. Therefore, sufficient time and financial resources are necessary for both the persons undertaking the monitoring and for the monitored actors.

## References

- Achterberg, S. & Dahmen, J. (2017) Bis hierhin und nicht weiter? Die Grenzen der Geschlechtergleichstellung in der Hochschulpraxis In: Dahmen, J. & Thaler, A. (Eds.) *Soziale Geschlechtergerechtigkeit in Wissenschaft und Forschung*. Barbara Budrich Verlag, Opladen.
- Acker, J. (2000) Gendered Contradictions in Organizational Equity Projects. *Organization* 7(4): 625-632.
- Ahnberg, E., Lundgren, M., Messing, J. & von Schantz Lundgren, I. (2010) Följeforskning som företeelse och följeforskarrollen som konkret praktik. *Arbetsmarknad & Arbetsliv* 16(3): 55-66.
- Arrizabalaga, E., Barnard, S., Clavreul, B. & Polo, L. (2014) *1st Periodic progress report on the implementation of action plans. Deliverable 2.2 GenderTime* (confidential report).
- Badaloni, S. and Perini, L. (2016) *A model for building a Gender Equality Index for academic institutions*. Padova: Padova University Press.
- Barnard, S., Bagilhole, B., Dainty, A. & Hassan, T. (2014) The gendered experience and impact of academic culture in the UK. Paper presented at the *8th European Conference on Gender Equality in Higher Education*, September 3-5, 2014, Vienna, Austria.
- Barnard, S., Hassan, T., Bagilhole, B. & Dainty, A. (2013a) A European Approach to Gender Equality? Paper presented at the *Work, Employment and Society Conference 2013*, 3-5 September, University of Warwick, UK.
- Barnard, S., Hassan, T., Dainty, A. & Bagilhole, B. (2013b) Interdisciplinary content, contestations of knowledge and informational transparency in engineering curriculum, *Teaching in Higher Education* 8(6): 748-760.
- Brouselle, A. & Champagne, F. (2011) Program theory evaluation: Logic analysis. *Evaluation and Program Planning* 34: 69-78.
- Brown, J. B., Hogg, W., Delva, D., Nanchoff-Glatt, M. & Moore, L. (1999) Peer consultation reflection exercise. Innovative strategy for personal and professional development. *Canadian Family Physician* June, 45: 1509-1515.
- Bustelo, M. (2003) Evaluation of Gender Mainstreaming. Ideas from a Meta-evaluation Study. *Evaluation* 9(4): 383-403.

Castaño, C., Müller, J., González, A. & Palmén, R. (2010) *Policy towards gender equity in science and research. Meta-analysis of gender and science research – Topic report*. Luxembourg: Publications Office of the European Union.

Chen, H.-T. (1990) *Theory-driven evaluations*. London: Sage Publications.

Chen, H.-T. & Rossi, P. H. (1983) Evaluating with Sense. The Theory-Driven Approach. *Evaluation Review* 7(3): 283-302.

Chen, H.-T. & Rossi, P. H. (1989) Issues in the theory-driven perspective. *Evaluation and Program Planning* 12: 299-306.

Coleman, G. & Rippin, A. (2000) Putting feminist theory to work: Collaboration as a means towards organizational change. *Organization* 7(4): 573-587.

Dart, J. & Davies, R. (2003) A Dialogical, Story-Based Evaluation Tool: The Most Significant Change Technique. *American Journal of Evaluation* 24(2): 137-155.

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

D'Eon, M., Sadownik, L., Harrison, A. & Nation, J. (2008) Using Self-Assessments to Detect Workshop Success. Do They Work? *American Journal of Evaluation* 29(1): 92-98.

DG (2011) *DG Research and Innovation Workshop on How to Advance Structural Change in order to improve Gender Equality in Research Organisations in Europe?* 30.6-1.7 2011, Brussels, DG RTD B6

Equality Challenge Unit (2014) *Measuring progress on equality: qualitative evidence*. Equality Challenge Unit, April 2014.

Frisby, W., Maguire, P. & Reid, C. (2009) The 'f' word has everything to do with it. How feminist theories inform action research. *Action Research* 7(1): 13-29.

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

Haylock, L. & Miller, C. (2016) Merging Developmental and Feminist Evaluation to Monitor and Evaluate Transformative Social Change. *American Journal of Evaluation* 37(1): 63-79.

- Kotter, J. (1995) Why transformation efforts fail. *Harvard Business Review* 73(2): 59-67.
- Krizsan, A. & Lombardo, E. (2013) The quality of gender equality policies: A discursive approach. *European Journal of Women's Studies* 20(1): 77-92.
- Kusek, J. Z. & Rist, R. C. (2004) *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*. Washington, DC: The International Bank for Reconstruction and Development/ World Bank Publications.
- Lindgren, L. (2012) En utvärderare om Högskoleverkets system för kvalitetsutvärdering 2011-2014. *Statsvetenskaplig tidskrift* 114(3): 471-482.
- Lindgren, L. (2011) *Hur används brukarundersökningar? En översikt av forskningen och perspektiv på kunskapsöversikter*. FoU i Väst, Rapport 3:2011.
- Linstead, S., Brewis, J. & Linstead, A. (2005) Gender in change: gendering change. *Journal of Organizational Change Management* 18(6): 542-560.
- Lipinsky, A. & Schäfer, M. (2016) *INTEGER Guidelines for the Self-Assessment of Transformational Gender Action Plans set up in Higher Education and Research Institutions*. GESIS Leibniz Institute for the Social Sciences, Cologne, Germany.
- McRoy, I. & Gibbs, P. (2009) Leading change in higher education. *Educational Management Administration and Leadership* 37(5): 687-704.
- Mullen, E. (2002) Multi-site Evaluation and Research. *Socialvetenskaplig tidskrift* no 2-3: 175-193.
- OECD (2002) *Glossary of Key Terms in Evaluation and Results-Based Management*. Paris: OECD (Organisation for Economic Co-operation and Development).
- Patton, M. Q. (2011) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation*. New York: The Guildford Press.
- Parsons, E. and Priola, V. (2013) Agents for change and changed agents: the micro-politics of change and feminism in the Academy. *Gender, Work and Organization* 20(5): 580-598.
- Pawson, R. & Tilley, N. (1997) *Realistic Evaluation*. London: Sage.
- Pawson, R. & Tilley, N. (2001) Realistic Evaluation Bloodlines. *American Journal of Evaluation* 22(3): 317-324.

- Pawson, R. & Tilley, N. (2004) *Realist Evaluation*. London: British Cabinet Office.
- Peterson, H. (2010) Women's Career Strategies in Engineering: Confronting Masculine Workplace Culture. Godfroy-Genin, A. (red.) *Women in Engineering & Technology Research*. Berlin: Lit Verlag.
- Reid, C., Tom, A. & Frisby, W. (2006) Finding the 'action' in feminist participatory action research. *Action Research* 4(3): 315-332.
- Rossi, P. H., Freeman, H. E. & Lipsey, M. W. (2004) *Evaluation. A Systematic Approach*. Thousand Oaks: Sage Publications.
- Scriven, M. (2013) *Key Evaluation Checklist*. Claremont Graduate University & The Evaluation Center, Western Michigan University.
- Scriven, M. (2002) Den nya utvärderingsvetenskapen. *Socialvetenskaplig tidskrift* no. 2-3: 104-117.
- Shaw, I. (2002) Directions in Qualitative Evaluation. *Socialvetenskaplig tidskrift* no. 2-3: 144-159.
- Siebenhandl, K. & Mayr, E. (2015) GenderTime: Evaluation of Measures for Women Encouragement in Science Organisations. *WomENCourage online proceedings*. ACM: Uppsala.
- Sridharan, S. & Nakaima, A. (2011) Ten steps to making evaluation matter. *Evaluation and Program Planning* 34: 135-146.
- Thaler, A. (2016) *Learning organisations in science and research: The role of transfer agents in gender equality change processes*. IFZ, University of Klagenfurt, Graz.
- Thomas, R. & Davies, A. (2005) What have the feminist done for us? Feminist theory and organizational resistance. *Organization* 12(5): 711-740.
- Tilley, Nick (2000) Realistic Evaluation: An Overview. Paper presented at the *Founding Conference of the Danish Evaluation Society*, September 2000.
- UNDP (2009) *Handbook on planning, monitoring and evaluating for development results*. United Nations Development Programme. New York.
- Vedung, E. (2010) Four Waves of Evaluation Diffusion. *Evaluation* 16(3): 263-277.



Whelan-Berry, K. S. (2013) Gender and change drivers: and exploratory study. *Gender in Management: An International Journal* 28(5), 256-277.

Williams, B. & Hummelbrunner, R. (2010) *Systems Concept in Action: A Practitioner's Toolkit*. Stanford: Stanford University Press.