

**Disappointing pioneers and surprising laggards.
Understanding the politics of ranking and reputation**

Anna van der Vleuten & Mieke Verloo

a.vandervleuten@fm.ru.nl m.verloo@fm.ru.nl

Paper for the Third Pan-European Conference on European Union Politics

21-23rd September 2006, in Istanbul

Panel 336 *Stretching and Bending Gender Equality*

1. Introduction

This paper has originated in the minds of two Dutch researchers of gender equality policies who are used to be approached as coming from a pioneer women-friendly country, and who keep being surprised that the positive image of the Netherlands travels so much better than criticism or even realistic evaluation of Dutch performance in this field. Our surprise translated into genuine academic curiosity when we realized that the Dutch reputation is supported by high scores in rankings computed by the UN and the EU.¹ Such rankings seem to proliferate lately, gaining in importance for the assessment of countries' performance.

Even though ranking has the appearances of a technical exercise, based on expert knowledge and depoliticized facts and figures, we argue that power is as central to it as to other policy instruments. This has led us to explore the origins and consequences of what we consider shifts in 'the politics of reputation' due to the increased importance of new policy methods such as ranking. We investigate the power-based mechanisms which operate between each pair of these three concepts, reputation, ranking and performance. How does reputation relate to ranking, as a good reputation seems to influence the way in which rankings are constructed as much as rankings are constitutive of a good reputation? How does reputation relate to performance, and to what extent do we see disparities between them, and how does ranking relate to performance? In this paper we try to conceptualize the consequences of the 'new' politics of ranking that these processes generate, and we investigate the institutional arrangements on ranking, in order to contribute to a political understanding of the quality of policy implementation.

¹ The Netherlands rank 12th in the global Gender-related Development Index GDI, and 8th on the Gender Empowerment measure GEM (Country Fact Sheets The Netherlands, UNDP, http://hdr.undp.org/statistics/data/country_fact_sheets/cty_fs_NLD.html). The Netherlands rank 3rd on Women members of single/lower houses of national parliaments in EU member states 2005 (Report on equality between women and men, 2006, European Commission Brussels) http://ec.europa.eu/employment_social/publications/2006/keaj06001_en.pdf

The case we study is gender equality policies in the European Union. In this field, a shift in policy and implementation style has taken place in the 1990s from legal instruments aiming at harmonization of national policies, enforced by the Commission and the Court, to policy instruments such as the coordination of national policies (Open Method of Coordination, European Employment Strategy), benchmarking and scoreboards (Héritier 2003). Here, the European level reflects changes at the national level, where a similar shift seems to take place in the field of gender equality (see Benschop and Brouns 2006 on the Glass Ceiling Index in the Netherlands, which ranks companies). Although the emergence of using ranking as a policy measure can be found in many domains at the European level the effects on reputation and performance can be assumed to be stronger if the policy domain is less ‘technical’ and more linked with power relations. For that reason, gender equality policies constitute the domain *par excellence* where the effects of these mechanisms – if any – should be most visible.

2. Reputation, ranking and performance

States value their reputation. It is part of their power position with regard to domestic society and to other states (Van der Vleuten 2005). Rönnblom notes how Sweden proudly presents itself in policy documents as the ‘most gender equal country in the world’ (Rönnblom 2004: 2). A good reputation legitimizes the course a government has chosen and may even serve to legitimize further inaction as a good reputation once acquired lasts a while. It therefore is an asset in international negotiations as well. Conversely, ‘being perceived as a cheater and free-rider undermines the bargaining power of a member state’ (Börzel 2003: 203). If others are convinced of the qualities of a state in the past, it is able to convince others more easily of the quality of its actual ideas. Generally spoken a state with a good reputation needs to invest less means to obtain a positive result than a state with a bad reputation, known for its poor performance and its lack of respect for previous agreements.

Reputation is based not only on *behavior* or performance in the past, but essentially on the way in which this behavior or performance is represented. This rhetorical dimension makes reputation malleable to a certain extent, though reputation has to remain rooted in facts to avoid a sudden loss of credibility when the facts are checked. In fact, rhetorical action (defined as the strategic use of arguments) follows some rules, described by Schimmelfennig as: a requirement of consistency, and credibility, and a need to appear convincing, with a resulting preference in favor of obscuring or hiding inconvenient facts or norms (Schimmelfennig 2003). Schimmelfennig also points at an interesting unintended consequence of rhetorical

action. Even if a state uses rhetoric to enhance its reputation in order to gain more leeway in negotiations at the European level, its *marge de manoeuvre* may be involuntarily limited by the ensuing rhetorical entrapment (Schimmelfennig 2003).

The concern for their reputation plays a role in explaining state compliance with EU policies. If a state prefers not to implement a policy, it will prefer this non-compliance to remain unnoticed, so its reputation as a reliable member state does not suffer. As the EU monitoring system increases the probability that non-compliance is discovered, the threat of ‘being named and shamed’ is supposed to influence the decision of a state to implement (Tallberg 2002). However, these ‘old’ politics of reputation not only consist of monitoring but also – in the case of ‘hard’ law – of enforcement measures. It is these enforcement measures that enable the European Commission to further investigate whether state rhetoric corresponds with the facts, and to unveil instances of non-compliance by taking a state to the Court. With each stage of the procedure, the publicity increases as well as the ‘reputational costs’ (Börzel 2003). Member states prefer not to have infringement proceedings in front of the Court against them, as usually the Commission wins the case and Court rulings receive much publicity, ‘further raising the social costs of non-compliance’ (Tallberg 2002) – damaging their reputation. The involvement of the Commission and the Court implies that an ‘absolute’ standard is applied, that is, the standard as laid down in the European policy at stake is interpreted by the Court for all member states in the same way. This mechanism enables not only the assessment but also the judgment of disparities between reputation and performance.

France, for instance, cherished its reputation as a social role model, defending gender equality from the very beginnings of the development of European integration. The European Commission, however, had to take France to the Court more than once because it did not comply fully with the gender equality directives. When in 2000, just when it occupied the EU presidency, France risked to be condemned to a daily fine of 142,425 euros if it did not implement the equal treatment directive, the government succeeded in convincing the French parliament of the necessity to act immediately by pointing, not at the financial consequences but at the consequences for its reputation. To be the first member state ever to be condemned to a penalty payment in social policy was unacceptable for France as would-be champion of ‘social Europe’ (Assemblée 2000).

The ‘old’ politics of reputation, linked to legal performance, are thus based on ‘absolute’ standards authoritatively interpreted by a third party, and the enforcement of this standard. We argue that the ‘new’ politics of reputation, linked to assessed performance and ranking, are based on relative standards and therefore exclude the possibility of judging and sanctioning low performance. While reputation is an old

concept in, among others, theories of state behavior, analyses of political leadership and management theories, ranking is a relatively new one. Since the 1980s, New Public Management and other similar trends in professionalizing bureaucracies have given birth to a panoply of indicators, procedures, measures, and other innovations meant to assess the performance of public administration systems, and to improve the performance and the political accountability of governments (de Vries 2005). Intended to be instruments for improving the quality and transparency of policies, they have given new meaning to the old politics of reputation by changing the basis on which reputation is built and the mechanisms by which reputation is judged.

The relation between reputation and ranking varies, as ranking refers to different practices. It includes the *establishment of indicators*, which translate or disaggregate policy goals in quantitative objectives. Based on its score on such an indicator, states or companies may then be ranked and classified as pioneers or laggards. Under conditions of measuring and ranking performance, reputation is no longer based mainly on shared understandings between state-society and among states, but has to be 'proven' in technocratic processes. The development of technocratic processes to measure performance implies that many decisions are being made as to which aspects of a phenomenon will and can be measured. In policy practice this has substantial consequences, as different states will be at an advantage or disadvantage dependent on the chosen ways of measuring. It then matters which procedures are chosen to select indicators and who has the power to decide that data have to be produced to fill them.

Another method of ranking consists of *selecting best practices*, pilot projects and benchmarks. Best or good practices are assumed to offer the best solution to a problem and are then presented as the model to be copied. This method does not result in an absolute standard which all have to apply and where non-conformity could lead to coercive measures. Even shaming will be of limited use, as a good reputation not necessarily presupposes the proposition of best practices and the development of different practices not necessarily leads to failure in solving the problem for which the 'best practice' was recommended. Nonetheless, there might be an interesting connection between this aspect of ranking and reputation which goes in the opposite direction: to deck out a new strategy, best practices may be collected although no real practice is in existence yet. In that case, practices proposed by states with a good reputation in the field will have more chances to be adopted than practices proposed by states which are perceived as laggards. Therefore the standard to implement, the model 'best practice', is neither a generalized absolute standard nor a relative one, but the specific standard of a

state considered to be a pioneer ('pioneer standard'), thus further confirming the reputation of the pioneer.

In the 1990s, the old politics of reputation based on 'legal performance' have resulted in scoreboards, showing whether states implemented EU directives on time. In spite of the strong monitoring and enforcement system in the EU, differences persist when it comes to the performance of member states. Looking at legal implementation, some states are considered pioneers, showing a good record in the field of EU gender equality policies, while others are labeled laggards, poor performers (Liebert et al. 2003; Falkner et al. 2005). This kind of labeling focuses solely on legal equality between women and men, which has been criticized for not always reflecting adequately de facto gender equality (Lombardo 2003). We are well aware that the link between ranking and performance based on legal implementation is problematic as well. We agree that the standard ('legal implementation') is no appropriate standard if one wants to measure and compare actual performance (the attainment of policy objectives), but at least it is an 'absolute' standard which is applied to all states in the same way, as opposed to rankings based on relative or pioneer standards.

We argue that the politics of ranking, against its presented ambitions of improving performance and policy learning (Trubek and Trubek 2005), works against the quality and even the transparency of policy making, in that it contributes to processes of gerrymandering and bureaucratic overload that may hinder innovation and improvement. One reason why this link between ranking and performance is problematic, results from the way in which objectives are unidimensionally quantified. Concepts and indicators are chosen on the basis of availability of data or because of the preference of member states, without any transparent democratic debate. The consequences are that concepts are stretched and bended to fit the available data and preferences. As a wicked or messy problem, gender equality is easily subjectable to such processes. As a result, the assessment of performance is of poor quality, it is unclear as to what specific kind of performance is actually measured and unable to offer insight in the ways in which performance could be improved. As the indicators chosen are not simply informative about progress on gender equality, but also tell a story about the power relations between member states and between member states and the European Union, this story needs to be unraveled before measures can be designed to counteract the perverse effects of the politics of ranking.

3. Developments in measuring and benchmarking gender equality

Gender equality is a domain where the EU has strongly influenced national policies, as since the 1970s it has developed directives which had to be transposed in national legislation, it has promoted the establishment of transnational networks and it has financed projects. Since the 1990s, the range of mechanisms to promote gender equality in European countries and at the level of the European Union has been expanded from equal treatment in legislation, focused on providing equal access and correcting existing inequalities in legislation, to a strategy of specific or targeted gender equality policies that often take the form of temporary projects for specific groups of women, and the strategy of gender mainstreaming (Verloo 2001). There is also an increasing use of gender indicators and benchmarks (Luxembourg Presidency Report 2005). The growing use of gender indicators and benchmarks has been fostered since the 90s. At the level of the European Union, the introduction was part of the development of the Open Method of Coordination (OMC) and the European Employment Strategy (EES), and in the field of gender equality policies these European developments received a strong impulse through the Beijing review process, especially the Beijing +5 and the Beijing +10 exercises in 2000 and 2005 (Rubery 2003).

Measuring and benchmarking gender equality in the OMC and EES

The OMC and EES processes have given member states many new opportunities to represent their performance towards gender equality, and hence their reputation. Gender equality as a European policy goal has become embedded as a key component of the EES (Luxembourg 1997, Lisbon 2000). In its ambition ‘to become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion’ (Lisbon 2000), gender equality is integrated as one of the ten guidelines on which member state progress on economic policy is to be assessed at EU level. For this assessment and the facilitation of progress, a new process of policy development, the open method of co-ordination (OMC), has been created. Interestingly, an important impulse to the inclusion of gender equality in the EES has been given by a country that does not exactly have a strong reputation in this field. According to Rubery (2003, p. 3), “the Austrian ministers for women and for labour in post during the 1998 Austrian presidency, both of them women ministers, had their sights set on promoting equal opportunities further within the EES. To that end they organised an informal joint council of ministers for labour and gender equality, commissioned an extensive report and used this momentum to push for the

successful inclusion of the commitment to gender mainstreaming in the EES guidelines at the December summit”².

The OMC is a semi-voluntary form of coordination (Scott and Trubek 2002) that is more flexible and participatory than the traditional Community legislative method, and that consists of a policy mix of targets, best practices and monitoring. Targets for member states are set at the meetings of the Spring European Council (using Structural Indicators, overseen by Eurostat). The Council organizes the annual monitoring of progress towards these targets through peer review and the Commission issues recommendations to individual member states. While the targets are set by the European Council, each member state can develop its own strategy to reach the targets, paying attention to its national specificities. For the purpose of mutual learning, a process of identification and exchange of best practices is organized between the member states, and guidelines are developed on the basis of best practices that have to be incorporated in the National Action Plans. The implementation of the guidelines is monitored in the JER (Joint Employment Report), and leads to a renewed identification of best practices, and to a new start of the OMC cycle (Biagi 2000).

For gender equality, the most important commitments made by the European Council are those concerning the narrowing of the gender gaps in employment (women’s employment rate to rise to 57% by 2005 and to 60% by 2010) and unemployment (Lisbon 2000, Stockholm 2001), as well as the provision of childcare for at least 90% of children between 3 years and school age and 33% for children under 3 (Barcelona 2002) and pay (Brussels 2003).

The main objectives of the OMC and EES processes are to promote mutual learning, and to foster the generation and diffusion of new ideas and practices across Europe. This postulates the member states as a collectivity that is eager to learn and committed to finding the ‘best’ solutions to any of the challenges ahead. While these postulations can be questioned on the basis of past performance of states in implementing gender equality directives (Van der Vleuten forthcoming), and there are fierce debates as to whether or not the OMC is delivering on its promises (Héritier 2003; Trubek and Trubek 2005), it is possible to draw some basic conclusions from the decade long existence of this method for the field of gender equality. One such basic result of the OMC and EES processes is that there is an ongoing production of ranking of the member states on employment and employment related issues. However, the targets only have ‘a general orientation function’ (Héritier 2003: 118) and member states are pressurized but

² Rubery notes that the later change in government brought an abrupt change in implementing gender mainstreaming in the EES in Austria.

not compelled to meet them. The peer review and the evaluation and recommendations by the Commission constitute a monitoring mechanism without sanctions other than ‘naming and shaming’ by publishing the performance schedules. For an underperforming state, the reputational costs of OMC and EES processes remain therefore limited as compared to the costs occasioned by the enforcement system linked to legal instruments.³

Another result is the production of best practices in the field. Member states are ‘encouraged to benchmark their performance against the best performer in the Union’ (Héritier 2003: 117). This implies that one of the national solutions is considered to be the best one, without questioning *all* practices and investigating which practice suits which national context best. National institutions are not called into question the way in which legal instruments could oblige a state to revise its central concepts or practices: the way in which, for instance, the equality directives required the Dutch government to fundamentally revise its breadwinner-centered system of pay and social security in spite of Dutch arguments that this system served well women’s needs.

In addition to the information given by member states, the Commission uses expertise from Networks of Gender Experts to assist in its assessment of member state progress in this area. They are another source of ‘labelling’ best practices, creating ‘pioneers’ in the process of doing so. In their review of all the pan-European gender impact assessments carried out, Rubery and Fagan (2000) position the Netherlands an early starter on gender impact assessment, together with the Nordic countries except Denmark. They then incorporate the Dutch conceptual framework for gender impact assessment as a model for what they recommend as a required further conceptual and practical elaboration of the gender impact assessment methodology. Interestingly, when doing that, they represent the conceptual framework as an academic one, rather than as a national Dutch practice (Rubery and Fagan 2000, p.5).

The 12 critical areas of concern and the Beijing review process

Another important impulse to the development of measuring and benchmarking gender equality has been given by the follow-up of the 1995 Beijing UN conference. Under the leadership of the United Nations, the Beijing Conference in 1995 has resulted in a Platform for Action to bring about gender equality in the world, which has been adopted by all European Union countries. At two points in time, in 2000 and in 2005, the UN took

³ This does not hold for the coordination procedure under the EMU and the Stability and Growth Pact, where sanctions are possible, although practice has shown that the application of sanctions even in this area is an essentially political decision, as the Court is not involved. Therefore, in spite of the possibility of sanctions, there is no absolute standard-setting and enforcement.

further action to monitor the progress made. On each occasion, governments were asked to report on their actions to implement the Platform for Action in the 12 critical areas of concern. In 2000, the European Parliament adopted a Resolution acknowledging ‘that various aspects of women’s life in the EU have improved since 1995, whilst deploring the lack of agreed indicators, gender-segregated data, benchmarks and a clear timetable in the Platform for Action, this being a major obstacle to proper evaluation of progress made over the last five years’ and urging ‘the participants at the New York Conference on Beijing +5 to establish indicators, benchmarks and a precise time schedule’ (European Parliament resolution on the follow-up to the Beijing Action Platform 2000/2020(INI)). This resolution clearly shows the importance of the Beijing review process for the development of pressure towards measurable output of gender equality policies at the European level.

In recent years, the European Council presents itself as playing a leading role in the development of indicators to operationalise the specific objectives concerning gender equality (Luxembourg Presidency Report 2005). In 1998, the Council committed the EU to the development of a simple suite of indicators to monitor progress on all 12 critical areas of the UN Platform for Action. Indicators have since been proposed by a series of EU Presidencies: Finland 1999, women in power and decision making; France 2000, reconciliation of family and working life; Belgium 2001, gender pay gap; Spain, Denmark, Ireland, Greece, Netherlands, 2002-4, violence against women. In addition, some indicators related to employment have been agreed by Councils as part of the EES, as described earlier. The EU now has indicators in five of the twelve critical areas: Women and the Economy (gender gaps in employment, unemployment and pay; provision of care for children and other dependents); Women in Power and Decision-Making (percentage of women among elected positions, ministerial positions, executive boards of top companies, and boards of central banks); Women and Poverty (percentage of women among those ‘at-risk-of-poverty’); Women and Education (educational attainment; life-long learning; science and technology graduates); and Violence Against Women (domestic violence: number of female victims, types of victim support, measures to end violence; sexual harassment at work).

[In the period after 1995, progress reports have been prepared by the European Commission services for the annual reviews of different critical areas of the Platform for Action by the UN Commission on the Status of Women (CSW).

4. Problems with measuring progress in gender equality

The increasing use of gender indicators and benchmarks as linked to the OMC and Beijing review processes has a number of problems that are relevant in the context of this paper. These are connected to the shift from absolute to relative standards of quality, the selections made in the production of targets, benchmarks and indicators that lead to a reduction of progress or performance to what is measured in these targets, benchmarks and indicators, the inherent normativity of benchmarks, targets and indicators as a means to monitor progress or good quality of policies, and the increasing understanding of quality as a matter of technical measurement as opposed to political judgment or opinion. All these problems result from power processes in the construction or use of indicators.

The shift to relative standards

Except for the targets in the EES, the indicators are used for relative ranking of member states, where the graphs presented mention the 'EU average' as an important benchmark for its member states. The absolute targets for the EES are not de facto targets for everyone, as some member states are out-performers on these targets from the start. The shift to relative standards automatically produces rankings. In spite of Europeanization, differences between states persist, so the annual reports show. The presentation of performance on the chosen indicators seems to 'prove' which states are doing well, and which states are under-performing. The ranking that is the result of the use of the indicators however can be shown to have many problems.

Mosesdottir (2005), in the final report of the WELLKNOW project, concludes that 'the emphasis made by the EES of ranking countries according to their performance on narrowly defined targets like, for example, the female employment rate runs the risk of making those countries that out-perform the others the unachievable standard for member countries falling behind. Economic, political and social contexts cannot be copied over night. Moreover, those nations performing well tend to focus more on their relative position than on their distance from the EU's employment and gender objectives' (p.53).

The feasibility study for an EU gender equality index commissioned by and presented to the European Commission made by Plantenga (et al 2003) shows how various indices can produce very different rankings. The point here is not to argue about the respective quality of these indices, but to accentuate that the choice for one or the other also has substantial implications for the reputation of states. The Annex to this paper shows two indices from the this study, showing the EO indicators that would put Portugal and Belgium ahead of Denmark and Sweden, and would put Finland on position

9 (EU 15), and then also the new index proposed by the authors of this study that present a ranking that fits the classical expectations of ‘good’ gender equality states, listing Sweden, Finland, Netherlands in the top three.

The implicit and explicit selections in the development of targets, benchmarks and indicators

The usual premise to defend the necessity of measuring gender equality and of developing gender equality indicators and the data to support them – whether done by policy makers or by academics - is that they are necessary to know whether gender equality policies are working (Plantenga et al 2003, Luxembourg Presidency Report 2005). There are complaints about the lack of indicators across Europe, and about the lack of data to support the indicators (Luxembourg Presidency Report 2005). The development of indicators at the European level has indeed resulted in an incomplete suite of indicators, where the most developed indicators are found in the field of the economy, encompassing aspects of education, social inclusion and poverty (Walby 2005). This points to a first problem of the benchmarking and measuring policy developments: the problem of the implicit and explicit selections that underline the chosen indicators. Constructing good indicators is a difficult task, the more so if measuring progress towards gender equality intends to measure structural progress and not simple inclusion, if they aspire to measure empowerment and if they intend to avoid the unitary category of women. Moreover, some domains of gender equality are given more importance in the construction of indicators than others, influencing the rankings that result from them. This is connected to the variety of framing gender inequality as a policy problem and gender equality as a goal across Europe that as such is a major element in the dynamics of developing ranking systems.

As the European experience shows, in this process certain fields are favored over others, and the ones that are favored are most close to fields that dominate policy making anyway. When it comes to gender equality policies, this means that we can expect to find (and we indeed do find) more indicators on gender inequality in the public sector than on gender inequality in the private domains, more indicators on numbers of women than on empowerment⁴ of women [=? short definition to avoid confusion]. Additionally, we can observe that the gender inequality problematic gets fragmented into a number of different fields, overlooking the complex interrelationships between gender inequalities in citizenship, labor, and intimacy.

⁴ Empowerment is used here in the sense of giving voice to non-hegemonic actors in a society (Verloo 2005; Fraser 1989; Fraser 1997).

In 2000, one year after the new right wing government that included Jorg Haider had been installed, Austria developed a new law on domestic violence based on an innovative model of conceptualizing state intervention in this field. Six years later, after several other countries had adopted (parts of) what has come to be called the Austrian model, the Netherlands, after long studies and extensive debates, introduced a similar law. This example shows that what is seen as a laggard country can be the origin of much needed innovation. Surprisingly, this does not mean that Austria is seen as a pioneer in gender equality policies. This could be connected to the fact that there are no targets, benchmarks or indicators on gender violence. What is not measured then does not count (as much) Moreover, Austria does not see itself to be a pioneer. In the Beijing +10 review process, the Austrian government had to be reminded by NGOs that they had omitted any reference to this innovative contribution to gender equality policies. They then took further action, and organized a special event at the CSW sessions for the Beijing +10 review in New York to 'promote' this law and its implementation⁵.

The inherent normativity of benchmarks, targets and indicators

A further difficulty related to benchmarks, targets and indicators is that they are inherently normative, and that most actors are well aware of this in their actions, even if they play down on it for the sake of strategy. [this may work as practical consciousness in the sense of Giddens (1984)]. To discourage gerrymandering and other forms of lying with statistics, the use of independent actors that make conclusions based upon the benchmarks, targets and indicators would be an option, as is the organization of decent transparent democratic procedures. None of these are available at the moment. Experiences with evaluation or monitoring independent from state actors are still rare. The Netherlands provide one of the few examples of both independent monitoring of the process and of the results of gender mainstreaming, and hence allow some reflection on the experiences with it. The bi-annual monitoring report on gender equality in the Netherlands however, monitors mainly progress towards gender equality in society, and does not relate this progress or regress to the efforts made in gender equality policies in general or gender mainstreaming in particular (Portegijs, Boelens and Oltshoorn 2004). The independent assessment committee currently assessing the quality of gender mainstreaming at the Dutch ministries has just submitted its critical mid term review to Parliament (Visitatiecommissie Emancipatie 2005) arguing that there is very little activity on gender mainstreaming, a lack of internal expertise and a lack of using external

⁵ Interview Rosa Logar, Chair shelter network Austria.

expertise too. This lack of expertise is used as an excuse for inactivity. Moreover, the whole infrastructure on gender equality policies has slowly been destructed, and international obligations are ignored. It remains to be seen whether this review will be able to influence the further direction of gender mainstreaming in the Netherlands, as it has not been well received at all by the Ministries criticized by it.

Understanding quality of performance as a technical matter

While it has been noted by earlier scholars of the OMC processes that there are problems with indicators, this has usually been understood as a matter of lack of expertise in a technical sense. Some of them are seen to be connected to the need for adequate and comparable data. Serious problems can arise when the data need to be collected to ‘fill’ the indicators. This often necessitates a process of gender mainstreaming at the level of official statistical institutions, and there are scarce documentations of gender mainstreaming in such institutions at this moment. This is not only a practical or technical problem, even if it is usually represented as such, but also a political one. While it implies that, starting from scratch, even under the best of conditions, it will take several years before substantial indicators can be available, and comprehensive monitoring can be possible, it also asks for an investigation of the decision making power of the different actors involved. It is unclear to many who exactly has the power to start implementation of agreed upon indicators, as the interaction between the Parliament, the Council, the Commission, Eurostat, and the member states is extremely opaque.

Rubery (2003), while also representing the problem to be of a rather technical matter, describes problems that occurred in the field of gender equality as “problems with respect to specific expertise in analysing developments with respect to equal opportunities or the use of indicators or benchmarks to assess progress” (p.16), gives two examples that show the political character of the use of benchmarks in the Joint Employment reports that provide the official review of the NAPs and progress towards achieving stated objectives. In one case, she notes, Greece was commended for achieving a closure of the gender employment gap, but without reference to the fact that this was only achieved through a fall in the male employment rate. Similarly changes in the segregation index were used to indicate progress in desegregating labor markets, without these changes being decomposed into affects related to structure and labor force change and into changes in the gender shares within occupational categories. She concludes rightly that there is evidence of the use and misinterpretation of misleading indicators, but does not analyse this in terms of political processes and power.

Plantenga et al (2003), who were commissioned to draft a composite gender equality index that has never been used, admit this political edge to it, but argue that it would not make much difference in the resulting ranking. They show that one different way of measuring gender gap in participation, from a headcount basis –as they did - to measuring the gender gap in FTE, would only affect the position of the Netherlands and the UK (i.e. lower this position) in the overall ranking (p.47).

5. Discussion – the Perverse Politics of Ranking

In this paper we set out to discuss the consequences of shifts in policy instruments for the ‘politics of reputation’. We have explored the power-based configuration of performance-ranking-reputation which has been created by these shifts and the perverse effects on performance. This section offers some tentative, preliminary ‘conclusions’ which definitely need further elaboration and empirical underpinning.

First, we have seen a shift from legal instruments to scoreboards and benchmarks. This shift favors states with a policy-oriented style. Thanks to its pilot projects and corresponding rhetoric, a state like the Netherlands will be considered a pioneer more easily than a state like Austria which has a legalistic governance style and less catchy projects – although pilot projects not necessarily result in better de facto performance than legal instruments. If analysed more closely, pioneers may therefore disappoint and laggards surprise: not gender-champion Sweden but laggard Austria took the initiative to include gender equality in the EES in 1998.

Ranking strengthens the position of certain states depending on policy style, not on performance.

Politics of ranking thus allows pioneers to do window-dressing instead of assuring real implementation whereas some so-called laggards perform really well when we focus on their de facto performance.

Ranking distorts the value of performance and therefore is problematic as instrument for the improvement of performance, as both pioneers and laggards lack an incentive to do better.

Second, we have seen a shift from political processes to so-called technical processes. Instead of negotiations between Commission, Council and European Parliament on new policy instruments to be approved by the national and European parliaments, and negotiations between governments, Commission and Court on the quality of performance (compliance with the requirements laid down in legal instruments), seemingly depoliticized facts and figures, translated in a position on the scoreboard, ‘prove’ the quality of performance. Of course, the fact that indicators do not assess correctly policy performance is not new and we have provided examples of the problems associated with the construction of indicators. It is also generally admitted that every indicator has its drawbacks. However, we argue that this is often erroneously considered to be a technical problem, which requires a technical solution and expert knowledge, instead of recognizing that the choice of an indicator always is a political decision, which therefore necessitates transparency as to who decides and the involvement of political actors who can be held accountable for these decisions.

Ranking is not merely a technical process but essentially a political process, which should be transparent and subject to democratic control.

It would be interesting to reconstruct the decision making process on the existing EU indicators to see to what extent not only the availability of data (a partly technical matter) but also the preferences of states have played a role in the choices that have been made. We have also noticed that states which have a good reputation are able to influence strongly the design of further indicators in the same area. This enables them to ensure that their national practice is optimally ‘graded’ by this indicator, thus further strengthening their reputation.

Third, there is a shift from ‘absolute’ to ‘relative’ standards. There is no authoritative third party for the interpretation and enforcement of a standard, equal for all, but a power-driven process where targets and standards are chosen in function of existing national practices. ‘Relative’ monitoring (peer review) replaces ‘absolute’ monitoring. There is no sanctioning mechanism apart from ‘naming and shaming’. Reputational costs therefore remain more limited. In addition, the third party (Commission and Court), have an interest in assuring compliance with EU (gender equality) legislation in order to keep up their reputation as guardian of the Treaties, whereas member states in a peer review process have an interest in keeping up their ‘joint’ reputation and show their commitment to gender equality. They do not have an interest in blaming each other.

The politics of ranking reduce reputational costs of poor performance as compared to the old politics of reputation.

We contend that the politics of ranking have perverse effects on the link between performance and reputation. In spite of their aim to improve performance through the potential consequences of ranking for reputation, they therefore hinder the improvement of the quality of performance.

The tragedy of gender equality policies is that it is unlikely that intentional intervention on a limited number of issues will be able to change unequal gender relations that are the result of a complex mix of historical, discursive and material, individual and institutional processes and actions, and that it is unavoidable that these intentional interventions will reinforce gender inequality as some point, while the option of doing nothing would only make things worse.

References

- Assemblée Nationale (2000), parliamentary documents, www.assemblée-nationale.fr
- Benschop, Y. and M. Brouns (2006), 'The Trouble with the Glass Ceiling. Critical reflections on a famous concept'.
- Biagi M. (2000), The Impact of the European Employment Strategy on Labor Law and Industrial relations in *International Journal of Labor Law and Industrial Relations*, vol 16, nr 2 pp. 155-173.
- Börzel, Tanja (2003), Guarding the Treaty : The Compliance Strategies of the European Commission, in T.A. Börzel and R.A. Cichowski (eds), *The State of the European Union* Vol. 6, 197–220.
- European Commission (2006) Report on equality between women and men, 2006, European Commission Brussels.
- European Commission (2005) Report on equality between women and men, 2005, European Commission Brussels.
- European Commission (2004) Report on equality between women and men, 2004, European Commission Brussels.
- Falkner, Gerda et al. (2005) *Complying with Europe. EU Harmonisation and Soft Law in the Member States*, Cambridge: Cambridge University Press.
- Fraser, Nancy (1989) *Unruly practices: power, discourse and gender in contemporary social theory*. Cambridge: Polity Press.
- Fraser, Nancy (1997) *Justice Interruptus*, London: Routledge.
- Giddens, Anthony (1984), *The constitution of society*.
- Héritier, Adrienne (2003) New Modes of Governance in Europe: Increasing Political Capacity and Policy Effectiveness?, in T.A. Börzel and R.A. Cichowski (eds), *The State of the European Union* Vol. 6, 105–126.
- Liebert, U. (ed.) (2003), *Gendering Europeanisation*, Brussels: Peter Lang.
- Lombardo, E. (2003) EU Gender Policy Trapped in the 'Wollstonecraft Dilemma', in *European Journal of Women's Studies*, Vol. 10, No. 2, 159-180
- Lilja Mósesdóttir (2006) Final Report. Research results of the project: From welfare to knowfare. A European Approach to Employment and Gender Mainstreaming in the Knowledge Based Society (see: www.bifrost.is/wellknow).
- Mosher J.S., Trubek D.M. (2003) Alternative Approaches to Governance in The EU: EU Social Policy And The European Employment Strategy, *Journal of Common Market Studies*, Volume 41, Number 1, pp. 63-68.
- Plantenga Janneke, Hugo Figueiredo, Chantal Remery, Mark Smith (2003), Towards an EU gender equality index. Feasibility study commissioned by and presented to the European Commission. Utrecht School of Economics/UMIST
<http://www.mbs.ac.uk/research/european-employment/projects/gender-social-inclusion/documents/EU%20index%20final%20report%20December%202003.pdf>

Rönblom, M (2004), *'The best place for women to be? A study of the construction of gender equality and gender mainstreaming in regional policies in Sweden and the EU'*, paper presented at the 2d Paneuropean Conference on EU Politics, Bologna 24-25 June 2004.

Rubery Jill and Colette Fagan (2000) Gender Impact Assessment and European Employment Policy. EGGE: European Commission.

Rubery, Jill (2003) Gender mainstreaming and the open method of coordination: is the open method too open for gender equality policy? Paper presented at Madison

Schimmelfennig, F. (2001), The Community Trap: Liberal Norms, Rhetorical Action and Eastern Enlargement of the European Union, in *International Organization* 55, 1, pp.47-80.

Scott, Joanne and David M. Trubek. 2002. "Mind the Gap: Law and New Approaches to Governance in the European Union." *European Law Journal*. 8(1): 1-18

Tallberg, J. (2002), 'Paths to Compliance. Enforcement, Management and the European union', *International Organization* 56, 3, 609-43.

Trubek David M. and James S. Mosher (2003) New Governance, Employment Policy and the European Social Model, in Zeitlin, Jonathan and Trubek, David (eds) *Governing Work and Welfare in a New Economy: European and American Experiments*. Oxford: Oxford University Press.

Trubek, David and Louise Trubek (2005) "Hard and Soft Law in the Construction of Social Europe: The Open Method of Coordination", 11 *European Law Journal* 343.

UNDP (200..) Country Fact Sheets The Netherlands,
http://hdr.undp.org/statistics/data/country_fact_sheets/cty_fs_NLD.html Accessed on 16 September 2006

Van der Vleuten, Anna (2005) Pincers and prestige: Explaining the Implementation of EU Gender Equality Legislation, *Comparative European Politics* 3, 464-88.

Van der Vleuten, Anna (forthcoming) *The Prize of Gender Equality. Member States and Governance in the European Union*. Ashgate: Aldershot.

Verloo, Mieke (2001) *Another Velvet Revolution? Gender mainstreaming and the politics of implementation*. IWM Working Paper No. 5/2001: Vienna.

Verloo, Mieke (2005) Displacement and Empowerment: Reflections on the Council of Europe approach to gender mainstreaming and gender equality, *Social Politics*, 12, 3, 344-366.

de Vries Michiel S. (2005) Generations of interactive policy-making in the Netherlands in *International Review of Administrative Sciences*, 71: 577 - 591.

ANNEX

Different rankings of EU member states (from Plantenga et al 2003).

The EO index is based on the indicators used in the JER Reports, and focuses only on paid labour.

C EO indicators

Table C.1 Actual values of the EO indicators and country ranking on the basis of EO composite index score, 2001¹

Country	EO1	EO2	EO3	EO4	EO5	EO6	EO7	Composite index score
Portugal	1.9	-15.8	26.5	21.5	94.1	1.08	7.6	0.56
Belgium	1.4	-18.7	26.1	18.1	92.7	1.15	11.4	0.53
Denmark	1.1	-8.3	28.1	19	89.6	<i>1.36</i>	<i>21.9</i>	0.46
Sweden	-0.3	-2.9	28	21.2	83.8	<i>1.36</i>	<i>21.9</i>	0.45
Netherlands	1.1	-17.5	25	18.1	78.9	1.23	16.3	0.43
Austria	1.3	-16.5	27.2	20.3	78.9	1.18	13.1	0.43
France	3.4	-14.1	26.6	17.4	89.2	1.34	21.2	0.41
Luxembourg	0.7	-23.9	26.8	19.6	83.8	1.26	16.4	0.41
Finland	1.1	-5.5	29.6	21.9	83.8	<i>1.36</i>	<i>21.9</i>	0.41
Italy	5.6	-27.5	21.9	15.2	91.4	1.3	18.8	0.38
Ireland	-0.2	-21.4	26.8	20.7	80.2	<i>1.36</i>	<i>21.9</i>	0.35
Germany	0.4	-13.7	27	18.2	80.6	1.53	29.5	0.33
UK	-1.1	-13.3	26.7	18.8	75.7	1.47	27.4	0.31
Greece	8.6	-29.8	21.7	15.5	86.8	1.29	19.4	0.30
Spain	8	-28.9	24.9	19.3	85.7	1.41	22.6	0.23

¹ numbers in italic refer to EU average

Source: Joint Employment Report 2002 (compendium-JER2002-updated-1121.xls)

The following ranking is based on the new index proposed by Plantenga et al (2003). This index is based on five dimensions that together should cover the relevant aspects of civil life, namely equal sharing of paid work, money, decision-making power, knowledge and time. In addition, each dimension is further specified in two sub-dimensions, as listed below.

Equal sharing of paid work: Labour force participation; Unemployment

Equal sharing of money: Pay; Income

Equal sharing of decision-making power: Political power; Socio-economic power

Equal sharing of knowledge: Participation in education and training; Educational attainment

Equal sharing of unpaid time: Caring time; Leisure

Table 4.6 The overall ranking of countries based on the gender equality index

	Participation	Unemployment	Pay	Income	Political power	Socio-economic power	Participation in education and training	Educational attainment	Care activities	Leisure	Composite index score
Sweden	0.90	0.94	0.42	0.75	0.86	0.40	0.57	0.77	0.55	0.61	0.68
Denmark	0.77	0.92	0.46	0.45	0.73	0.10	0.73	0.86	0.89	0.40	0.63
Finland	0.85	0.91	0.31	0.56	0.70	0.32	0.57	0.73	0.88	0.35	0.62
Belgium	0.49	0.87	0.50	0.44	0.41	0.47	0.92	0.97	0.57	0.38	0.60
Netherlands	0.50	0.93	0.29	0.95	0.64	0.27	0.85	0.68	0.69	0.23	0.60
Germany	0.60	0.99	0.15	0.86	0.58	0.29	0.91	0.53	0.55	0.23	0.57
Portugal	0.55	0.84	0.68	0.47	0.34	0.43	0.93	0.74	0.37	0.23	0.56
France	0.61	0.73	0.50	0.68	0.16	0.55	0.95	0.77	0.57	0.02	0.55
Luxembourg	0.31	0.92	0.42	0.75	0.26	0.38	0.88	0.45	0.55	0.23	0.52
Ireland	0.39	0.98	0.36	0.42	0.18	0.57	0.99	0.67	0.27	0.23	0.51
Italy	0.21	0.53	0.77	0.60	0.14	0.04	0.97	0.99	0.48	0.23	0.50
Spain	0.16	0.34	0.43	0.40	0.52	0.48	0.89	1.00	0.25	0.23	0.47
UK	0.62	0.91	0.13	0.27	0.30	0.39	0.24	0.72	0.55	0.34	0.45
Austria	0.55	0.92	0.22	0.08	0.48	0.39	0.77	0.33	0.37	0.23	0.43
Greece	0.15	0.28	0.42	0.62	0.03	0.24	0.97	0.91	0.18	0.23	0.40