International Comparisons in Science Studies: What and Why Do We Compare?[1]

Anne-Sophie Godfroy

Abstract

This article is based on experiences with EU-funded research projects over the last ten years. They have posed many methodological difficulties: many classifications do not fit the objects and actual issues are not measured or observed. They question the effectiveness of comparative research in the social sciences as it is being conducted now. We need new methodologies to make comparisons without overlooking specific contexts and effective tools to analyze large amounts of data, taking into account translation problems and mixing various types of data: qualitative, quantitative, more or less certain, first-hand and second-hand, statistical, from case studies, etc.

Furthermore, international comparisons assume that we can identify good practice, common indicators, etc. Scientific governance is based on such assumptions, but what happens to that model if the comparative methodology is not so relevant? Rethinking methodology for SSH in science studies could lead to rethinking the governance of science itself.

Keywords: *comparative research, international comparisons, methodology, research policy, European research, Social Sciences and Humanities*

Introduction

This article is based on the experience of research in EU-funded projects over the last ten years. These projects related to gender and engineering education: INDECS, WOMENG, HELENA; gender and science images at secondary school: MOTIVATION; and women in SET professions: WOMENG, PROMETEA.

The projects had a common topic: gender and SET (Science Engineering & Technology) and a common aim: to compare national settings, to develop a better understanding and to identify good practice.

They had common methodological issues: to collect and combine various levels of data, qualitative and quantitative; to analyze the data; and to propose effective measures to policy makers.

The usual methodological issues posed by international cross-comparative research were addressed in these projects. The main issues were:

(1) The cultural and linguistic issues posed by any international and interdisciplinary project. It requires specific guidelines for translation in order to avoid misunderstandings and bias.

(2) The specific issues of cross-comparisons, either a variable-oriented approach, or a case-oriented approach (Ragin 1987). Both approaches were used in complementary ways, with the quantitative approach for defining case studies and the case studies for better understanding the causal relations behind quantitative data.

(3) The standard problems of research design in human and social sciences: first, the choice of qualitative, quantitative or mixed methodology, according to the requirements of the research topic; mixed methodologies were experimented with, allowing the use of various collections of data.

(4) Very important was the issue of scale, as international comparative research generates a huge amount of data. If there are seven countries in the project, data is multiplied by seven. This leads to a paradoxical situation: on the one hand, the number of case studies may be too limited to go beyond exploratory research, but, on the other, there is too much data, so a large quantity of that data is not fully examined in the final analysis due to a lack of tools.

We also faced methodological issues specific to these projects: definitions of 'SET' and 'gender' and definitions of research activity and research jobs.

Underlying the projects was the assumption that engineering & technology have a specific culture, explaining that there are fewer women in technology than in science, which is true considering the statistical data we have at the European level. Intuitive perception seems to con-

firm the assumption. Projects were about defining and changing this SET culture. But if the separation between science and SET is clear in statistical data, drawing the border between engineering & technology and science is far from being clear in the field.

Within SET, multiple classification problems emerge, as SET are not defined the same way in different countries. Architecture, for example, may be considered engineering (thus part of SET) or art, agronomy may belong to science or to engineering, ICT may be grouped under mathematics or technology, etc.

Experiences from EU-funded projects

If we take a look at our past projects, we experienced various situations.

In the INDECS project, the first one in 1991–1992, we experienced a lack of harmonization mostly due to the underestimation of cross-cultural issues concerning research practices from one country to another and from one discipline to another. We planned 'questionnaires'. This was interpreted in various ways, from the full transcription of open interviews to lists of multiple-choice questions. Translation issues were not addressed specifically, leading to some strange results: Greek male engineers were 'emotional', because 'emotional' had been understood to mean 'with a strong character'. 'Critical thinking' was considered very positive in some countries, very negative in others. At the end of the project, we decided to write a 'synthesis', which was interpreted in different ways, like the rest of the guidelines of INDECS.

The pitfalls of comparative research as experienced in INDECS led to the design of a framework for comparative research developed in the following projects, WOMENG and PROMETEA. The methodology consisted of a four-step 'toolbox' to achieve harmonization and propose comparison (Godfroy & Pinault 2006):

– The first step consisted of designing common research tools. Collaboration is essential to ensure that all questions are meaningful in all contexts and no key-issue is omitted. Questions must be understood the same way.

– The second step consisted of defining common detailed procedures for documented fieldwork.
– The third step was reporting to a common database. It allows direct access to the data and common presentation of the information as well as practical tools for browsing the results.
– The fourth step was analyzing and interpreting. Even if the database provides tools for retrieving the data, there is still a need for tools to analyze huge quantities of data and better methodologies to identify the most interesting elements for comparison. An iterative mixed methodology was experimented with, but only a small quantity of the data was used in the analysis. This step would require more research and experiments with CAQDAS[2] software.

Thanks to the toolbox, results in terms of methodological coherence were very satisfactory in PROMETEA and WOMENG, but, at the same time, the outcome of the research was very disappointing. If comparisons were properly conducted, most of the results were already known from previous literature. Not enough attention was paid to each specific national case due to the obligation to compare more or less similar situations, and large quantities of data remained unused for analysis. After INDECS, on the one hand, and PROMETEA and WOMENG, on the other, it was obvious that a balance had to be found between harmonization and attention to specificities.

The MOTIVATION project allowed more methodological freedom in order to pay more attention to specificities. Harmonization procedures were much easier. This led to very interesting national studies, but inadequate comparative analysis, due to the lack of harmonization, and therefore poor comparative added value.

The current project, HELENA, will try to achieve the right balance and survey tools in order to better explore the data collected and identify meaningful comparisons. Testing possible treatments with CAQDAS software nVivo is under way. Coping with translation issues is at the heart of this research project.

Comparative methodology issues

The issue of scale is not often focused on, even if it is one of the most important challenges we face in EU projects. Literature usually addresses comparison issues going beyond three or four terms (Lallement & Spurk 2003). Some authors, as Janet Harkness, have studied large-scale cross-comparisons, but under a quantitative aspect and the use of multiple choice questionnaires (Harkness, Van de Vivjer, & Mohler 2002). There is no methodology for large scale qualitative comparisons dealing with semi-structured interviews or document analysis.

Beyond the existing statistical data, which is easy to manage, an EU project with six to fifteen countries involved generates tons of data, so that simply browsing results is difficult. There is a multiplicative effect: data in seven countries means seven times the amount of data. In WOMENG, for example, we had about 1,400 questionnaires (200x7) with 600 items and 4 different samples, plus 700 to 1,000 pages of English interview summaries.

In addition to the huge amount of data, different persons in different contexts and languages did the fieldwork. This meant that there was no direct access to the data, except for the team which had collected the data.

Data was only available through many filters: translation, poor knowledge of national settings, although we had procedures for reporting and providing background information on the context.

Access also depends on the technical tools required for browsing the database, especially for transversal reading (e.g. the same question in different settings). If the database has not been designed to cross the given parameters to retrieve the data, it is almost impossible to do it manually, because of the time it would involve.

Timing and expansion are another challenge in such projects. Not everybody is available at all times of the year. We need to identify the *kairos* in each country. This is especially true of studies in an academic context: students are not available during periods of examinations, training, and holidays. If the right moment is missed, the entire project is delayed, because all the results are connected. Finding equivalent settings from one country to another is a further challenge.

Collegial research design is an obligation in order to be sure that the research makes sense in different contexts, but it leads to expansion. Additions lead to expansion. The use of common instruments increases the length of each instrument. An effort has to be made to be less cumulative and more integrated.

If the 'toolbox' (Godfroy & Pinault 2006) proposes satisfying solutions to the problems already mentioned, the difficulties of analysis remain challenging. Beyond the technical difficulty of browsing the data and the fact there is no immediate, intuitive overview of results, we face an epistemological problem: which comparisons are meaningful? How can they be identified?

In our experience, the comparative methodology mixed quantitative and qualitative approaches. Mixed methodologies are still emerging (Cresswell 2003). This methodology was an opportunity to use all data available, even if heterogeneous, as we worked with national or European statistical data, quantitative data from the project (data from case studies or specifically collected), and qualitative data (interviews, focus groups). The idea was to maximize the potential of iterative mixed methodologies, allowing the concurrent progress of qualitative and quantitative strategies in an iterative approach.

The following example illustrates the methodology. We have 3 levels of data collection:

– *Set 1*: an overall statistical framework built from existing gendered statistical data and harmonized in international classifications (e.g. data from *She Figures* or Eurostat).
– *Set 2*: specific quantitative data collected by the project on specific issues (e.g. results from a multiple choice questionnaire on a large scale, data collected from national statistics, data from a given institution, etc.).
– *Set 3*: qualitative data from interviews, focus groups, participant observation and document analysis.

The basic idea is to address the same issues through the different sets of data and to move from one level to another in an iterative way. If there

is a correlation between two parameters in set 1, the idea is to explore set 3 to find what could be the causal relation between the parameters. Conversely, if a specific causal relation is noticed in interviews, comparison with set 1 and set 2 tests the hypothesis. The hypothesis can be confirmed. If not, we conclude that we have a very specific case study, one which is not representative of the general pattern.

This kind of methodology is an attempt to go beyond the case-oriented and variable approach. There are two main traditions in cross comparison (Ragin 1987):

– On the one hand, a variable-oriented tradition, based on a quantitative approach. It neglects social bases and empirical phenomena, but provides easy large-scale comparison.
– On the other hand, a case-oriented tradition. It is very sensitive to complexity and specificity, but almost impossible to extend to a large-scale comparison because of complexity and cultural diversity.

Ragin's recommendation is to go beyond qualitative and quantitative, to 'formalize qualitative comparative methods', and to examine 'constellations, configurations and conjunctures'.

The two methodological traditions identified by Ragin correspond in some way to the two well-established traditions of comparison: OECD benchmarking and comparative literature.

– The OECD has a benchmarking approach. Quantitative data is collected from a large number of cases or exhaustively. Indicators are formed from this data and used for benchmarking, a tool created for business and marketing. It was experimented with at Xerox at the beginning of the 1980s (Bruno 2008). From business, it was imported as a tool for EU policy or OECD policy. Many research projects are based on the assumptions produced by a benchmarking approach to statistical data. These may be the percentage of women researchers, the percentage of GDP invested in research, etc. All countries are supposed to compete with one another and to try to achieve the norm of the 'best', without any in-depth consideration of the context or the definitions behind the numbers.

– Comparative literature has developed methodologies which are mostly qualitative and sensitive to language and cultural issues (Chevrel 1989). But they are complex and difficult to extend to large comparisons with ten or twenty terms.

The challenge in EU research would be to find a methodology as sensitive to cultural issues and qualitative aspects as comparative literature, together with the straightforwardness of comparative benchmarking and indicators.

What and why do we compare?

What do we compare?

A first question about comparisons should be 'what do we compare?' We know that cultural contexts are different and that we never compare similar objects from one country to another. We compare oranges with tangerines, not oranges with oranges (Godfroy & Pinault 2006). Beyond the unavoidable cultural differences, classifications themselves may be questionable as categories for accurately describing facts.

If we go back to our experience in European projects relating to gender and scientific research, let us take the classification 'researcher' as an example. What is a researcher? In this category, we find persons who work in the research sector or persons who do research. Are technicians preparing experiments researchers? Are research managers or deans researchers because they used to be researchers? In industrial research, is an R&D department a synonym for a research department? Is a doctoral student a researcher or is it necessary to have a Ph.D. to be considered a researcher? Should we consider an academic who is an associate professor a researcher or a part-time researcher as he/she also teaches?

Sometimes existing quantitative data mix different definitions. This is the case in the *Livre blanc 2004. Les Femmes dans la recherche privée en France*, a report published in France, where only permanent researchers holding a Ph.D. are counted as researchers in some tables, and all people working in the research sector are considered researchers in other tables

from the same publication. This is not because of a lack of awareness on the part of the authors of the brochure, who are very professional statisticians. This is because a flexible definition is the only way to address the questions which this study sought to answer. A fixed definition would produce an irrelevant picture of the research field.

'Research' itself raises the same kind of issues about its own definition despite the fact that there is a lot of quantitative data about research. Research combines research proper, research management and administration, and sometimes teaching. Classifications use positions to define research activities: 'A' stands for full professors or equivalents, 'B' for associate professors or equivalents, 'C' for post-doctoral fellows, 'D' for doctoral students. But positions and activities are not always correlated. A full professor may be a research manager, a dean or a researcher. On the other hand, where are technicians or engineers in this classification, although they play an important role in laboratories?

Classifications by qualifications are not always relevant either. In the PROMETEA project on women in engineering and technology research, defining who the women researchers were in engineering and technology research was tricky. Should we consider women with a Ph.D. in engineering and/or sciences or women with an engineering degree? From one country to another, some definitions were applicable, some not. Flexibility was essential for identifying the right persons.

The category 'engineering & technology' is itself not so well defined. Classifications are not the same in different countries: ICT may be classified with mathematics; agronomy is sometimes considered an engineering discipline, sometimes not. The same is true of architecture. Many classifications present physics and engineering in one group. In that case, how can we separate them? Classifications may even vary within the same country, or there may be no classification at all, or the classifications are too broad. Where are the borders between engineering and technology or between E&T and science?

Interviews mean identifying a field and comparing this specific field with an average situation. In our projects, it meant identifying a research laboratory in engineering and technology and interviewing the members. The same laboratory might belong to many disciplines either due

to multidisciplinary research interests or many research interests; or because people have different academic backgrounds, sometimes many. The common idea is to choose a 'clear-cut' case, one clearly embedded in a single discipline, not a borderline one. Such cases are not easy to find, and do not represent the usual situation. This idea is practically inapplicable.

If we take a laboratory we studied in PROMETEA, the main research interest was in photonic devices and technologies, materials and nanostructures for photonics, nonlinear nanophotonics and biophotonics, and quantum nanophotonics. The first two items are perhaps more technological because they contain the words 'technologies' and 'device', but 'quantum nanophotonics' evokes basic scientific research. If we look at the current research projects in this lab, we find the same diversity: polymer-based devices for optical telecommunications; laser microcavities; hybrid nanostructures for laser emission; material engineering for photonics; photo-induced orientation in molecular media; nonlinear nanophotonics and nonlinear bio-imaging; coherent control of molecular motions; quantum nanophotonics. Some of the projects are easily classified under engineering & technology research such as optical telecommunications, laser emission, material engineering. On the other hand, nonlinear nanophotonics, molecular motions or quantum photonics belong to basic scientific research. In this case, how can we describe the activity of the laboratory in the usual categories? Is it science or is it technology, and in which disciplines? Officially, the laboratory is classified neither under nanotechnology nor under photonics; it belongs to the discipline '*sciences et technologies de l'information et des communications*'.[3] This classification may be explained by the history of the laboratory, but does not correctly describe the actual research activity. If we have a closer look, we notice a close collaboration inside the same institute with laboratories for chemistry, biology and applied physics on molecular photonics applied to biotechnologies and telecommunications. The institute was financed by an agreement between three partners: a governmental research agency, an institution of higher education and a telecommunications company. Even there, the borders between governmental, industrial and higher educational research were completely blurred.

Researchers perceive and describe themselves under multiple identities. Sometimes, they refer to their academic backgrounds, sometimes to their actual research topics. The identity of engineer is privileged in industry; the identity of researcher is privileged in governmental and academic research, even if people have exactly the same degrees.

Studying industrial research presents tricky methodological issues. The classification of industrial activities is not relevant either. Classifications by sectors (NACE)[4] are not very useful for studying research in industry. The existing classifications are product-oriented, not sensitive to research in many cases. If a company produces textiles, it may be manufacturing classical cotton or wool fabric or it may be a high-tech firm producing Kevlar or other sophisticated synthetic materials.

A company may have different activities and different locations in different countries. Therefore, the corporate culture is sometimes stronger than the sectored or national cultures. In addition, there is no centralized data collection for the industrial research sector. Furthermore, if we consider R&D, the borderline between research and development is impossible to draw. Companies usually have no data separating the two activities, which are intertwined. Finally, it is always possible empirically to study one company, but very hard to provide national or sectored pictures or to compare them. The connection between the case study and the overall data is very complicated.

These projects were based on an apparently 'simple' assumption: a lower number of women researchers in engineering and technology than in natural science in the existing statistical data. Even in such a 'simple' topic, we experienced a need to form new classifications in order to describe more accurately what we wanted to study.

Classifications are not relevant because activities change. This is the first reason to change them.

In addition, they were created before the development of comparative approaches. Amending national classifications to adapt them to a common international definition produces international classifications. This is the way EUROSTAT works. New classifications could be created in the perspective of comparative research without neglecting local cultures.

Anyway, this process could have drawbacks. If classifications are changed, this will create a gap in statistical data, preventing any approach mixing old and new data. It happened in Eastern European countries when they switched from a Soviet-friendly statistical system to the European Union statistical standards. There is a need to balance the benefits of new and better-adapted tools for describing and comparing the real world and the drawbacks.

Nonetheless, we must be aware that perfectly harmonized definitions would be useful, but would not solve all problems because SET is embedded in a cultural context that statistical data cannot take into account. Eurostat claims to compare 'apples with apples, not apples with pears' (presentation on Eurostat website). Despite constant and effective efforts to produce harmonized classifications, this is not exactly the situation we have experienced: even if categories are well defined, which is not yet the case, an engineer in Germany is not an engineer in France, even if they have the same degree in engineering, as defined by the Bologna process. The same observation can be applied to any object.

Why do we compare?

Besides the exact definition of the objects we try to compare, we must clarify why we compare. Comparing has become virtually an obligation for the majority of European-funded research.

Two reasons may be at the origin of this alleged obligation. First, many issues are the results of comparative benchmarking among European or international statistical data, so problems are formulated as comparative issues. This was the case with the proportion of women in engineering and technology, but this reasoning can also be applied to the proportion of poor people among the population, access to higher education, or other concerns.

Second, because partners come from various European countries, as is required in the composition of the consortium, these projects almost automatically generate data suitable for comparative analysis.

If comparison is assumed as a research strategy in most cases, the reasons for comparing and the aims of comparison may differ.

Benchmarking and identifying good practices are the usual reasons for comparing. The underlying theory comes from the OECD and the European Commission. For years, they developed statistical tools for harmonizing classifications and comparing the performance of different countries. On the Eurostat website, the institution's mission is clearly described:

> Eurostat's mission is to provide the European Union with a high-quality statistical information service. (…) Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions. This is a key task. Democratic societies do not function properly without a solid basis of reliable and objective statistics. On one hand, decision-makers at EU level, in Member States, in local government and in business need statistics to make those decisions. On the other hand, the public and media need statistics for an accurate picture of contemporary society and to evaluate the performance of politicians and others. Of course, national statistics are still important for national purposes in Member States whereas EU statistics are essential for decisions and evaluation at European level.[5] (…) The bottom line is we try to provide you with data that are comparable because apples have to be compared with apples – not with pears (…).[6]

From this perspective, translating the real world into numbers is a key task for providing policy makers and governments with appropriate monitoring tools. Defining the indicators is not only a technical issue, but also implies philosophical considerations about policy making. Moreover, replacing a description in words by a description in numbers creates a new social reality (Desrosières 2008, Chapitre 4: Pour une politique des outils du savoir: le cas de la statistique). It also produces new possibilities for comparison through the commensuration process (Espeland & Stevens 1998). New spaces for equivalence and comparison are created, where ranking and benchmarking become possible. The effects of such equivalence making may be the idea of equal opportunities between the different terms, as between gender or countries; it may also be competition, ranking and the requirement to achieve a given norm. Statistical data is therefore used as evidence and as an instrument of governance (Porter 1995). The development of 'OMC' or 'Open method of co-ordination' illustrates this attempt at governing through numbers. The aim of the OMC is to influence social policies in Europe by urging governments to improve their national ranking in tables of indicators (Bruno 2008).

Such a perspective must be challenged at different levels. At the level of the construction of classifications, translation into numbers does not construct a reflection of the world; it transforms the world and reconfigures it a different way. This process requires discussion and consensus on adopted conventions (Desrosières 2008). We have seen how existing classifications are not relevant for describing phenomena. At the level of policy making, it creates decisions to improve indicators and rankings, which does not always imply improving the experience under measurement. Examples are the guidance of the economy in order to increase GDP, which is supposed to be a synonym for increasing well-being and/or wealth, but this is not the case, as recent debates on this indicator have shown.[7] Another case is the ranking of American law schools studied by Espeland and Sauder (2007).

Other reasons for comparison have long traditions, but have not been applied systematically to comparative European research. It might be fruitful to try them out.

One of these traditions is comparative literature. The idea of comparative literature is to compare different objects to study degrees of similitude and reveal new conclusions, not apparent through singular analysis, especially concerning what is original in each object (Chevrel 1989). This approach reveals new hypotheses and has a heuristic value. It is sensitive to qualitative data and to alterity, as the indirect route through alterity is considered the best way to identify and understand local or national singularities. This approach is familiar with issues such as translation, cultural context, borders and fuzzy borders. Besides, it has developed a strong awareness of the transformation and circulation of ideas, reception issues, reconstruction in other contexts, etc. This knowledge could be very useful to study the circulation of social policies or good practice across the European space. Instead of creating competition and the obligation to comply with given norms, it creates a dialectic between the same and the other in order to better understand local issues and for mutual enrichment. In comparative literature, apart from the focus on local singularities, the comparative activity creates a 'world literature' or 'universal literature', distinguishing important works with an impact on the global history of literature. This theory allows imagining inter-

actions between cultural spheres without denying awareness to local complexity. Such a model could be very fruitful applied to social phenomena and should be explored methodically.

The work of Sheila Jasanoff in *Designs of Nature* (Jasanoff 2005) might be considered an example of methodology close to the comparative literature philosophy. She compares politics and policies of the life sciences (debates on GMO, cloning, stem cells, etc.) in the United States, Britain, Germany and the European Union. Her objective is 'to explain as fully as possible why new developments in the life sciences were differently received into three national political systems, and what the implications of these stories are for the future democratic control of biotechnology' (Jasanoff 2005, 39). Then 'prospects for supranational harmonisation, both at the EU level and globally, will be deduced to some extent from the comparison of national cases' (ibid.).

She studies the policies of the three states as three 'controlling narratives' and investigates 'why they were differently received and institutionalized within each national political system' (ibid). This methodology is very close to comparative approaches in literature, political studies or politics, as Jasanoff explains: 'The methods I adopt (…) owe much to the history and sociology of knowledge and the anthropology of technological cultures as they do to comparative politics, policy studies, or law' (ibid., 15).

She openly criticizes the usual comparative approaches to science policies and 'faith in the possibility of melioration through imitation' (ibid.) which is behind benchmarking and the replication of good practices. She proposes 'a new kind of comparative analysis' because 'with a growing awareness of the culturally embedded character of both knowledge and policy, there are reasons to be sceptical of the unproblematic learning from other's experience'. She calls for new methods and new conceptual approaches. The aim is to identify 'styles of regulation' embedded in specific cultural and political contexts.

> To understand how policy domains are carved out from the political sphere and rendered both comprehensible and manageable, we must employ analytic categories different from those of decision makers operating within the policy process. We need a conceptual language that can grapple with both continuity and change, while rejecting some of the rigidities of the structure. (Jasanoff 2005, 23)

Conclusion

From an analysis of our unsatisfactory experience with European research, from a methodological point of view, we conclude that there is a need for new comparative approaches to science policies or social policies at the European level. The Social Sciences and Humanities can contribute to this programme to a large extent. Comparative literature and comparative history methodologies might be fruitful inspiration, as Jasanoff's work has demonstrated.

In this article, the methodological issues have been listed that should be addressed in depth. Beyond the technical aspects, they carry new political and cultural assumptions, as the existing benchmarking and commensuration approaches are not neutral and objective, but transform social reality.

New classifications would be useful, as existing ones do not properly describe the objects we study. They should reveal new research questions, and, as policies are linked to existing comparative research, new classifications mean new policies.

There is also a need for appropriate tools to manage large-scale projects and large amounts of data. Appropriate databases and a better use of CAQDAS software should be investigated methodically. One of the limitations of this kind of software for comparisons is the translation issue. Experiments on ways to use CAQDAS when dealing with different languages could be planned, testing different options: CAQDAS after translation, CAQDAS before translation, identifications and correspondence between key words in different languages.

Mixed methodologies (Creswell 2003) could be very interesting, as they allow the exploitation of different collections of heterogeneous data, quantitative or qualitative. This is obviously a path to explore in order to bridge the gap between case studies and statistical data.

A crucial issue is the balance between harmonization and singularities. Changing perspective and taking inspiration from comparative literature could be a way of creating a better balance. The use of statistical data as a first inspiration tends to create norms and commensuration processes that erase specific local and cultural issues essential for understanding phenomena. Prevalent results remain, but such findings are already known

and not specific. Therefore, such outcomes are finally of little interest in terms of policy making, as overall perspectives do not guarantee that a given measure will be successful locally.

In addition to methodological issues, the aims of comparative research need to be clarified. If the Social Sciences and Humanities contribute to research for evidence-based policies, this issue has to be addressed seriously. What and why do we compare? Is legitimating numbers and benchmarking the aim of SSH or should they offer something different, such as methodologies for dealing with qualitative objects?

In the future, it is evident that SSH could play an important role in understanding and evaluating science policies, health policies, social policies, etc. This role is already visible in existing EU research, and it is growing: The 'Science in Society' chapter is now a visible issue in the framework programme, and more money is being spent on it, even if the amounts are still ridiculous compared with the money spent on the natural sciences. Topics addressed by SSH are essential for the future of Europe: gender issues, an ageing population, poverty, citizenship, etc. An 'SSH for science in use' must be developed, with expert results disseminated to policy makers. SSH can play this role, however, if they propose specific perspectives and new results different from the results expected from OECD statistical data.

A programme for SSH in the future could be summarized in two research interests: first, addressing epistemological issues, clarifying what and why we compare; second, renewing methodological tools and creating new comparative methodologies mixing qualitative and quantitative data and allowing studies on a very large corpus of data. After developing such a programme, SSH could contribute massively to the definition of European research and EU policies – including alternative EU policies – as pure statistical approaches and numbers have reached their limits.[8]

Notes

[1] This paper was originally published in *Innovation – The European Journal of Social Science Research*, volume 23, number 1 March 2010, p. 37–48, Routledge.

[2] Computer-Assisted Qualitative Data Analysis

[3] ICT

[4] Statistical classification of economic activities in the European Community. See http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2&StrLanguageCode=EN&IntPcKey=&StrLayoutCode=HIERARCHIC&CFID=2253871&CFTOKEN=13f9f99cb06dcc56-DA005703-B164-3CEA-024126280EA286D3&jsessionid=f90044688be36d6c4a62

[5] http://epp.eurostat.ec.europa.eu/portal/page/portal/about_eurostat/corporate/introduction

[6] http://epp.eurostat.ec.europa.eu/portal/page/portal/about_eurostat/corporate/introduction/harmonization

[7] A commission co-ordinated by J. P. Fitoussi, A. Sen, and J. Stiglitz has been commissioned by the French State to propose a better alternate indicator: International Commission on Measurement of Economic Performance and Social Progress. See http://www.stiglitz-sen-fitoussi.fr/.

[8] Opposition to the Bologna process and governance through quantitative evaluation is growing. *The Black Book of the Bologna Process*, published by students unions, was a starting point of the movement, which is spreading across Europe. Similar concerns have been expressed in other social settings.

References

Bruno, I. (2008), *A vos marques, prêts, cherchez! La stratégie européenne de Lisbonne, vers un marché de la recherche*, Bellecombe-en-Bauges: Editions du Croquant.

Chevrel, Y. (1989), *La littérature comparée*, Paris: PUF.

Creswell, J. (2003), *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, London: Sage.

Desrosières, A. (2008), *L'argument statistique I. Pour une sociologie historique de la quantification*, Paris: Presses de l'Ecole des Mines.

Espeland, W. and Stevens, M. (1998), 'Commensuration as a social process', *Annual Review of Sociology* 24: 313–343.

Espeland, W. and Sauder, M. (2007), 'Ranking and reactivity: How public measures recreate social worlds', *American Journal of Sociology* 113 (1): 1–40.

Godfroy-Genin, A. S. and Pinault, C. (2006), 'The benefits of comparing grapefruits and tangerines', *European Journal of Engineering Education* 31 (1): 23–33.

Godfroy-Genin, A. S. and Sagebiel, F. (2007), 'Möglichkeiten und Schwierigkeiten internationaler und multimethodologischer Forschung über Gender in den Ingenieurwissenschaften', *Zeitschrift für Frauenforschung und Geschlechterstudien* 25 (2): 27–43.

Godfroy-Genin, A. S. (2009), 'Women's academic careers in technology, a comparative European perspective', *Equal Opportunities International* 28 (1): 80–97.

Harkness, J., Van de Vivjer, F. J. R., and Mohler, P. (Eds.) (2002), *Cross-Cultural Survey Methods*, Hoboken: John Wiley and Sons.

Jasanoff, S. (2005), *Designs on Nature*, Princeton: Princeton University Press.

Lallement M. and Spurk J. (Eds.) (2003), *Stratégies de la comparaison internationale*, Paris: Editions du CNRS.

Ministère délégué à la Recherche et aux Nouvelles Technologies (Ed.) (2004), *Livre blanc 2004. Les femmes dans la recherche privée en France*, Paris: La documentation française. http://lesrapports.ladocumentationfrancaise.fr/BRP/044000105/0000.pdf.

National Unions of Students in Europe (2005), *The Black Book of the Bologna Process*, http://www.bologna-bergen2005.no/Docs/02-ESIB/0505_ESIB_blackbook.pdf.

Porter, T. (1995), *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*, Princeton: Princeton University Press.

Ragin, C. (1987), *The Comparative Method, Moving Beyond Qualitative and Quantitative Method*, Berkeley / Los Angeles: University of California Press.

Ragin, C. and Becker H. (1992), *What is a Case?*, Cambridge: Cambridge University Press.

Sagebiel, F. (2005), 'Using a mixed international comparable methodological approach in a European commission project on gender and engineering', in Hoffmeyer-Zlotnik, J. and Harkness, J. (Eds.), *Methodological Aspects in Cross-National Research*, Mannheim: ZUMA-Nachrichten spezial 10, 47–64.