High Educ (2014) 67:409–422
DOI 10.1007/s10734-013-9680-2

Challenges in higher education research: the use of quantitative tools in comparative analyses

Emanuela Reale

Published online: 9 November 2013

Abstract Despite the value of the comparative perspective for the study of higher education is widely recognised, there is little consensus about the specific methodological approaches. Quantitative tools outlined their relevance for addressing comparative analyses since they are supposed to reducing the complexity, finding out and graduating similarities and differentiations, and improving the generalization of the results. Nevertheless, the limited availability of comparative data hampers statistical comparative analyses, and data and statistics might be seen just as complementary analytical tools. This paper has a twofold aim. Firstly, it addresses key methodological problems for accurate international comparisons in higher education studies. Secondly, the article discusses the difficulties encountered in using measurements in comparative studies, and issues to be addressed in order to improve the robustness of the method and the possibility of using quantitative tools.

Keywords Comparative research · Qualitative analysis · Quantitative analysis · Statistics

In a seminal work published in 1996, Teichler explained that the notion of comparative higher education (HE) research means dealing with "research addressing phenomena of higher education in more than one "culture", "society" or "nation" systematically or in a single one in comparative perspective". It pursues the typical logic of comparison which is universal for research striving to identify common elements and differences as well as to test hypotheses on causal relations." (Teichler 1996, 448–449).

The comparative approach is seen as one of the most fruitful in higher education studies, since it allows researchers to broaden their observation base and to achieve a more extensive and reliable understanding of the phenomena observed. Nevertheless, this rec-

E. Reale (✉)
Institute for Research on Firm and Growth CERIS CNR, Via dei Taurini, 19, 00185 Rome, Italy
e-mail: e.reale@ceris.cnr.it

ognition is sometimes combined with a poor understanding of the methodological needs of comparative analysis, as well as with confusion about the difference between "methods" and "techniques", between concepts and statistics.

This paper has a twofold aim. Firstly, it addresses key methodological problems for accurate international comparisons in higher education studies. Secondly, the article discusses the difficulties encountered in using measurements in comparative studies, and issues to be addressed in order to improve the robustness of the method and the possibility of using quantitative tools. The questions to deal with are: what is the meaning of comparative research and what are the advantages of adopting a comparative approach? What are the peculiarities of comparative research in the higher education field? How can methodologies for comparative research be improved using measurements? And finally, how and under what conditions are indicators useful for comparative analyses?

The paper is organised as follows. The first section discusses the methodological choices researchers have to make when using a comparative approach. Section "Peculiarities of comparative higher education research" presents the most relevant limitations and peculiarities of comparison in the field. Sections "Improving the quality of comparative research: the emerging use of quantitative tools" and "Data and indicators for studying higher education" deal with the problems of quantitative tools, and the use of data and indicators for analysing higher education systems and organizations. The last section outlines some recommendations and concluding remarks.

Comparative research and methodological choices

Comparative research is an unavoidable mean for researchers, and for researchers in higher education too. Nowadays it is difficult to find anyone who disagrees with the above statement. Nevertheless, the meaning and the consequences of doing comparative research still need to be discussed in order to improve the soundness of the methodological approach of the research work.

Challenges in comparative research

Discussing the challenges of comparative higher education research is not a simple task, since one cannot consider the problems encountered to be general issues. In several cases, specific problems in disciplines belonging to the positivist tradition of Western science might be practically irrelevant in other intellectual traditions, such as the humanistic ones. In the former case, presenting and testing hypotheses in order to provide causal explanations is a crucial matter, but it is not necessarily the same in the latter case (Valimaa 2009).

The positivist tradition of Western science sees scientific knowledge as the sole valid knowledge, which can be obtained through the verification of general laws using empirical evidence. Like the physical world, also society operates according to some general laws, which cause the effects we can see or perceive. Hence, objectivity, rationalism, and causality are the means to investigate social facts. Main criticisms to this approach have come from several scholars in the social and human sciences, who argue that research on society should focus on human and cultural norms, values, symbols, and social processes. Trying to investigate these kinds of social phenomena to identify unique and precise causal relationships among facts is almost impossible, because said relationships work at a given time, in a given historical context and geographical location, and, unlike what happens

with the natural sciences, they are neither invariable not generalizable (Smelser 1976; Ragin and Zaret 1983). The "revolution" (Verba 1967) in comparative research initiated in the 1960s is mainly grounded in the positivist tradition, although different perspectives and views exist, as well as differences and fierce debates among contrasting approaches (e.g. the generalised law approach vs. the configurative-idiographic approach, Verba 1967), which also implies a different focus on the use and importance of quantitative and qualitative evidence.

Interestingly enough, the existence of causal inferences is assumed in comparative research from the very beginning: "the whole idea of comparative research emanates from a clearly quantitative tradition, the rules for experiments in science" (Allardt 1990, 189). This assumption has given rise to what has been termed a 'deterministic quantitative approach'. However, it is acknowledged that information which is unique and characteristic of the observed variables cannot be eliminated from the analysis: "whatever theoretical notions we have about societies, they are products of history and they contain many unique patterns, and memories. These unique occurrences… cannot be unveiled except by qualitative research directed at the singular and unique" (Allardt 1990, 189). Moreover, it is generally recognised that collecting data and testing causal hypotheses in order to develop generalisations is not an easy task, and the final result can often be either a generalisation removed from reality or a set of explanations tailored to each case (Verba 1967).

In recent times, the positivist approach has enjoyed growing popularity due to a general tendency in the social sciences to borrow modes of knowledge production from the natural sciences and to develop general, theoretical works. Quantitative researchers have been focusing on nomothetic reasoning, in order to capture similarities and repetitions of phenomena across time and to generalise findings to a large number of cases belonging to different systems. It is complemented by the heavier use of quantitative tools for cross-country and/or cross-disciplinary research in order to maximise the variation of the explanatory variables or to verify some hypotheses based on a general theory (Fairbrother 2007). This tendency also affects comparative studies about HE, and it is linked to the request coming from the State for measurements, accountability and causal explanations, which can support evidence-based policy-making (van der Ploeg and Veugelers 2008).

In this paper, we deal with the challenges of comparative research on higher education, which are related to the positivist perspective of Western sciences and the prevailing idea that any kind of causal inference or descriptive analysis cannot be performed without comparing empirical evidence emerging from different cases, providing concrete recommendations and suggestions to policy makers and managers. Comparisons are supposed to allow scholars to find robust explanations and answer research questions, and policy makers to have empirical evidences to justify the reforms or managers to drive changes of the organizations. In the last few decades, comparative research on higher education has become increasingly focused on comparing cases from different countries, due to the growing trans-national dimension of higher education policies, and the emergence of policy objectives, like the European integration, which affect the strategies, objectives, governance and management of the higher education institutions (HEIs). Nevertheless, several constraints and methodological problems have emerged as well.

Methodologies in comparative research

When one deals with comparative research, it is important to identify the methodology, which best suits the achievement of the declared objective. Yet, a preliminary question

ought to be asked: what is a methodology? In a seminal work dating back to 1970, Sartori answers that a methodology is the logical structure and procedure followed in research activities; it differs from techniques, which refer to surveys and statistics used for the empirical analysis. "If one does not make this distinction, one may be a fantastic manipulator of data but will remain an unconscious thinker, a researcher affected by methodological unawareness." (Sartori 1970, 1033). On the contrary, a "conscious thinker" is a researcher aware of the limitations of his/her findings, although he/she knows that said findings can be managed successfully in spite of existing limitations.

Sartori also distinguishes between "conceptual stretching" (vague conceptualisations useful to extend coverage but lacking precision "saying less in a far less precise manner") and the need to use "empirical universal" concepts, which can be applied at any time and in any place and can be empirically tested or represented. Conscious thinkers question the reasons for comparing and their main focus is not the extension of knowledge but rather how to create and validate new knowledge. In this sense, Sartori agrees with Lijphart's statement "to compare is to control, testing hypotheses and laws against as many cases as possible" (Lijphart 1971).

The comparative method implies the usual sequence of research steps: selecting precise research questions, identifying alternative explanatory hypotheses, developing the empirical testing, and searching for causal explanations.

During the 1960s and 1970s, the comparative method was investigated in depth, and several norms and practices were developed. The starting point was Lijphart's contribution, which sets the comparative method apart from the experimental one, the statistical one and case studies, highlighting its peculiarities, and posing the problem of "many variables, small N" (Lijphart 1971; Id., 1975). The main limitation is that there might be more complementary explanations to consider than cases to observe: "as the number of explanatory factors approaches the number of cases, the capacity to adjudicate among the explanations through statistical comparison rapidly diminishes" (Lijphart 1971, 686).

During the 1980s, other scholars tried to put forth methodological alternatives, and several innovations in the comparative method were presented and discussed (Collier 1993). The solution to the problem of a limited number of cases against many variables was found in increasing the number of cases, reducing the variables, and choosing comparable cases. The general approach is to compare a limited number of cases belonging to the same *genus*, thus comparability is a mode of investigation that uses taxonomies *per genus* (the class predicting similarities) *et differentiam* (the species of the genus or subspecies). Sartori (1970) proposed to replace the question "what do we compare?" with the question "how much do we compare?", addressing differences in degree (more or less) rather than differences concerning type (either-or). Adopting this perspective would make it possible to overcome the shortcomings of the "*per genus et differentiam*" method of analysis, as it would concentrate on quantifying degree. However, this does not solve the problem of what to compare. The dilemma is: if you select two identical cases, you have no clues for their comparison. If you select two items with nothing in common, again you have no room for comparability. Comparability really emerges when you have two or more items that are "similar enough"; adopting this perspective, "one can say that comparison is aimed at discovering fundamental similarities, which are hidden by secondary dimensions (making similar things look alike), but this operation too might be tricky (Sartori 1970, 1035, see Prospect 1).

At present, comparative research must meet several basic requirements related to its robustness, which are: (a) the research design should focus on concepts and classifications (not on mere descriptions); (b) the investigation should be based, at least in principle, on

Prospect 1 The 'what is' question in Sartori, 1970

What is?	Qualitative question
Either-or type of logic (concept)	Classification based on universally exclusive and jointly exhaustive categories (characteristics which an object either has or lacks)
More or less logic	Graduation applicable only to objects in the same class

testing causal hypotheses by means of various possible explanations; (c) cases can be selected on the basis of their great similarities or great differences, or because they are a representative sample; (d) it is best to deal with more observations but fewer variables, through the formulation of more targeted hypotheses. The presence of the abovementioned requirements is customised according to the results the scholars wish to achieve, which obviously change, and the field being investigated. The aim might be to discover the specificities of a system as far as the phenomenon analysed is concerned, or to detect convergences and differentiations across cases or categories assumed to be similar and unified. Comparisons might also be made for the purpose of formulating a theory and, in this case, they should, at least in principle, try to point out causal relationships generating similarities and differences among the observed cases, moving beyond the limited scope of the analysis at hand.

A final remark concerns the importance of historical reconstruction in comparative analysis, which sheds light on the origin of phenomena and on their patterns of change. Because of "the absence of enough independent, self-contained cases to be compared in order to identify causal patterns, the researcher is thus often left with a substitute, namely that of narrating a story instead." (Azarian 2011). Nevertheless, using history in comparative research does not mean becoming a "storyteller". On the contrary, the systematic analysis of key turning points in time and of their causes is a powerful means to understand the generative mechanisms underlying the evolution (changes, transformations) of the observed phenomena (Neave 1996; McMichael 1990). In this line, Valimaa (2009) recalled the Bereday's lesson of two main approaches to comparative studies, namely the area studies and the thematic studies, as well as the permanent tension between the theoretical ambition to find causal relationships "and the need to develop studies historically well-rooted" (Valimaa 2009, 142).

Several lines of comparative research have been applied to the study of higher education in order to identify the different elements making up HE systems or institutions. The objects of investigation are generally strongly linked to the most important problems the policies and the organizations addressed, looking for explanations, converging patterns, and performance. In this respect, it is useful to recall the normative dimension of comparative research in HE, "not only in the sense that an authority defines what is 'good' comparative research …. (but) because the idea of comparison easily turns into the idea of competition, when the research outcomes are interpreted to… the policy makers" (Valimaa 2009, 143).

Peculiarities of comparative higher education research

Despite being widely adopted by scholars, comparative research is affected by serious methodological limitations, which hinder the attainment of rigorous and explicative

results. Some limitations are closely linked to its method—like the issue of autonomy of the units chosen for comparison: "as various species of entities are selected for the aim of comparison, there is often an underlying and tacit assumption about their autonomy and a tendency to ignore the complex interplays and mutual influences among the units"(Azarian 2011, 120). Problems might also derive from an asymmetric understanding of the cases, since scholars sometimes have no direct experience of certain country-specific features, which are important for the meaningfulness of the analysis (Azarian 2011).

The above limitation is often overcome by bringing together scholars from different countries, in order to integrate the knowledge and various experiences they have had. This happens, for instance, in research projects, which receive some kind of financial support from non-national funding sources, first and foremost from European sources. However, working in international teams might bring about other problems—such as integrating different views, cognitive schemes, ideas, languages and meanings, which can influence the robustness of comparative work in relation to the interpretation of empirical results and the outcome of observations. The literature suggests principles to be adopted so that the validity of conclusions drawn from cross-national survey data is not undermined (Kuechler 1987) and indicates strategies to promote effective collaboration and conflict resolution in qualitative research efforts, in order to avoid any cognitive misunderstanding of the conceptualisation (Bagshaw et al. 2007).

Teichler (1996) outlines other specific features of research on higher education, which are likely to affect the comparative method when it is applied to this field. First of all, higher education is an object-focused area, much more than an analysis revolving around a single discipline. Secondly, studies on higher education are strongly driven by the social relevance of the themes addressed. Research in the HE field is linked to solving practical problems, and this determines shifts in focus depending on the relevance of the various themes within the social and political debate. Moreover, the importance of practitioners and of professional problem-solving as well as an incomplete knowledge of the field—which requires collaborative efforts in order to overcome knowledge asymmetries, make it necessary for research on HE to consider theories, paradigms, and methods from other disciplines and, at the same time, to use various "knowledge spheres" (quantitative and structural aspects, subject-related aspects, person-related, research-related and teaching-related aspects, governance and organisation, Teichler 1996). The mentioned features do not impact only on the comparative method, but more generally they affect all research methodologies in this field. In this sense, the remark that "comparative research does not differ in its logic from research undertaken within a country" (Teichler 1996, 448) is correct. At the same time, the peculiar features of higher education are shared by other sectors of the public administration and one can argue that, as a rule, the use of the comparative method has to be adapted to the specific characteristics of the field under investigation.

Also the role of hypotheses in HE comparative research is limited. There are no "sufficient regularities in social experience for them to be capable of being incorporated into 'overarching' frameworks and hypotheses" (Kogan 1996). In sum, causal explanations are the main objective, but many authors suggest that one does not necessarily need hypotheses for theorising (Page 1995).

Other two characteristics can be outlined, which contribute to the complexity of the concrete research work in the HE field, and are relevant when the use of quantitative tools is concerned:

- Comparative research in HE needs to take into account the different actors involved at system (the national and supranational state, intermediary layers, institutions) and institutional levels (stakeholders, government bodies, faculties, departments, groups, individuals), and their interactions occurring in multi-layered and multi-level policy space, in order to understand similarities and differences between the phenomena under analysis. This is extremely important in the new global environment in which the higher education systems and institutions are involved;
- Many of the observed variables in higher education research refer to individual beliefs, such as intentions, values, or cognitive components, rather than to objective facts. Moreover, the role of ideas shaping behaviours and driving institutional changes cannot be underestimate. Using them comparatively is possible, but it is crucial to consider the limitations deriving from the fact that scholars often ignore how said beliefs emerged and where values and cognitive frameworks come from (Kogan and Bleiklie 2006).

In sum, before deciding about the use of measurements, the "conscious scholar" has to ask general questions about what he/she is looking for when comparing HE systems or HEIs. Reflecting on this would allow to specifying the type of analytical endeavour to be developed, and the extent to which it is possible to develop formal hypotheses or the objective to pursue through a sound conceptualization. Afterword one can check how far the phenomenon under investigation can be effectively addressed using a quantitative approach, and what kind of limitations and caveats have to be recognised.

Improving the quality of comparative research: the emerging use of quantitative tools

The third question posed in this paper is: to what extent quantitative techniques might help to better address problems of comparative research in the HE field? This question is not asked in order to investigate structural and epistemological differences in methods affecting the research design, the object and aims of the analysis, and the data and technical tools used. Instead, the aim is to speculate about the possibility and the consequences of using measures in comparative research, and the extent to which such measures can improve the quality and robustness of results.

Quantitative tools and measures are becoming more relevant not only as an answer to the problem of "many variables, small N", though they can help to reduce it. The growing tendency towards measuring is linked to the neo-liberal assumption that 'we cannot manage what we cannot measure', and measuring is a way to make results more precise, general, robust, and objective. Moreover, measures and statistical analyses might support the possibility to highlight causal relationships among the phenomena observed and, as a consequence, they make it possible to theorise on the basis of hypotheses.

Nevertheless, in comparative research, the term 'quantification' is often applied to the attribution of numerical values to items, indicating rank order, differences or the distance between them. Sartori and other scholars recall that this is not properly "measurement", rather a method to transform qualitative elements into quantitative ones, in order to treat them using statistical tools. In fact, it is important to distinguish between measures used for calculations (addressing the formulation of ordinal scales aimed at demonstrating that A is double than B), and measures used for addressing the building of rough scores aimed at determining to what extent A is similar to or different from B. The former is a case of real "measurement", whereas in the latter case measures are used as tools for the numerical representation of qualitative characteristics (Sartori 1970).

The mentioned discussion let understand the importance of conceptual formulation, in order to avoid the risk of having treatments of qualitative items with statistical methods but no theoretical significance. When the comparative method coincides with the statistical treatment, one can have complicated models, which are supposed to demonstrate and predict the functioning of the systems or of the organization behaviour, whose final results are either trivial or meaningless, and in no way able to contribute to creating new knowledge. In this case, the statistical treatment becomes more important than the underlying concepts, and the relevant result is the functioning of the statistical tests, regardless of their explanatory capability or the actual findings of the research effort. Thus, independently of the control technique used, in comparative research the preliminary phase of concept formation is the most important because, "before measuring the objects under investigation in terms of some variable, one has to form the concept of that variable" (Lazarsfeld and Barton 1951; Sartori 1970).

Two examples might better illustrate the problems of quantitative tools, and the extent to which they contribute to solve, at least partly, the constraints of HE comparative research.

Classifications as tools for quantitative comparative analysis

Classifications are actually powerful tools for the creation of an order among several variables, which can be graduated using numerical values, and then treated with the support of statistical techniques. Classifications can show how an observed phenomenon varies among systems and among institutions (Van Vught et al. 2005), hence, their great relevance in comparative research for capturing the role of different actors at different levels.

Lepori and Bonaccorsi (2013) state that classifications are a method to combine extensive and precise information in order to make it comparable; the attribution of specific observations to one or another category is a way to simplify the description of some properties which different objects may have. Classifications can follow a binary approach (identifying conditions for attribution to one or another type), an approach based on similarities (i.e. a set of features usually characterising different objects), or a functional approach (categorisation based on the presence of types of activities performed). Whilst it is difficult to identify clear and objective criteria to adopt one of the above classifications, since none of them is truly effective when actually applied to empirical observations, multiple classifications might be needed simultaneously. Moreover, conflicts and controversies concerning whether the content is suitable for attributing a given status, or for representing power relations or distribution, might affect the legitimacy of research results. Here "a commonly used practice can be defined damage avoidance: two things are considered comparable if comparison does not generate harms, distortions, or unacceptable consequences, to (almost) all the actors involved" (Lepori and Bonaccorsi 2013).

Classifications are also steps towards the creation of typologies, which "represents a (generally implicit) statement or theory about how systems operate and about the trade-offs between different values. … Typologies must go beyond the classification stage to establish empirical propositions and explain or predict features of system behaviour that are logically independent of the type-defining characteristics." (Lange and Meadwell 1985, see also Manardi 1990). In a recent work about the typologies of higher education institutions, Van Vught and his colleagues also argue that typologies can be developed on an a priori (similarities and differences) or an a posteriori basis (effective data about

the performance of systems). They can focus on one dimension or several dimensions, can be hierarchical (tree-like) or non-hierarchical (rank based on the outcomes), and can be built using subjective data or objective evidence (Van Vught et al. 2005). In all the cases above, the ultimate goal is to have types or ideal types of systems, which allow researchers to formulate hypotheses and to test them against expected outcomes.

Actors' role and the measurements of logics and believes

Another important issue to be considered is at what level the quantitative tools for comparative research should be used—an issue related to the so-called macro–micro dilemma. The macro level deals with what governments do (or do not do), with explanations focusing on their relationships with other political entities and the external context. The micro level investigates how decisions are taken, what relationships exist among the actors involved, what logics of action are at play. Difficulties in using measurements for comparative studies at the micro level are generally linked to the following: (a) the interactions among the actors are strongly influenced by the context of action; (b) it is difficult to separate causes and effects of the phenomena observed; (c) many of the variables observed refer to individual beliefs (intentions, values, cognitive components) the origins of which are hard to understand.

Aldricht (2009) outlines various approaches used in comparing organisations, ranging from extreme de-contextualisation (organisations are all alike) to extreme contextualisation (organisations are all unique). Considering how environments shape organisations implies taking into account time (clock time and socially constructed time), space (geographical location, laws, rules, regimes, practices, resource constraints), and history (traditions and cognitive frames).

In comparative higher education research, the same problem arises. In practical terms, when researchers use, for instance, data from a survey, in order to understand roles, hierarchies, logics of actions, rationales and motivations, said data provide evidence on the perceptions of different actors operating in different contexts. It is then up to the researchers to eliminate the context effect, or retain it, when the assumption is that the context does not influence the specific features under investigation. In fact, several biases can occur: (1) measurements biases due to poor data collection or wrong procedures that pressures respondents for answering—with the possibility of having distortion in the answers, (2) sampling biases and systematic biases in the response to one specific question, (3) the absence of a sufficient level of agreement between respondents, and iv) the significance of the observed differences (van Mierlo et al. 2009).

Therefore, measurements can help the comparative method as they make for more robust explanations (when they are not merely descriptive), but this is true only when they are tailored for addressing the appropriate level, and when researchers are aware of limitations and constraints of quantitative methodologies. In order words, the closer the level of analysis is to the micro level, the more the differences among contexts and cognitive frames have to be carefully considered. As a consequence, the mentioned cases would greatly benefits from the combination of quantitative and qualitative tools, not only for the validation of the one's results with the other, rather as mixed method or triangulation "to map out or to explain more fully the richness and complexity of human behaviour by studying it from more than one standpoint" (Cohen and Manion 2000, 254). Developing methodological approach using mixed method is actually a key challenge in HE comparative research.

Data and indicators for studying higher education

Despite the advantages, measurement (i.e. understanding differences existing in the dimensions of the observed item) is a critical issue. Numbers must be attached to facts, and the problem is how facts are defined under a numerical representation, identified, and collected. In other words, when quantitative elements are used in comparative research, it is crucial to understand what the numbers actually represent and under what conceptual framework they are produced. This aspect is linked to the meanings and limitations of indicators and data collection.

Indicators

Indicators are social constructs based on conceptual frameworks, providing definitions and normative understanding of the underlying reality. Indicators give a synthetic representation, not a complete and objective description of reality; in this sense they are proxies of the phenomenon they want to represent. They are built upon a conceptual model (definitions, state-of-the-art, delimitation of the elements to be measured), and their specific characteristic is not their quantitative nature—S&T indicators might be based on categories applied to qualitative information, such as survey results –, but their epistemological status: indicators are constructs whose aim is to provide evidence on some properties of the reality which cannot be observed directly because of their complex and abstract nature (Lepori et al. 2008).

Besides this structural limitation, indicators tend to become the phenomenon they address instead of being proxies (indicators indicate, they do not measure, Barré 2010); as to the nature of indicators, they "are intrinsically dependent on a representation (or model or theory) of the topic at stake and are therefore debatable. … Indicators are neither truth nor fallacy, but a common language with a high potential for collective deepening of issues with their underlying values, as long as certain methodological and procedural rules are respected" (Barré and Régibeau 2009).

Keeping these warnings in mind, indicators are a further resource to enhance our knowledge of the HE field, improving the opportunities for comparisons. One important item is the purpose of the indicators use, whether they are just describing some aspects of the reality without leading to further representations, or they assume a connection between the quantities and not observable qualitative properties. They former can be used as simple counts, in order to understand the magnitude of a phenomenon addressed through the synthetic representation provided by the selected measure. The latter case is much more valuable for knowledge production, but it needs a careful work of construction, which implies conceptualization, data collection, and a participatory process, allowing different users to provide insights for a more reliable result.

Indicators can be either measures of entities, or rank orders among chosen variables. Three examples can be outlined from research efforts in the HE field.

One is the case of classifications. The most important classification of higher education institutions is the Carnegie Classification developed in US on the seventies (Van Vught 2009), and conceived as a tool for researchers, then transformed into a league table for several types of users, aimed at depicting the graduate and undergraduates program, the enrolment and the undergraduate profiles, the size and settings of the universities. A different approach has been experimented in the U-Map project (U-Map Report 2010), whose focus was the institutional diversity (between institutions) and horizontal diversity (differences in missions and profiles) allowing different users (scholars, stakeholders,

managers, policy makers) to understanding the HEIs profiles across countries and within countries despite their complexity. Taking into account the seven categories of HEIs diversity outlined in the literature (Birnbaum 1983), U-map classification is based on selected dimensions (teaching and learning profiles, student profiles, research involvement, knowledge exchange, international orientation, regional engagement) that are operationalized through specific indicators (i.e. degree levels, expenditures on teaching and on research, size of student body, peer review publications, start-up firms, foreign degree, etc.); U-map is one example of how quantitative tools based on a sound conceptualization can improve comparison, allowing to create profiles of institutions and to characterize HE systems.

The second example is the U-Multirank tool for higher education institutions, which is based on the U-Map classification, and uses a multidimensional and multi-level user-driven approach, aimed at examining institutions' performance across a wide range of higher education missions. It provides HEIs performance profiles at two levels: for the institution as a whole; and at the level of different disciplinary fields, enabling to compare the institutions in terms of the activities they are engaged in, improving comparative method based on similar cases (Van Vught and Ziegele 2012).

Finally, it is useful to point out the work developed by Huisman et al. (2007), which outlined three caveats for comparing diversity across HE systems, namely the loss of information on institutional variety, the availability of institutional data and the gathering of similar data across different systems. The authors dealt with the mentioned constraints through: (a) focusing on comparisons of a limited set of variables; (b) selecting a number of countries that guarantee a certain comparability of the HE systems; (c) providing a definition of diversity built upon the concept of diversity in biology and ecology; (d) building two indicators of diversity based on the selected concept of diversity: number of organizational types in the HE system and dispersion of HEIs across the organizational types (Huisman et al. 2007, 569).

In the mentioned examples, the value of indicators is that they allow for comparison of diversity across systems and over time, based on rank orders between different variables that constitute types of institutions, whose dynamics and changes can be observed with the support of empirical data, reinforcing the possibility to formulate provisional hypotheses and to generalize the results.

Data collection

Indicators grounds on data, and the availability of comparable data is a crucial item for reliable quantitative approaches. Any data set is the result of a process that might involve several actors—social and political actors as well as scholars and indicator designers, and all the decisions producing the final results derive from several processes of negotiation and mediation among ideas, values, and interests that are put in place in order to reach a consensus on definitions, underlying concepts, and technical issues related to data construction (Reale et al. 2012). Moreover, one has to understand what data actually allows him/her to do. Census data provide detailed information and, in some cases, measures which allow counting. In principle, they are characterised by completeness in terms of coverage, and the data collection is based on a huge work devoted to building definitions, perimeters, timeframes, qualitative checking, and techniques for estimating missing items. On the contrary, ad hoc surveys are characterised by sampling strategies, which cannot always be considered representative in statistical terms; for the most part, they are just a numerical representation of several cases. Besides this limitation, it is important to

recognise that the transformation of qualitative items into numerical ones might be extremely useful to better explore correlations among variables and linkages among different items in contexts, even though the actual testing of causal hypotheses is not always guaranteed.

Relevant examples of data collection in HE are the experiments regarding the collection of data on individual HE organisations pursued within the PRIME Network of Excellence through the AQUAMETH project, which demonstrated the feasibility of an European data collection at university level, or through the European University Observatory Project (OEU), which tested the possibility to develop a common framework for the characterization of the research activities undertaken in universities, gathering data under different institutional conditions.

In more recent time, the European Commission funded the EUMIDA project, whose aim was to create a methodology for data collection on individual universities, to be integrated into regular statistical data collection (Bonaccorsi et al. 2010). The interest of such kind of census data for HE studies is that they provide a basic research infrastructure of comparable micro data, incorporating different types of classification of the HEIs, which can be used for monitoring and mapping purposes, using dimensions such as history/foundation of university, dynamics of growth, specialization pattern, subject mix, funding composition, offer profile and productivity. Interestingly enough, the efforts to develop common higher education data standards are expanding both within and across countries as US, Australian, and European experiences show (Borden et al. 2013). The research community is strongly engaged in these efforts, a fact that is supposed to allow the understanding of the different facets of HE systems and to promote responsible comparisons.

Conclusions

This paper deals with several questions about the advantages of using the comparative method for research in higher education, and the conditions under which it is possible to attain a sound methodology by adopting this approach. Some concluding remarks can now be presented on how to improve methodologies for comparative research in the higher education field.

Firstly, when the comparative method is adopted, higher education should be treated like other public administration sectors. Specific rules apply to comparative research depending on the characteristics of the field—like in several other cases—and the "special" features highlighted in the literature do not affect the rules of comparison.

Secondly, in comparative research quantitative tools, statistical techniques and mathematical treatments are important for reducing the complexity, finding out and graduating similarities and differentiations. When measuring, one can compare using numbers as values that show the different magnitude of the selected entities, or can attribute numerical values to qualitative items in order to identify rank order or the distance between them. In all the cases the conceptualization of what to investigate and to compare is unavoidable, as well as the conceptualization of the proposed quantitative tools, otherwise one can only have sterile statistical treatments that do not contribute to enlarge the body of knowledge in the HE field.

Third, some elements emerge as critical issues to be addressed in order to improve the robustness of the method and the possibility of using measures for comparison. They are: (1) accuracy in the construction of the research design and the extent to which the

comparative method can be applied; (2) defining concepts and eventually theories guiding the investigation and outlining methodological problems for the empirical investigation, recognising the limitation of the selected approach; (3) appreciation of the extent to which comparable data are available and empirical evidence can be collected to test the selected hypotheses or to develop new knowledge within a given conceptual framework; (4) country/institutional specificities as well as field/sector/discipline specificities must be carefully considered; the decision to take these aspects into account for contextualisation or to ignore them is an important element in the comparative method.

When comparable data for robust statistical analyses and for the design of reliable indicators are available, quantitative resources cannot be seen just as complementary analytical tools of qualitative evidences; in this respect, one important item to be explored is how far mixed methods and triangulation techniques can improve the empirical base of the comparative analysis widening the scholars' understanding.

Another matter for further analysis concerns the extent to which the use of quantitative techniques might be used for broader comparisons, addressing linkages, similarities and differences in the content and policy logics of higher education and other related fields of policy action, namely education and culture.

References

Aldricht, H. (2009). Lost in space, out of time: Why and how we should study organizations comparatively. In B. King, T. Felin, & D. Whetten (Eds.), Studying differences between organizations: Comparative approaches to organizational research. *Research in the Sociology of Organizations, 26*, 21–44.

Allardt, E. (1990). Challenges for comparative social research. *Acta Sociologica, 33*, 183–193.

Azarian, R. (2011). Potential and limitations of comparative method in social science. *International Journal of Humanities and Social Science, 1*(4), 113–125.

Bagshaw, D., Lepp, M., & Zorn, C. (2007). International research collaborations: Building teams and managing conflicts. *Conflict Resolution Quarterly, 24*(4), 433–446.

Barré, R. (2010). Towards socially robust S&T indicators: Indicators as debatable devices, enabling collective learning. *Research Evaluation, 19*(3), 227–231.

Barré, R., & Régibeau, P. (2009). *Report of the expert group "ERA indicators and ERA monitoring"*. Luxembourg: EC.

Birnbaum, R. (1983). *Maintaining diversity in higher education*. San Francisco: Jossey-Bass.

Bonaccorsi, A., Brandt, T., De Filippo, D., Lepori, B., Molinari, B., Niederl, A., et al. (2010). *Feasibility study for creating a European University Data Collection. Final Study Report*. Brussels: European Commission, DG Research.

Borden, V. M. H., Calderon, A., Fourie, N., Lepori, B., & Bonaccorsi, A. (2013). Challenges in developing data collection systems in a rapidly evolving higher education environment. *New Directions for Institutional Research, 157*, 39–58.

Cohen, L., & Manion, L. (2000). *Research methods in education*. London: Routledge.

Collier, D. (1993). The comparative method. In A. W. Finifter (Ed.), *Political science: The state of the discipline II* (pp. 105–119). Washinghton DC: American Political Science Association.

Fairbrother, G. (2007). Quantitative and qualitative approaches to comparative education. In M. Bray, et al. (Eds.), *Comparative education research: Approaches and methods* (pp. 39–62). Dordrecht: Springer.

Huisman, J., Meek, L., & Wood, F. (2007). Institutional diversity in higher education: A cross national and longitudinal analysis. *Higher Education Quarterly, 61*(4), 563–577.

Kogan, M. (1996). Comparing higher education systems. *Higher Education, 32*, 395–402.

Kogan, M., & Bleiklie, I. (2006). Comparison and theories. In M. Kogan, et al. (Eds.), *Transforming Higher Education: A comparative Study* (pp. 11–33). London and Philadelphia: Jessica Kingsley Publishers.

Kuechler, M. (1987). The utility of surveys for cross-national research. *Social Science Research, 16*(3), 229–244.

Lange, P., & Meadwell, H. (1985). Typologies of democratic systems: From political inputs to political economy. In H. J. Wiarda (Ed.), *New directions in comparative politics*. Westview: Boulder.

Lazarsfeld, F., & Barton, A. (1951). Qualitative measurement in the social sciences: Classifications, typologies and indices. In D. Lerner & H. D. Lasswell (Eds.), *The policy sciences: Recent developments in scope and method* (pp. 155–192). Stantford: Stantford University Press.

Lepori, B., Barré, R., & Filliatreau, G. (2008). New perspectives and challenges for the design and production of S&T indicators. *Research Evaluation, 17*, 33–44.

Lepori, B., & Bonaccorsi, A. (2013). *The socio-political construction of a European census of higher education institutions. Design, methodological and comparability issues*. Minerva. doi:10.1007/s11024-013-9235-9.

Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review, 65*(Sept), 682–693.

Manardi, A. (1990). Classification, typology, taxonomy. *Quality & Quantity, XXIV*(2), 129–157.

McMichael, P. (1990). Incorporating comparison within a world-historical perspective: An alternative comparative method. *American Sociological Review, 55*(3), 385–397.

Neave, G. (1996). Higher education policy as an exercise in contemporary history. *Higher Education, 32*, 403–415.

Page, E. C. (1995). Comparative public administration in Britain. *Public Administration, 73*(1), 123–141.

Ragin, C., & Zaret, D. (1983). Theory and method in comparative research: Two strategies. *Social Forces, 61*, 731–754.

Reale, E., Inzelt, A., Lepori, B., & van den Besselaar, P. (2012). The social construction of indicators for evaluation: Internationalization of funding agencies. *Research Evaluation, 21*, 245–256.

Sartori, G. (1970). Concept misformation in comparative politics. *American Political Science Review, 64*(4), 1033–1053.

Smelser, N. (1976). *Comparative methods in the social sciences*. Englewood Cliffs, NJ: Prentice-Hall.

Teichler, U. (1996). Comparative higher education: Potential and limits. *Higher Education, 32*, 431–465.

Valimaa, J. (2009). On comparative research in higher education. In A. Amaral, et al. (Eds.), *From governance to identity* (pp. 141–155). Dordrecht: Springer.

Van der Ploeg, F., & Veugelers, R. (2008). Towards evidence-based reform of European Universities. *CESifo Economic Studies, 54*(2), 99–120.

van Mierlo, H., Vermunt, J. K., & Rutte, C. G. (2009). Composing group-level constructs from individual-level survey data. *Organizational Research Methods, 12*(2), 368–392.

Van Vught, F. A. (Ed.). (2009). *Mapping the higher education landscape: Towards a European classification of higher education*. Dordrecht: Springer.

Van Vught, F. A., Bartelse, J., Huisman, J., & van der Wende, M. (2005). *Institutional profiles, towards a typology of higher education institutions in Europe*. Enschede: CHEPS.

Van Vught, F. A., Kaiser, F., File, J. M., Gaethgens, C., Peter, R., & Westerheijden, D. F. (2010). *U-Map-The European classification of higher education institutions*. Enschede: CHEPS.

Van Vught, F. A., & Ziegele, F. (Eds.). (2012). *Multidimensional ranking: The design and development of U-multirank*. New York: Springer, Science+Business Media.

Verba, S. (1967). Some dilemmas in comparative research. *World Politics, 20*, 111–127.