**ECPR General Conference 2014, Glasgow, 3-6 September**

Panel on:
**The role of ideas and indicators in science policies and research management**
**Chairs: Laura Cruz-Castro and Luis Sanz-Menéndez, CSIC IPP**

Session S018. Governing Knowledge: Policy and the Politics of Knowledge Production and Use

# The transformative power of research evaluation: effects on University governance and practices

Emanuela Reale*
Institute for research on firm and growth CERIS CNR, IT
Via dei Taurini 19 Rome 00185 Italy
e.reale@ceris.cnr.it

Giulio Marini
CERIS CNR, IT and CIPES, PT
g.marini@ceris.cnr.it

**Abstract**

The paper investigates how research evaluation, and the use of specific quantitative tools for the assessment purposes, is likely to transform universities into more complete organizations, affecting hierarchy and rationality. The research questions are: is evaluation, by the way of research evaluation, transforming the hierarchies within the universities and the control mechanisms? How evaluation is affecting the vertical diversity within and between universities? We expect research evaluation emerges as a powerful means used by the leadership in order to reinforce the universities as professional organizations, reducing the collegiality; at the same time we expect signals of internal resilience aimed at securing the maintenance of the existing power distribution.
The paper comparatively investigates how the different levels perceive the relationships of the universities with the national agencies for evaluation; these perceptions will be compared to other information about the use, and the degree of relevance of evaluation in terms of leadership (Rector level), government (Board, Senate and Administration), and in terms of results in practices (Middle Management level). Besides this scheme, some light could be shed on the modes of impact of research evaluation to steer the universities.

* Corresponding author. Draft paper, do not quote without the authors' permission.

University is changing. New perspective and ideas about science as commodity, and education being useful (Nedeva, 2004), impact the relationships between the government and the Universities. To this regard HEIs have increasing their accountability (in a context of an increasing formal autonomy that pushes to find new and different patterns possibly linked to some tool such as evaluation), and the latter trying to cope with the new challenges, for instance by adopting different strategies. Moreover, the idea of science as being accountable and contributing to economy and society, goes with the introduction of several policy instruments, such as competitive funding and evaluation, suitable to improve the capability of steering and control, as well as the strategic capabilities of the organizations.
The aim of the paper is to focus on how evaluation is likely to transform universities into more complete organizations (Brunsson N., Sahlin-Andersson K., 2000), affecting hierarchy and rationality. Our research questions are: is evaluation, by the way of Quality Assurance and research evaluation, transforming the strategy of universities and their control mechanisms? How evaluation is affecting the vertical diversity within universities? We consider evaluation as a policy instrument aimed at realizing more accountable universities. So forth we are interested to investigate how Universities implement evaluation and to what extent evaluation among itself and among other tools has a transformative power.
We assume that evaluation is an instrument that Universities use more as a soft steering tool, with a formative role, rather than as a command and control tool, with summative effects. Universities implement evaluation differently, according to their internal characteristics (i.e. size and age), and according to characteristics that derives from the national contexts (namely level of «formal autonomy» as it is stressed here). We assume that the impact of evaluation on hierarchy is much more evident than on rationality. Moreover, we consider "evaluation" as a word with a broad meaning, including instruments such as research evaluation and quality assurance, with very different, often idiosyncratic, effects.

**1. Theoretical background**

The emergence of universities as organizations has been extensively discussed in the literature, with several models proposed and several empirical investigations upon how they were implemented in different countries (Weik, 1976; Clark, 1983; van Vught, 1993; De Boer et al., 2002). The process of transformation of universities in stronger organizations was mainly pushed by the managerial principles linked to the NPM paradigm, which largely inspired the national reform of the HE systems from the eighties (Paradeise et al, 2009), impacting on Universities and producing several, often unexpected, changes (Hood, 2004; Bleiklie, 2007). To reduce the internal fragmentation and to overcome the loosely coupling features are both goals in most of the government reforms in Europe, usually associated with the recognition of formal autonomy and the quest for accountability toward the government and external stakeholders, as well as with incentives toward introducing market-like mechanisms in order to improve efficiency, effectiveness and bettering the competitiveness of universities.
Evaluation is one of the instruments that is supposed to enhance the command and control capability of the government, both the State toward the universities and the internal central government of the Universities toward their lower hierarchy layers. With the setting of a new steering-at-the-distance approach, a different repartition of power inside the universities are expected, as well as the emerging of new actors influencing the universities choices combined with an increasing capability of the universities themselves in planning, controlling and measuring achieved results. The evidences collected in different studies indicate that, on the contrary, evaluation in many cases produces more effects in terms of legitimation and prestige than in terms of coordination and steering, because of the difficulty to assess the complexity of the academic profession from one side, and the limited possibility to impact on the research agenda of the individuals on the other side (Whitley, 2007).
Reale and Seeber (2012), analysing the evolution of evaluation as a policy instrument, show that evaluation rose as an independent steering instrument mostly as to QA and ex-post research assessment exercises. However the origin and the drivers of change were different, since the European processes played a leading role as to the former, while the latter were mostly linked to the national political initiatives. The common narrative was then declined in very different ways among countries, and the instruments reveal the extent to which they were adapted to the existing characters of the national HE systems.

In all the cases, the expectation was that evaluation should contribute to the construction of universities as "complete" organizations, overcoming the characteristics that make them specific organizations (Musselin, 2007), and transforming their distinctive elements of identity, hierarchy and rationality (Brunsson and Sahlin-Andersson, 2000; Bleiklie et al., 2011; Seeber et al., 2014). Among them, hierarchy and rationality are those of interest for this paper and the consequences of evaluation upon them are inquired.

Hierarchy is the capability of coordinating actors and establishing devices for control and management in order to create "a collective entity that is engaged in a common project and aiming at shared priorities" (De Boer et al. 2007, p.33). As a matter of fact, hierarchy and "hierachyzation" are not new issues in tertiary education system, nor in the recent discourse, nor in scientific literature about higher education (Birmbaum, 2004).

The organizational literature (Ferlie et al., 2007) assimilates higher education institutions to professional organization, built for the purpose of specific groups of professionals, and devoted to assure the dominance of some groups on others, through a well-developed vertical differentiation among groups. Diefenbach T. Sillince (2011) stated that in professional organizations formal hierarchy is strong and it is based on the principles of seniority. Thus actors rank themselves on the base of being senior or junior, of exerting dominance versus subordination: "vertical differentiation is achieved through a variety of means typical for the profession… junior professionals can only become fully accepted if they obey to written and unwritten rules of the profession, if they accept their status as 'apprentice' and the nature of the career path." Professional organizations also work according to the principle of professional autonomy, focusing on the content of the work and intrinsically based on self-regulation, which contributes to shape the hierarchical order. This clashes to some extent with the principle of seniority because it is against the idea of dominance, subordination and obedience. To this respect informal hierarchies emerged as ways for convergence, and the informal principle of hierarchical order becomes the domination among semi-autonomous professionals. In the paper we assume that the pressures towards excellence exerted by the European governments in the last two decades (Bleiklie, 2011), allowed universities to use evaluation as a mean for restructuring the intra-organizational university hierarchies, impacting the vertical differentiation and transforming the formal and informal hierarchies. The mode and the directions of the mentioned transformations are not homogeneous between countries and universities, nor stable, rather they tend to create diversities, which generally have been evolving following path-dependency patterns and is still now not in depth inquired.

Rationality deals with the development of instruments for setting intentional objectives, strategic planning and efficient and effective decision making based upon hard facts. As a consequence, constructing organization would imply evaluation becoming more prominent as policy instrument and strategic mean, as well as the development of indicators, standards and performance assessment (Brennan and Shah, 2000; Segerholm C., 2001). On the one hand, it might create the possibility of developing instruments for social control inside the universities; on the other hand, it entails a potential for external control on the organizational processes and for social control inside the universities (Bleiklie et al, 2011, Reale et al., 2011).

To this regard this work contributes to the empirical testing of the mentioned potentialities adopting a comparative perspective focused on the Universities perceptions and behaviours. As to hierarchy, it explores the importance of evaluation inside the universities, the impact on the activities, and on the distribution of decision-making power. It also studies if evaluation is associated with a stronger central coordination and control from one side and a more precise allocation of responsibility and accounting rules to superiors or external stakeholders on the other side. As to rationality, evaluation can drive the setting of objectives and goals, measuring results and impacting on resource allocation, thus introducing a 'management by objective' practice (de Boer et al, 2007).

**2. Empirical setting and methodology**

The paper uses the dataset developed within the EUROCORE-TRUE Project - Transforming Universities in Europe - through a survey directed to different organizational levels within a sample of twenty-six Universities in eight European countries (Germany, France, Italy, Netherlands, Norway, Portugal, Switzerland, United Kingdom). We collected 696 answers, with a response rate of 48% overall, with no

particular differences in terms of response rate for role (Rector, Board, Senate, Central Administrators, Middle management: e.g Deans, Directors of Departments, or Centres[1]), university and country. The analysis is carried out at the organizational levels[2]. The country specificities are included in the analysis by the way of some institutional characteristics (see infra).
The survey used five different questionnaires for position held in the University: Rector/President; Board member; Senate member; Central Administrator; Middle Managers (deans & heads of departments) with some identical questions[3]. Shop floor, and in general academics that are not in charge of government roles, were not asked to participate Beside this limitation, the dataset allows to have a large set of information about how academics involved in actual decision-making power perceive and judge their role and the influence of evaluation on the University strategies and concrete activities.
Board's, Senate's and Rectors' questions have been analyzed in order to understand to what extent the different levels of government are reckoning and realizing the actual implementation of evaluation. By investigating the different government levels, we intend to look at the effects produced at the policy level, dealing with the relationships of universities with national agencies for evaluation and quality assurance. Also we want to look at practices and influence, strategies and management. This perception level is compared to other information about the use and degree of relevance, especially in terms of manifest strategies and leadership (Rector level), and in terms of results in practices (Middle Management level). The last two levels determine respectively the learning features about evaluation in strategies (designed objectives) and outcomes (activities concretely carried out). Evaluation is supposed to impact on the central coordination and control, the allocation of responsibilities to leaders and units, constructing management and reinforcing Middle Managers (Deans, Directors of Departments, Centres, etc.) position, influencing the distribution of decision making power (DMP).
Some methodological details now follows. All questions and variables are assumed as valid since many questions are about perceptions and opinions. As to the reliability, Cronbach Alphas have been computed to test the reliability of Likert scales that are the main scale of the questionnaires; in few sets of variables this test was not satisfactory and the variables have been expelled. In some cases a *deflation of variables*[4] has been done: the 12 variables with evaluation as an issue of a set of other topics have been deflated on the basis of the assumption that importance and relevance play a "zero-sum game". In this way, the values obtained have a stronger soundness in terms of empirical evidence of the relevance of evaluation with respect to other tools or factors (evaluation is not the only tool that may steer HEIs). For instance, in two different questions "evaluation" might have the same score in terms of importance in the same scale, let's say from 1 to 5. If the mean of all the items of the two questions are not the same, those equivalent scores have actual different level of importance. At least, the data on the distribution of the actual decision-making power (DMP.X.X) with a total of 39 variables for one question have been normalized for levels (Rectors, Senate, Board, Central Administrator, Middle Management) so that the sum of values for levels has to be 100 for each university. We assume that in this case normalized data are sounder since power is assumed as a zero sum game within the levels analyzed.

**3. What does descriptive analysis tell?**

Descriptive analysis presents some general characteristic of the sample of Universities surveyed through a selection of variables. (Table 2).
Universities in the sample have a good level of formal autonomy in national context (0,574 in a range from 0 to 1), thus a good room of manoeuvre is guaranteed by the national laws and regulations; among the different issues considered, setting autonomously the steering instruments is lower than the average value

[1] Since Rectors/Presidents and Central Administrators are usually chaired by only one persons, in these cases data have been always collected.
[2]. Regarding hierarchy, data have been computed in order to let count 1 any pair of HEIs and role (i.e. Rector of HEI1 = 1; MM of HEI1 = 1/30 if we collected 30 valid questionnaires among MM of HEI1).
[3] See annex for further details about the merging of the 5 row datasets.
[4] Deflation has been computed as the quotient between the Delta of a mode from the mean of the items all over the standard deviation of all the items in the variable.

(0,477), but as far as evaluation is concerned, Universities have the highest level of formal autonomy in defining internal evaluation of research (0,875), which is even much higher than for QA (0,666).
According to the Rectors, evaluation is important but does not play a direct and strong role for the definition of the University strategy (R.24.12_Def = -0,757[5]) if compared with other issues such as leadership (R.24.2_Def = -1,706), or the cooperation with academics; being connected with evaluation agencies (R.6.3_Def = 0,834) is not so much determinant as other institutional relationships may be, such as funding agencies (R.6.4_Def = -1,455). The other central government levels also confirmed the Rectors' view. Facstr_DEF and Pfacstr_DEF show the importance of the evaluation as a matter to be taken into account for the definition of the strategy among other 11 items. The first is about the present, the latter looks back at a span of 5 years. In this comparison of time evaluation has become a more relevant factor for the definition of HEI's strategy (Factstr.12_Def = 0,360; Pfactstr.12_Def = -0,383), even though currently other factors play a major role, funding first and foremost (Pfactstr.2_Def = -1,668).
Consequently, Evaluation agency is important (Actor_inf.3_Def = -0,188) is a relevant actor with respect to other external actors such as Ministries or local governments, but not the very most influent actor for the definition of the HEI's strategy, funding agencies being much more important (Actor_inf.2_Def = -0,583). Moreover, when evaluation is external, the influence on financial issues (e.g. determining the budget of institutions and academic units), and on research priorities is higher than personal impacts and teaching programs.
The descriptive analysis confirms that evaluation is a broad term, including different items such as quality assurance, teaching evaluation and research evaluation, not all having the same effects on the concrete activities carried out by the Universities. Central government level of Universities perceives evaluation of research as the most influent factor on the concrete academic activities (to define and in case change them). Evaluation of teaching, although relevant (Imfitem.3_Def = -0,330), is much less influent for the University activities than research evaluation (Imp.1_Def = -1,417), while the influence of the rankings is generally low (Infitem.4_Def = 1,295) even though with a high standard deviation (0,9297). Accordingly, Rectors and Boards consider research evaluation a much more influent steering instrument on the University decisions than QA (Imp.3_Def = 0,748; Imp.4_Def = 0,385), and the outcome of evaluation more suitable to change the content of internal decision-making.
The distribution of decision making power differs according to the issues: when the issue of setting the evaluation rules and procedures for the Units evaluation is concerned (DMP.6), the decision making process is mainly allocated at the central level (DMP.6c = 0,479), but, as far as the assessment of academic individual performance is concerned (DMP.11), the decision making process is more likely in the Faculty hands (DMP.11f = 0,772). The Shop Floor level seems to have a very little role shaping the rules or the practices of evaluation.
Nevertheless, when the effectiveness of evaluation is concerned, Middle managers[6] think evaluation impacts much more on reputation, both internal (MM.11.5 = 1,860; MM.10.5 = 2,510) and external (MM.11.6 = 1,934; MM.10.6 = 2,599), than on funding (MM.11.1 = 2,060; MM.10.1 = 3,264) and other concrete issues such as recruitment (MM.11.2 = 2,258; MM.10.2 = 3,092) and career (MM.10.3 = 3,014; MM.11.3 = 2,265), number of post-docs and phD positions (MM.10.4 = 3,492; MM.11.4 = 2,341), or content of teaching programmes (MM.10.7 = 2,590; MM.11.7 = 2,296), although they consider the impact on the mentioned issues not negligible. As a whole, the impact of research evaluation is generally considered stronger than teaching evaluation.
The importance of several effects of evaluation and QA has been investigated through a set of question regarding the effects of evaluation and quality assessment (EVQA). We coded the nine effects of the questionnaire into four main types as follow: a) generating conflicts among academics and between managers and academic, b) generating opportunistic behaviours among academics, c) generating efficiency and effectiveness improving quality, and transparency of strategic decision making, d) generating bureaucracy with more norms, rules and constraints to the academic freedom. It is interesting to note that

[5] In Deflated variables, scores are meant as follows: the less the score, the more influent is the item. Negative scores means that the item in question is more important than the mean; positive scores mean that that item is less important than the mean of all the items.
[6] In these variables scores are means with a rage from 1 (more important) to 5 (less important).

the most important effect is on bureaucracy, although it is not so high (mean value of 2,567), the less important is on conflicts (EVQA.c = 3,215). Universities recognized the bettering of efficiency and effectiveness as a more important effect (EVQA.ee = 2,594) than opportunistic behaviours (EVQA.b = 2,805).

As to the influence of evaluation and QA when they are external, teaching programs and personal impact (e.g. career of individual academics – 2,550 and 2,543) score higher than the financial impacts (e.g. budgets of institutions and academic units – 2,406), while the impact on research priorities present the lowest value (2,260).

Summing up, there are evidences that evaluation is an important tool for universities, whose influence is growing by time, producing stronger effects on concrete University activities than on setting strategies. So we can expect a deeper effect of evaluation on internal relationships of Universities than on their capability to become more rational organizations. Furthermore for the Central government the relevance of relationships with QA and Evaluation agencies as external actors is not so prominent (Instit.3_Def = 0,830). Thus, it is not surprising that the impact of evaluation is not so strong on the allocation of resources both financial and human ones. The strength of the instruments lies more precisely on the capability of being influential on internal and external reputation as seen from the opinions by the middle managers.

Two important distinctions emerge: one is between external and internal evaluation, the other is between teaching and research evaluation, as well as between QA and research evaluation. External evaluation is more able to influence resource allocation than internal evaluation; internal evaluation importance and influence is visible more on reputation and on concrete University activities, rather than on central government steering. Research evaluation produces more important impact than teaching evaluation on several issues of the academic life, including the content of teaching programmes and the enrolment of students. Similarly QA has a weaker correlation on University strategic decisions than research evaluation.

**Tab. 2 - Descriptive statistics at HEI level**

| Variable | Label | Items | Mean | S.D. |
|---|---|---|---|---|
| AU_total * | General degree of formal autonomy in national contexts | | 0,574 | 0,1402 |
| AU_steering.tool * | Degree of formal aut. in determination of steering instruments | | 0,477 | 0,1036 |
| AU_ev_int.QA * | Formal autonomy in defining internal QA | | 0,666 | 0,2888 |
| AU_ev_int.RE * | Formal autonomy in defining internal evaluation of research | | 0,875 | 0,2321 |
| Accout_EXT_DEF | Extent to which an HEI has to account to external actors | | 0,695 | 1,2448 |
| DPM.6c | Actual decision making in setting the rules for research evaluation of teams | Central level | 0,479 | 0,1005 |
| DMP.6f | | Faculty level | 0,391 | 0,0671 |
| DMP.6s | | Shop Floor level | 0,130 | 0,0634 |
| DMP.11c | Actual decision making in individual research evaluation | Central level | 0,261 | 0,1486 |
| DMP.11f | | Faculty level | 0,472 | 0,0875 |
| DMP.11s | | Shop Floor level | 0,268 | 0,1348 |
| Infitem.1_DEF | Influence on the activities carried out by the HEI | Evaluation of research quality | -1,41 | 0,7246 |

| | | | | |
|---|---|---|---|---|
| | | | 7 | |
| Infitem.2_DEF | | Evaluation of teaching quality | -0,330 | 0,71000 |
| Infitem.3_DEF | | Accreditation | 0,453 | 0,83386 |
| Infitem.4_DEF | | Rankings | 1,295 | 0,92297 |
| Actor_inf.3_DEF | Influence of evaluation agency for the strategy over other 13 actors | | -0,188 | 0,69049 |
| Actor_inf.2_DEF | Influence of funding agencies for the strategy over other 13 actors | | -0,583 | 0,64026 |
| R.24.12_DEF | Importance of evaluation as a factor in developing the current institutional strategy | | 0,757 | 1,32121 |
| R.24.3_DEF | Importance of leadership as a factor in developing the current institutional strategy | | -1,706 | 0,92013 |
| R.6.3_DEF | Importance to be connected with quality assurance agencies (accreditation/evaluation) | | 0,834 | 0,90062 |
| R.6.4_DEF | Importance to be connected with funding agencies (accreditation/evaluation) | | -1,455 | 0,51021 |
| Factstr.12_DEF | Importance of evaluation for the definition of the current strategy | | 0,360 | 0,90087 |
| Factstr.3_DEF | Importance of funding agencies for the definition of the current strategy | | -2,364 | 0,55523 |
| Pfactstr.12_DEF | Importance of evaluation as a factors compared to 5-10 years ago | | -0,383 | 3,37134 |
| Pfactstr.3_DEF | Importance of funding agencies as a factors compared to 5-10 years ago | | -1,668 | 0,67001 |
| Imp.3_DEF | Influence of the steering instruments on the decisions | Quality assurance | 0,748 | 0,85707 |
| Imp.4_DEF | | Research evaluation | 0,385 | 0,81787 |
| Instit.3_DEF | Level of interaction with the Q.A. agencies over other 5 actors | | 0,830 | 0,91159 |
| Instit.4_DEF | Level of interaction with the Funding agencies, over other 5 actors | | -0,735 | 0,76047 |
| MM.10.1 | | Funding and budget | 3,264 | 0,48054 |
| MM.10.2 | Impact of teaching evaluation on... | Recruitment of new staff | 3,092 | 0,4955 |
| MM.10.3 | | Careers of academic staff | 3,014 | 0,47037 |
| MM.10.4 | | Number of post-docs or PhDs | 3,49 | 0,47 |

| | | | | |
|---|---|---|---|---|
| | | | 2 | 61 |
| | | | 2,51 | 0,41 |
| MM.10.5 | | Unit's reputation in the HEI | 0 | 99 |
| MM.10.6 | | Unit's reputation outside the HEI | 2,59 | 0,40 |
| | | | 9 | 1 |
| MM.10.7 | | Content of the teaching programmes | 2,59 | 0,47 |
| | | | 0 | 66 |
| MM.10.8 | | Enrolment of students | 2,80 | 0,40 |
| | | | 7 | 47 |
| MM.11.1 | Impact of research evaluation on... | Funding and budget | 2,06 | 0,43 |
| | | | | 44 |
| MM.11.2 | | Recruitment of new staff | 2,25 | 0,43 |
| | | | 8 | 46 |
| MM.11.3 | | Careers of academic staff | 2,26 | 0,46 |
| | | | 5 | 22 |
| MM.11.4 | | Number of post-docs or PhDs | 2,34 | 0,46 |
| | | | 1 | 07 |
| MM.11.5 | | Unit's reputation in the university | 1,86 | 0,37 |
| | | | 0 | 04 |
| MM.11.6 | | Unit's reputation outside the university | 1,93 | 0,35 |
| | | | 4 | 86 |
| MM.11.7 | | Content of the teaching programmes | 2,29 | 0,38 |
| | | | 6 | 2 |
| MM.11.8 | | Enrolment of students | 2,12 | 0,42 |
| | | | 2 | 13 |
| EVQA.1 | How would you rate the effects of evaluation and quality assessment as regards the following issues? | Increasing conflicts among academics | 3,23 | 1,06 |
| | | | 5 | 4 |
| EVQA.2 | | Increasing conflicts between managers and academics | 3,17 | 1,08 |
| | | | 2 | 5 |
| EVQA.3 | | Academic activity subjected to more norms and rules | 2,56 | 1,03 |
| | | | 7 | 4 |
| EVQA.4 | | Increasing opportunistic behaviours in teaching and research | 2,71 | 1,07 |
| | | | 2 | 8 |
| EVQA.5 | | Increasing constraints on academic freedom | 2,89 | |
| | | | 0 | 1,14 |
| EVQA.6 | | Improving teaching quality | 2,62 | 0,95 |
| | | | 7 | 9 |
| EVQA.7 | | Improving research quality | 2,47 | 0,97 |
| | | | 2 | 8 |
| EVQA.8 | | Increasing transparency | 2,67 | 1,06 |
| | | | 3 | 4 |
| EVQA.9 | | Improving strategic decision-making | 2,60 | 1,02 |
| | | | 4 | 1 |
| EVQA.c | EVQA 1+2 | evaluation generating conflicts | 3,21 | 0,98 |
| | | | 5 | 83 |
| EVQA.ch | EVQA 4+5 | evaluation generating changing behaviors | 2,80 | 0,99 |
| | | | 5 | 11 |
| EVQA.ee | EVQA 6+7+8+9 | evaluation generating efficiency and efficacy | 2,59 | 0,83 |
| | | | 4 | 87 |
| EVQA.b | EVQA 3 | evaluation generating | 2,56 | 1,03 |

| | | | | |
|---|---|---|---|---|
| | | boureaucracy | 7 | 38 |
| EX.EVA.1 | How would you rate the influence of external evaluation and quality assessment on your university in terms of: | personal impacts (e.g. career of individual academics) | 2,543 | 0,4262 |
| EX.EVA.2 | | financial impacts (e.g. budgets of institution and academic units) | 2,406 | 0,4467 |
| EX.EVA.3 | | teaching programs (content impact) | 2,550 | 0,4180 |
| EX.EVA.4 | | research priorities | 2,260 | 0,3645 |

*Source: Own elaboration on TRUE dataset*

**Refer to national contexts 0 = non autonomous, 1 = autonomous*
*AU_ev_int.QA and AU_ev_int.RE are part of AU_steering.tool with other variables. AU_steering.tool with other 7 dimensions compose AU_total. See annex for further details*
*With the exception of DMP, the less is the value, the strong is the item*

**4. The main significant correlations at HEI level**

Correlations provide empirical evidence of variables that are correlated in a positive or negative ways. The evidences outline how the importance of evaluation on hierarchy and rationality is linked to the presence or the absence of certain characteristics of the national context or of the universities themselves, as well as how they are associated to other specific items or tools. We intend to investigate four main items, which are supposed to produce effects on hierarchy and rationality: namely the actors' constellation influencing Universities, the linkages between evaluation and the centralization of decision-making power, the role of Faculties and middle managers, and the factors influencing the impact that evaluation might have. Hereafter the main results are described.[7]

**Tab. 3 – Selection of correlations**

| *Pairs of variables* | | *R* | *p* |
|---|---|---|---|
| Size | R24 | 0,766 | 0,000 |
| DMP11.C | RECT.2 | 0,568 | 0,002 |
| MM10.1 | INFITEM1 | -0,590 | 0,002 |
| MM11.3 | DMP.S | -0,570 | 0,003 |
| DMP11.F | IMP.3DEF | 0,548 | 0,004 |
| MM11.3 | UNIMAN | 0,557 | 0,004 |
| ACTORINF | Age | -0,538 | 0,005 |
| DMP 6F | MM10.1 | 0,544 | 0,005 |
| DMP11S | R10 | 0,547 | 0,005 |
| Instit | ExEv.3 | 0,524 | 0,006 |
| MM10.6 | Accountability_EXT | 0,530 | 0,006 |
| DMP11.F | INFITEM.1 | -0,510 | 0,008 |
| DMP.F | Accountability_EXT | 0,506 | 0,008 |
| MM10.6 | UNICOL | 0,521 | 0,008 |
| MM11.8 | EXTEV4 | 0,515 | 0,008 |

[7] Correlations from now to the end of the paragraph show the actual values of the correlation matrix. According to the Likert scales and other variables as appeared to the questionnaires, a negative sign may correspond to a "the more the more" meaning, such as a positive correlation might have a "the more the less" interpretation. Table 3 fosters this information.

| | | | |
|---|---|---|---|
| DMP11F | MM10.1 | 0,509 | 0,009 |
| STR_BUILDING1 | DMP6.C | 0,499 | 0,010 |
| MM10.6 | DMP.C | -0,507 | 0,010 |

Actors' constellation reveals three major strong correlations:

- The more the universities are old the less the relationships with the Evaluation Agencies are important (ACTORINF*AGE -0,538);
- The more the Universities have strong relationships with the QA Agencies, the more evaluation is reckoned to have an impact over teaching programmes (INSTIT*EXTEV.3 0,524).

Seniority is a characteristic relaxing the need of Universities to create a close relationship with Evaluation agencies, probably due to the fact that old organizations have strong identities and capability to adapt to the rules autonomously. If the University decide to have a close interaction with the Agency, the impact on teaching programmes is much more consistent. Thus, the choice of a University to strengthening the relationships with external actors is an issue producing effects in terms of impact of the steering tool eventually associated with the actor itself.

As to the centralization of decision-making power, three main results can be outlined:

- The more the universities are big, the more Rectors consider evaluation an important steering instrument (R24*SIZE -0,766);
- The more the university strategy is developed at the highest central level the more the power of central decision making on setting the rules and procedures for evaluation of units (STR_BUILDING1*DMP6.C -0,499);
- The more the decision making power at central level on evaluation of academic individual performance, the more the Rector's power on selection appointment and dismissal of academic staff (DMP11.C*RECT.2 -0,568).

Therefore size matters: rectors of big universities are more likely to use evaluation as a steering instrument than those leading small organizations; the mentioned result can be read together with some interesting observations on conditions under which the universities deals with strategies. When the strategies are developed at the central government level, Universities tend to allocate at the same level also the decision-making about the instruments to implement the strategies themselves, e.g. in the case of evaluation, rules and procedures for assessing the internal units. It means that the tendency toward reinforcing the centralization of the strategies does not necessarily go with decentralization processes as to the implementation of the instruments for realizing the strategies themselves. The discussed conclusion is coherent with the strong linkages between the centralization of the decision making power on the assessment of the individual performance and the concentration of power in the rectors' hands on decisions related to the human capital of the University.

Then, we can outline some elements about the role of Faculty and middle managers:

- The more the decision making power at Faculty level on evaluation of individual performance, the more the influence of QA as steering instrument (DMP11.F*IMP.3DEF -0,548)
- The decision making power at the Faculty level on evaluation of individual performance become stronger as the influence of research evaluation on the universities' activities become weaker (DMP11.F*INFITEM.1 0,510)
- The more the decision making power at Faculty level, the more the external accountability of the universities (DMP.F*Accountability_EXT -0,506);
- When Faculties have strong power on the individual evaluation of academic performance and on setting rules and practice, then the impact of teaching evaluation on the funding level become stronger (DMP11F* MM10.1 and DMP 6F-MM10.1 respectively -0,544 and -0,509)

The mentioned results show that the repartition of power inside the university follows very specific paths, and tend to resist to changes imposed by the managerial paradigm. When Faculties have strong decision-making power on evaluation of individual performance, several other characteristics can be expected, such as an enhanced effect of QA as steering instrument of the University, or even a strong role of the external accountability. At the same time, a substitution effect emerge between QA and research evaluation, which

is linked to the strengthening of the decision making power of Faculties on the assessment of individual performance. In that case Universities performance is more impacted by indicators and standards coming from QA than from research evaluation. It also implies the possibility to have a more prominent role of external actors on University strategies, since the more important is QA, the more the influence of the Evaluation Agencies, and even the more the possibility of stakeholders to influence the internal decision-making. Finally, we can outline the absence of strong correlation between evaluation issues and the decision making power at the shop floor level, except the case when the University is involved in external evaluation because national regulations require this sort of activity or because the University decides to submit itself to an external assessment (DMP11S*R10 0,547). Saying differently, evaluation is an instrument managed by the Central government level and by the Faculties, and the meaning of collegiality as distinctive element of the University is transformed into a sort of repartition of power between different decision-making levels of internal government, according to the specific aim of the evaluation tool. Hierarchy is supposed to become stronger than before, but the internal organization of the university does not necessarily assume a real "vertical" configuration as if any level of power must be referred to the upper level of internal government. In the case evaluation thas is not an internal process, but rather an external exercise, so that here the shop floor level gains a significant power to influence the University choices.
The results about the impact of teaching and research evaluation complete the aforementioned picture:

- The impact of teaching evaluation on the Department reputation outside the university becomes stronger when the internal accountability of the University and the collegialism are stronger (MM10.6*Accountability INT, MM10.6*UNICOL respectively 0,530; 0,507). On the contrary the impact of teaching evaluation on the Department reputation outside the university become stronger when the decision making power at central level is weaker (MM10.6*DMP.C 0,507)
- The more teaching evaluation impacts the funding level, the less the evaluation of research influence the University activities (MM10.1*INFITEM1 -0,590)
- The impact of research evaluation on the career of academic staff is stronger when the university is perceived as managerial, and the decision making power at the shop floor level is weaker (MM11.3*UNIMAN; MM11.3*DMP.S 0,557; 0,570)
- The more research evaluation impacts the university capacity to raise external funding, the more the external evaluation influences the research priorities (MM11.8*EXTEV4 0,515)

Strong linkages between the impact of teaching evaluation, internal accountability and collegialism emerge, which seems to be self-reinforcing issues. Moreover, some elements impeding the production of the mentioned results come out as well: Universities where the central level has a strong decision-making power cannot expect to have a growing impact of teaching evaluation on the Departments reputation. The correlations show very different results in the case of research evaluation: for instance, when funding is strongly driven by the outcome of teaching evaluation the influence of evaluation of research declines, but the reverse effect does not emerge. The explanation might be linked to the different role attributed to teaching and research evaluation, the latter considered more linked to driving the allocation of resources than the former. Research evaluation impacts strongly the research priorities when it substantially enhances the University capability of funding raising, but this circumstance is mainly linked to external evaluation; the effects on academic career are associated with the presence of a managerial culture and the substantial weakening of the decision making power of the shop floor level.

**5. How HEIs throughout Europe gather themselves in groups? Results from the cluster analysis**

Assuming that 26 HEIs in 8 European countries could not be groups necessarily for country or other basic characteristic (size, specialization and the like), it was interesting to stress some national patterns (formal autonomy variables) with a selected bundle of variables about evaluations meant as steering-at-the distance tool. The null hypothesis is to assume 26 universities as European cases and not as cases belonging to different traditions or systems, like the main stream of literature about policy suggests (Gosta Andersen). As will be discussed, it has found out that country affiliation actually still matters since HEIs are basically nested in national contexts where different systems of evaluation play peculiar and different situations. Hierarchical Cluster Analysis – which was chosen to don't have to select a priori the number of clusters – fosters groups of Universities on the basis of different features, which characterize them in a

comparative perspective. In our tests, a dozen of different selection of restricted set of variables has been launched and the soundest is here discussed: 23 cases over 26 have been considered due to a list wise exclusion for missing values (namely NO4, UK2 and UK3 have been automatically dropped out). The final list of variables[8] and the criteria for selection for the cluster analysis are:

- Formal autonomy variables. This assures a strong country feature so that whether HEIs of same countries are in different cluster it means that universities of the same country are really different. As expected, this may happen more easily in those country where formal autonomy in evaluation is higher, so that differentiation can happen;
- The main interesting variables according to issues and correlation matrix have been chosen in order to capture similarities between HEIs;
- Decisions making power, and some perceptions of Middle Managers have been included to evidence the impact of evaluation at different levels within HEIs.

Based on the mentioned criteria, the variables taken into account are:

- 2 DMP (those with the highest S.D. between HEIs): actual decision-making power on evaluation academic individual performance (dmp_11_c), and actual decision making power on setting the rules and procedures for evaluation of units (faculties, institutes, etc.) (dmp_6_f);
- Perception of HEI as managerial or collegial organizations (uni_man / uni_coll);
- External accountability of the University as perceived by the different University levels (Accountability_ext);
- 6 items over the 16 of those related to the impact of teaching evaluation and research evaluation on several activities (the MM10&11 series): funding and budget; Recruitment of new staff; Careers of academic staff; Unit's reputation in the University; on Unit's reputation outside the University; Enrolment of students;
- All Deflated variables (11 variables);
- 4 Formal Autonomy variables related to the following issues:
  - internal quality assurance (au_ev_int_qa);
  - internal research evaluation (au_ev_int_re);
  - setting steering instruments (au_steering_instrument);
  - total formal autonomy (au_total).

Thus, we considered national features of formal structure of HE system (as it is supposed to be on the basis of rules and regulations), such as the role that evaluation plays at national level, and how internal governance of Universities and their hierarchies matter for evaluation as at-a-distance-steering-tool. Fig. 1 shows the result of the Cluster analysis as a dendrogram (full data as separated mean for groups are presented in Appendix 1).

Considering more launches with several selections of variables and the present dendogram, the soundest results looked to be the cut at 6 groups in the current selections of variables, even though three clusters are formed by one HEI only. This odd output can be justified by the actual specificity of IT2 and of UK1 as the only British case, and – regarding IT3 – as a consequence of the strong autonomy that Italian system has in terms of steering instruments. So forth, three main clusters emerged, which are characterized by a specific set of strong variables.

The largest cluster (Cluster IV in the Dendogram and Appendix 1) includes all the Universities of Norway, all the Universities of Germany and four out of five of the Swiss Universities -the excluded is a generalistic one (NO1, NO2, NO3, DE1, DE2, DE3, CH1, CH2, CH3, CH4). Universities of this cluster consider evaluation a strategic tool whose importance grew up significantly in the last five years; the influence of research evaluation and evaluation of teaching quality on the University activities is important (the latter more than the former), the decision making power at the central level is relevant when evaluation is concerned, and it is more important on individual academic evaluation than on setting the rules and procedures for evaluation. Here the country emerged as a central feature, suggesting the presence of a clear hierarchy and vertical repartition of power and responsibility. Interestingly enough, the Universities of this cluster have a lower level of total Formal autonomy than the other clusters, but the formal autonomy on setting the steering tool and on internal QA are the highest. We can synthesise the characteristic of this cluster saying

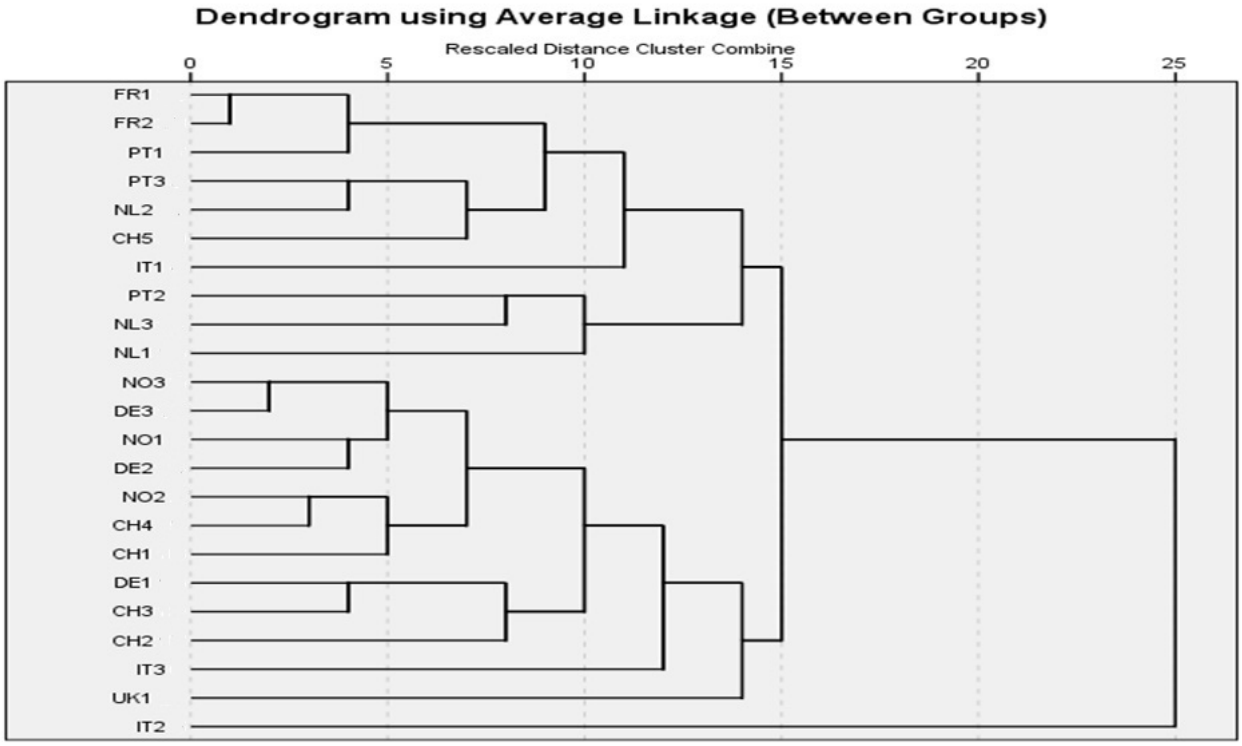
[8] The table of the appendix displays the full list of variables and their means for clusters.

that Universities use their large autonomy to adjust governance power toward the central level to get ready to face and manage the impact of evaluation as an at-the-distance steering tool, without excluding the relevance of the shop floor level, thus pairing the managerialism with the maintenance of a strong collegiality.

A smaller cluster (Cluster V) includes two Universities of the Netherlands and one Portuguese University (NL1, NL3, PT2). Here Universities tend to define themselves as collegial organization, where the Rectors consider very important to be connected with the QA Agencies, but with no actual strong role of evaluation as steering tool is emerging; only accreditation may matter. Even in this case the role of the country characteristics, as to the low formal autonomy on setting the steering instruments, and the low autonomy on internal QA, are important.

The other big cluster (Cluster VI) groups seven Universities of different countries (IT1, PT1, PT3, NL2, FR1, FR2, CH5), which are characterized by institutional behaviours less effective as to the use of evaluation as steering tool, and as to the effects produced by evaluation on the activities developed. Here the Universities perceived themselves less managerial, which have not yet to account to external context, but realize evaluation is an important tool, so that they must find a way for coping with this factor.

**Fig. 1 – Cluster analysis**




The three outlier-cases of the cluster analysis show different characteristics:

IT2, where evaluation has led the institutional change by embodying the tool to anticipate government reforming patterns;

UK1, where evaluation is already part of community since years, thus it has already exerted the transformative power and it is now an ordinary steering instrument;

IT3, a technical university, where rankings emerged as the most influent items on the activities of the University, and quality assurance mechanisms the most influential steering instrument on the decisions made within the university.

To sum up, most of the Universities in the sample can be split from:

- Ones where the concentration of power in a central governance does not imply the end of collegialism nor a strong presence of NPM practices but, as a result, allow Universities to avoid adverse selections and allow to better plan one's development;

- Other HEIs that either may be affected only by QA as an alternative steering-at-the-distance tool or may be defined as "late comers" as to the use of evaluation as an instrument for strategies, and as to the impact on concrete activities developed – in an opposite pole for instance with the IT2 that is a "first mover" case.

Country affiliation does matter, but evaluation (and more generally the presence of steering tool) let emerge the possibility of an increasing differentiation of Universities in a more internationally based comparison based on how the universities use the level of autonomy (room of manoeuvre) granted by the State.

**Conclusions**

The paper analyses the use and the impact of evaluation on Universities in order to understand how to what extent an instrument is able to impact on hierarchy and rationalization, thus transforming universities in more complete organizations. The aspiration is to contribute to the literature about how the HE systems are changing, looking inside the University decision-making power and the impact of evaluation on concrete activities and on strategies, adopting a comparative perspective on 26 universities in eight European countries.

Several limitations of the evidences collected can be outlined, since they cannot deal with different disciplines, gender differences, age/seniority of respondents, and opinions of the scholars in "shop floor level" (people who are not in charge of anything). Thus, the survey is able to supply evidences on how the academics, which are entitled of the internal formal power, perceive the role of evaluation in the different forms of research evaluation, teaching evaluation and quality assurance. The data analysis includes descriptive statistics, correlations and cluster analysis, in order to figure out the basic characteristics of the sample explored, the most significant influences between factors and groups of homogeneous Universities in terms of use and impact of evaluation, beside their location in a specific country.

In this sample of European Universities, evaluation plays an increasing role in governance and definition of strategy, though other factors, both external and internal, remain more relevant (i.e. funding and leadership).

QA, teaching evaluation and research evaluation differ each other for the degree of influence and the different changing dynamics they trigger. In the first case evaluation has a stronger impact than teaching evaluation on academic activities and strategies, as well as on funding issues; in the second aspect teaching evaluation and QA tend to reinforce the decision making power of Faculties and the role of Middle management much more than research evaluation, which is used as a steering instrument managed by the central level, moreover the former is more associated with external accountability than the latter.

Evaluation affects the hierarchies inside the university since it reinforces the Middle managers' position in the case of teaching evaluation. It has a major impact on Departments and University reputation both internal and external, thus, it is able to change the hierarchical key figures of the organization in terms of actual power, and to challenge it as to the different distribution of decision making power, reinforcing both the central government and the middle management. To this regard evaluation produces transformative effects, and strengthens a division of roles between the central and the medium levels. At the same time the importance of the shop floor level seems to be more and more compressed by the use of the evaluation tools, especially when the initiative is an internal one.

Although the shop floor capability to influence the decision-making is weakened, it does not necessarily means that people not involved in middle management positions or central government positions do not have any relevant role in power games. Our findings allow to detecting where the power is allocated within universities, nothing to say about the way by which academics exert their influence through informal relationships. One interesting evidence in this respect is the high value the Universities attribute to the cooperation with academics for the development of the actual institutional strategy: this issue was listed as the second most important factor just after leadership, well above the importance of evaluation. It means that further researches are needed especially for the exploration about the informal relationships and processes inside the Universities as professional organizations. It also shows that beside the effects evaluation produces Universities confirm the tendency to remain specific organizations rather than transforming "completely into complete ones".

Evaluation may go with the concentration of power in the hand of the Rector, but it is not a common case, and this circumstance happens when research evaluation and the setting of rules and procedures for the assessment of the units are concerned. The centralization of the power for evaluation emerges as a self-reinforcing issue: the presence of a strong decision making power of the central level on the evaluation of individual academic performance is generally associated with the strengthening of the power of the Rector on selection and appointment of academic staff. Evaluation impacts the content of teaching programmes, and the influence of research evaluation on strategies is significant. Nevertheless the importance of the tool is lower than other steering instruments, and generating bureaucratization of the internal activities is perceived even more relevant than efficiency and efficacy. Thus, the effects of evaluation on rationalization of Universities emerged as more problematic and questionable; furthermore, they are more related to external evaluation (thus on external steering) than on an internal steering capabilities, and the effectiveness is weaker than funding instruments or the internal leadership. Surprisingly enough, rankings do not emerge as tools influencing strategies or activities: here an interesting research question rises as to the actual effects of instruments for measuring the performance on concrete decision making processes and activities carried out by the organizations.
Universities tend to use evaluation according to the national characteristics of formal autonomy; however, in some cases the influence of the country characteristics can be overcome by individual initiatives of the organizations, which elaborate specific internal dynamics in order to cope with (or to react against) the changes evaluation is supposed to produce. When this occurs, evaluation becomes a source of differentiation between Universities within the same countries.

**References**

Birnbaum, R. (2004). **"**The End of Shared Governance: Looking Ahead or Looking Back.**"** New Directions for Higher Education 127, 5**–**22

Bleiklie I. and Kogan M. (2007), "Organization and Governance of Universities", *Higher Education Policy*, 20

Bleiklie I. (2011), *Excellence, quality and the diversity of higher education systems*, in Rostan M. and Vaira M (eds), Questioning Excellence in Higher Education, Sense Publisher, Rotterdam/Boston/Taipei

Bleiklie,I Enders J, Lepori B., Musselin C. (2011), *Universities as penetrated hierarchies: organizational, rationalization, hierarchization and networking in higher education*, paper presented at the EGOS Conference, Gothenborg, 2011

Brennan J: and Shah T. (2000), *Managing Quality in Higher Education: An International Perspective on Institutional Assessment and Change, Buckingham*, OECD, SRHE and Open University Press

Brunsson N., Sahlin-Andersson K. (2000), Constructing Organizations: The Example of Public Sector Reform, *Organization Studies*, 21, (4), 721-746

Clark B.R.(1983), *The Higher Education System Academic Organisation in Cross National Perspective*, Berkeley, University of California Press

De Boer H., Huisman J., Klemperer A., van der Meulen B., Neave G., Theisens H., van der Wende M. (2002), *Academia in the 21st century*, AWT, n. 28

De Boer H., Enders J., Leisyte L. (2007), Public sector reform in Dutch Higher Education: the organizational transformation of the university, *Public Administration*, 85, 1, 27-46

Diefenbach T. Sillince J.A.A.(2011), Formal and Informal Hierarchy in Different Types of Organizations, *Organization Studies*, 32 (11), 1515-1537

Ferlie E., Lyinn J., Pollit C. (2007), *The Oxford Handbook of Public Management*, Oxford University Press

Hood C., James O., Peters B.G., Scott C. (2004), *Controlling Modern Government*, Edward Elgar, UK – USA

Musselin C. (2007), *Are Universities specific organizations?*, in Krucken G, Kosmutzky A., Torka M. (eds), *Towards a Multiversity? Universities between global trends and national traditions*, Transaction Publishers, New Brunswick (USA)

Nedeva M. (2006), Changing Science: The Advent of Neo-Liberalism, *Prometheus*, 24, 3, 269-281

Paradeise, C., Reale E., Bleiklie I. Ferlie E. (eds) (2009) *University Governance: Western European comparative perspectives*, Springer, Dordrecht

Reale E., Rebora G., Marini G., Turri M. (2011), *Emerging evaluation in higher education institutions as steering instrument: implementation and expectations*, paper presented at the EGOS Conference, Gothenborg, 2011

Reale E., Seeber M. (2012), Instruments as empirical evidence for the analysis of Higher Education policies, *Higher Education*, 65 (1), 135-151

Seeber Marco, Benedetto Lepori, Martina Montauti, Jurgen, Enders, Harry de Boer, Elke Weyer, Ivar Bleiklie, Kristin Hope, Svein Michelsen, Gigliola Nyhagen Mathisen, Nicoline Frølich, Lisa Scordato, Bjørn Stensaker, Erica Waagene, Zarko Dragsic, Peter Kretek, Georg Krücken, António Magalhães, Filipa M. Ribeiro, Sofia Sousa, Amélia Veiga, Rui Santiago, Giulio Marini & Emanuela Reale (2014): European Universities as Complete Organizations? Understanding Identity, Hierarchy and Rationality in Public Organizations, Public Management Review, DOI:10.1080/14719037.2014.943268

Segerholm C. (2001), National Evaluations as Governing Instruments: How Do They Govern? *Evaluation*, 7, 4, 427-438

Segerholm C., Åström E., Governance through Institutionalized Evaluation: Recentralization and Influences at Local Levels in Higher Education in Sweden, *Evaluation*, 2007, 13, 48-67

Van Vught F. (1993), *Patterns of Governance in Higher Education Concepts and Trends*, Unesco, Paris

Weick K. E. (1976, Educational Organizations as Loosely Coupled Systems, *Administrative Science Quarterly,* 21, 1-19

Whitley R. (2007), *Introduction*, in Whitley R. and Glaser J., *The changing governance of the sciences*, Springer, 3-30

## Appendix

### List of questions and merging of the five questionnaires

The full list of variables used for this paper in terms of short name (new labels) and labels (wording) is displayed below. The labels of the variables in SPSS are just the wording of the questionnaires, so they may be long (where too long to be accepted by SPSS, the wording has been slightly adapted thus trying to preserve the original structure of the question). For each variable, the number of the question in each questionnaire can be easily found. If for a variable there is no match in a questionnaire the symbol **ø** is used.

| NEW LABEL | B | S | R | CA | MM | Wording |
|---|---|---|---|---|---|---|
| ACTOR_INF | ø | ø | 25 | 29 | 27 | To what extent do the following actors influence university strategies? |
| COMPARE | ø | ø | ø | 20 | 22 | For which of the following activities, does your university systematically compare different units? |
| COPE | 36 | 28 | ø | 26 | ø | To what extent has your university been able to bring about changes in the following areas during the last 5-10 years ? |
| DECMAK | 31 | ø | ø | 34 | 32 | Which of the following statements best describes how decisions on the budget in your university are taken ? |
| DMP | 39 | 31 | 26 | 46 | 43 | Please indicate the extent to which actors within your university have actual decision-making power for the listed issues. |
| EXTEVA | 18 | 14 | 12 | ø | ø | How would you rate the influence of external evaluation and quality assessment on your university in terms of … |
| FACSTR | 37 | 29 | ø | 27 | ø | How important have the following factors been in realizing the current institutional strategy? |
| IMP | 11 | ø | 7 | ø | ø | How would you rate the influence of the following steering instruments on the decisions made within your university? |
| INFITEM | 17 | 13 | 11 | ø | ø | How would you rate the influence of the following items on the activities carried out by your university? |
| INSTIT | 10 | ø | 5 | ø | ø | How would you rate your level of interaction with the following institutions? |
| PRFUND_1 | ø | ø | ø | 36 | 34 | Which of the following statements best describes how past level of funds influences the budget of institutional units (e.g. faculties) in the next year? |
| MM10 | ø | ø | ø | ø | 10 | How would you rate the impact of teaching evaluation on the following items? |
| MM11 | ø | ø | ø | ø | 11 | How would you rate the impact of research evaluation on the following activities? |
| PERFOR | ø | ø | ø | 18 | 21 | Does your unit have procedures for subunits and/or individual academics to report on their performances? |
| P-FACSTR | 38 | 30 | ø | 28 | ø | How important are the following factors compared to 5-10 years ago? |
| POW.T&R | ø | ø | ø | 13 | 16 | In recent years, how did the power of the following actors within your university change in relation to academic affairs (teaching and research)? |
| R10 | ø | ø | 10 | ø | ø | Over the last five years, did your university participate in evaluations? |
| R24 | ø | ø | 24 | ø | ø | How important have the following factors been in developing the current institutional strategy? |
| R6 | ø | ø | 6 | ø | ø | How important is it for your university to be connected to each of the following institutions? |
| RECT | 21 | 17 | 15 | ø | ø | How would you rate the influence of the university leadership (e.g. rectorate) on the following issues? |
| STR_BUILDING | 34 | 27 | ø | 25 | ø | Please rate your level of agreement with each of the following statements. |

| | | | | | | |
|---|---|---|---|---|---|---|
| **UNI** | **1** | **1** | **1** | **1** | **5** | **To what extent do you agree with the following statements as regards this university?** |
| **W.R.ALLO** | **29** | **ø** | **ø** | **31** | **29** | **Which of the following statements best describes the way resources are allocated to institutional units (e.g. faculties)?** |

Other variables used are: formal autonomy dataset; information about country and description of single HEIs (secondary data and literature information); combination of these data (i.e. accountability).

### Use of "Formal Autonomy" dataset

In order to enlarge the set of variables that may predict the behavior of HIEs, further variables have been taken into account from a descriptive data set filled in 2011 by eight experts of the eight countries involved in TRUE project. These data represent the effective feature by countries. Actual performances of HEIs therefore might be reasonably influenced from these patterns.
Here a comprehensive description and recode of variables are fostered.

**Table. Recode of variables into new dimensions for the characterization of countries**

| Name of variable | Label of the variable | # of values | Dimension |
|---|---|---|---|
| V01 | FREE TO DECIDE ON LEGAL STATUS | 3 | Legal |
| V02 | FREE TO APPOINT ACADEMIC STAFF | 3 | HR Management |
| V03 | FREE TO DETERMINE SALARIES | 4 | HR Management |
| V04 | FREE TO DETERMINE PROCEDURES FOR ASSESSMENT INDIVIDUAL STAFF | 4 | HR Management |
| V05 | FREE TO DETERMINE PROCEDURES FOR ACADEMIC PROMOTIONS | 4 | HR Management |
| V06 | FREE TO DETERMINE HOW TO SPEND PUBLIC GRANT | 3 | Financial |
| V07a | ESTIMATE AVERAGE PROPORTION OPERATIONAL GRANT | Continuous | --- |
| V07b | ESTIMATE AVERAGE PROPORTION TUITION FEES | Continuous | --- |
| V07c | ESTIMATE AVERAGE PROPORTION THIRD PARTY FUNDING | Continuous | --- |
| V07d | CATEGORISED PROPORTION OPERATIONAL GRANT | 4 | --- |
| V08 | FREE TO BORROW FUNDS ON CAPITAL MARKET | 3 | Financial |
| V09 | FREE TO BUILD UP RESERVES | 3 | Financial |
| V10 | FREE TO CHARGE FEES BACHELORS | 4 | Financial |
| V11 | FREE TO CHARGE FEES MASTERS | 4 | Financial |
| V12 | FREE TO CHARGE FEES DOCTORAL STUDENTS | 4 | Financial |
| V13 | FREE TO CHARGE FEES NON-EU STUDENTS | 4 | Financial |
| V14 | FREE TO SET TARIFFS FOR CONTRACT ACTIVITIES | 3 | Financial |
| V15 | OWNERSHIP OF BUILDINGS AND PROPERTIES | 3 | Legal |
| V16 | NUMBER OF CATEGORIES FOR GENERATING PRIVATE FUNDING | 3 | Financial |
| V17 | FREE TO ENTER PARTNERSHIPS WITH HEIs | 3 | External Governance |
| V18 | FREE TO ENTER PARTNERSHIPS WITH NON-HE ORGANSATIONS | 3 | External Governance |
| V19 | FREE TO SELECT OWN BACHELORS STUDENTS | 2 | Policy |
| V20 | FREE TO SELECT OWN MASTERS STUDENTS | 2 | Policy |
| V21 | FREE TO DECIDE ON NUMBER OF STUDY PLACES FOR BACHELORS PROGRAMMES | 3 | Policy |
| V22 | FREE TO DECIDE ON NUMBER OF STUDY PLACES FOR | 3 | Policy |

| | | | |
|---|---|---|---|
| | MASTERS PROGRAMMES | | |
| V23 | FREE TO DECIDE ON RESEARCH PROGRAMMES AND MAJOR THEMES | 4 | Policy |
| V24 | FREE TO START NEW BACHELORS PROGRAMMES | 2 | Policy |
| V25 | FREE TO START NEW MASTERS PROGRAMMES | 2 | Policy |
| V26 | FREE TO START NEW DOCTORAL PROGRAMMES | 2 | Policy |
| V27 | FREE TO SET UP INTERNAL EVALUATION SYSTEMS FOR TEACHING | 4 | Steering instruments |
| V28 | FREE TO SET UP INTERNAL EVALUATION SYSTEMS FOR RESEARCH | 4 | Steering instruments |
| V29 | FREE TO TAKE PART IN EXTERNAL QUALITY ASSESSMENT FOR TEACHING | 4 | Steering instruments |
| V30 | FREE TO TAKE PART IN EXTERNAL QUALITY ASSESSMENT FOR RESEARCH | 4 | Steering instruments |
| V31 | FREE TO DETERMINE INTERNAL GOVERNANCE STRUCTURE | 4 | Internal Governance |
| V32 | FREE TO SELECT EXECUTIVE HEAD | 3 | Internal Governance |
| V33 | FREE HOW TO EVALUATE THE EXECUTIVE HEAD'S PERFORMANCES | 3 | Steering instruments |
| V34 | FREE TO APPOINT MEMBERS OF GOVERNING BOARD | 3 | Internal Governance |
| V35 | FREE HOW TO EVALUATE THE PERFORMANCES OF GOVERNING BOARD | 3 | Steering instruments |
| V36 | REQUIRED TO PRODUCE STRATEGIC PLAN | 2 | Steering instruments |
| V37 | REQUIRED TO ESTABLISH MULTI YEAR CONTRACT WITH MINISTRY | 3 | Steering instruments |
| V38 | OBLIGED TO TO PUBLISH ANNUAL REPORT | 2 | Steering instruments |
| V39 | OBLIGED TO SUBMIT AUDITED FINANCIAL STATEMENT | 2 | Steering instruments |
| V40 | OBLIGED TO SUPPLY INFORMATION DEMONSTRATING COMPLIANCE WITH OTHER NATIONAL POLICIES | 2 | Steering instruments |
| V41 | OBLIGED TO PUBLISH OUTCOUMES EVALUATIONS | 2 | Steering instruments |
| V42 | OBLIGED TO PROVIDE DATA NATIONAL DATABASES | 2 | Steering instruments |
| V43 | NUMBER OF REPORTING OBLIGATIONS | 3 | Steering instruments |
| V44 | DIVISION CORE GRANT FEDERAL AND NATIONAL LEVEL | 4 | --- |
| V45 | CHARACTERISATION NATIONAL FUNDING SYSTEM | 3 | Funding |
| V46 | CHARACTERISATION DEGREE OF DIVERSIFICATION THIRD PARTY RESEARCH FUNDING | 3 | Funding |
| V47 | IMPORTANCE DIRECT NEGOTIATION ON FUNDING 2010 | 5 | Funding |
| V48 | IMPORTANCE DIRECT NEGOTIATION ON FUNDING 1995 | 5 | Funding |
| V49 | IMPORTANCE INCREMENTAL BUDGETING 2010 | 5 | Funding |
| V50 | IMPORTANCE INCREMENTAL BUDGETING 1995 | 5 | Funding |
| V51 | IMPORTANCE FORMULA-BASED QUANTITATIVE INDICATORS 2010 | 5 | Steering instruments |
| V52 | IMPORTANCE FORMULA-BASED QUANTITATIVE INDICATORS 1995 | 5 | Steering instruments |
| V53 | IMPORTANCE PERFORMANCE CONTRACTS 2010 | 5 | Steering instruments |
| V54 | IMPORTANCE PERFORMANCE CONTRACTS 1995 | 5 | Steering instruments |
| V55 | IMPORTANCE NUMBER OF STUDENTS AS FUNDING DRIVER 2010 | 5 | Policy |
| V56 | IMPORTANCE NUMBER OF STUDENTS AS FUNDING DRIVER 1995 | 5 | Policy |
| V57 | IMPORTANCE DEGREES AS FUNDING DRIVER 2010 | 5 | Funding |

| | | | |
|---|---|---|---|
| V58 | IMPORTANCE DEGREES AS FUNDING DRIVER 1995 | 5 | Funding |
| V59 | IMPORTANCE THIRD PARTY GRANTS AS FUNDING DRIVER 2010 | 5 | Funding |
| V60 | IMPORTANCE THIRD PARTY GRANTS AS FUNDING DRIVER 1995 | 5 | Funding |
| V61 | IMPORTANCE RESEARCH OUTPUT AS FUNDING DRIVER 2010 | 5 | Funding |
| V62 | IMPORTANCE RESEARCH OUTPUT AS FUNDING DRIVER 1995 | 5 | Funding |
| V63 | IMPORTANCE RESEARCH TRAINING AS FUNDING DRIVER 2010 | 5 | Funding |
| V64 | IMPORTANCE RESEARCH TRAINING AS FUNDING DRIVER 1995 | 5 | Funding |
| V65 | IMPORTANCE OUTCOMES EVALUATION EXERCISES AS FUNDING DRIVERS 2010 | 5 | Steering instruments |
| V66 | IMPORTANCE OUTCOMES EVALUATION EXERCISES AS FUNDING DRIVERS 1995 | 5 | Steering instruments |
| V67 | IMPORTANCE TRANSFER TO ECONOMY AND SOCIETY AS FUNDING DRIVERS 2010 | 5 | Funding |
| V68 | IMPORTANCE TRANSFER TO ECONOMY AND SOCIETY AS FUNDING DRIVERS 1995 | 5 | Funding |
| V69 | FREE TO DECIDE ON INTERNAL ALLOCATION PUBLIC FUNDS | 5 | Financial |
| V70 | FINANCIAL ACCOUNTNG REQUIREMENTS | 3 | Financial |

*Source: own elaboration on Formal Autonomy dataset (TRUE project)*

Questions 7 are dropped because information about percentages of annual budget are known at HEIs level and differences from the national averages may not be significant as the main value itself. Question 44 is not relevant since few country, namely Germany, are in a federal system, hence transnational comparisons are not so viable. Question 46 is not taken into account since it does not represent an ordinal variable but just a nominal one.
Variables from 47 to 68 included are about the change over time about key assets of HEIs.
No inversion of polarity is required since when the question is about "freedom", the first values are labeled as "no", viceversa whenever there a question about "obligations" the first values are "yes".

| Dimensions | UK | FR | IT | NL | NO | PT | CH | DE |
|---|---|---|---|---|---|---|---|---|
| **external governance** | 1,000 | 0,750 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 0,500 |
| **Financial** | 0,614 | 0,409 | 0,818 | 0,750 | 0,450 | 0,867 | 0,455 | 0,313 |
| **Funding** | 0,250 | 0,250 | 1,000 | 0,500 | 1,000 | 1,000 | 0,500 | n.a. |
| **HR management** | 1,000 | 0,333 | 0,375 | 0,611 | 0,778 | 0,292 | 0,250 | 0,625 |
| **Internal governance** | 0,704 | 0,306 | 0,889 | 0,204 | 0,556 | 0,704 | 0,000 | 0,204 |
| **Legal** | 0,500 | 0,250 | 0,750 | 0,500 | 0,500 | 0,750 | 0,000 | 0,250 |
| **Policy** | 0,875 | 0,125 | 0,125 | 0,500 | 1,000 | 0,375 | 0,514 | 0,625 |
| **Steering instrument** | 0,615 | 0,397 | 0,607 | 0,278 | 0,476 | 0,365 | 0,577 | 0,399 |
| **Total** | 0,757 | 0,415 | 0,821 | 0,605 | 0,845 | 0,794 | 0,474 | 0,416 |

*Source: own elaboration on Formal Autonomy dataset (TRUE project)*

### Deflated variables about evaluation: an example

Example for one of deflated variables. "Instit_DEF": How would you rate your level of interaction with the following institutions? Quality assurance agencies (accreditation/ evaluation)

| Name of the institution | How would you rate your level of interaction with the following institutions? National parliament | How would you rate your level of interaction with the following institutions? National ministry(ies) | How would you rate your level of interaction with the following institutions? Quality assurance agencies (accreditation/ evaluation) | How would you rate your level of interaction with the following institutions? Funding agencies | How would you rate your level of interaction with the following institutions? Local authorities | How would you rate your level of interaction with the following institutions? National associations of universities | Mean of 6 items | Instit_DEF (referred to the third issue) |
|---|---|---|---|---|---|---|---|---|
| **NO1** | 3,63 | 2,88 | 4,00 | 3,75 | 3,63 | 2,88 | **3,461667** | **1,253702** |
| **…** | … | … | … | … | … | … | **…** | **…** |
| **Survey** | 3,511538 | 2,668846 | 3,185 | 2,688846 | 2,755385 | 2,721538 | **2,921859** | **0,829972** |
| **S.D.** | 0,916212 | 0,616163 | 0,780878 | 0,599701 | 0,881416 | 0,626109 | **0,374138** | |

*Source: own elaboration on TRUE dataset*
*Note: In red the values of the issues that are deflated to the other issues.*

**Mean values for the 6 discussed clusters (only variables that entered the analysis; in red the values characterizing the cluster)**

| | IT2 | UK1 | IT3 | NO1 NO2 NO3 DE1 DE2 DE3 CH1 CH2 CH3 CH4 | NL1 NL3 PT2 | IT1 PT1 PT3 NL2 CH5 FR1 FR2 |
|---|---|---|---|---|---|---|
| dmp_6_c | 0,549 | 0,490 | 0,507 | 0,528 | 0,391 | 0,482 |
| dmp_6_f | 0,313 | 0,451 | 0,378 | 0,364 | 0,448 | 0,379 |
| dmp_6_s | 0,138 | 0,058 | 0,115 | 0,108 | 0,161 | 0,140 |
| dmp_11_c | 0,420 | 0,087 | 0,399 | 0,331 | 0,095 | 0,275 |
| dmp_11_f | 0,428 | 0,493 | 0,497 | 0,423 | 0,554 | 0,496 |
| dmp_11_s | 0,152 | 0,421 | 0,103 | 0,246 | 0,351 | 0,229 |
| uni_man | 2,530 | 2,270 | 2,580 | 2,629 | 2,537 | 2,800 |
| uni_coll | 2,280 | 1,870 | 2,260 | 2,374 | 2,130 | 2,273 |
| accountability_ext | 3,340 | 3,470 | 2,780 | 3,155 | 3,133 | 3,150 |
| au_ev_int_qa | 1,000 | 0,333 | 1,000 | 0,900 | 0,333 | 0,619 |
| au_ev_int_re | 1,000 | 1,000 | 1,000 | 0,800 | 0,778 | 0,952 |
| au_steering_instrument | 0,607 | 0,631 | 0,607 | 0,505 | 0,333 | 0,436 |
| au_total | 0,696 | 0,713 | 0,696 | 0,514 | 0,602 | 0,538 |
| actor_inf_def | -1,029 | -0,545 | 1,244 | 0,199 | -0,904 | -0,637 |
| r6_def | 0,000 | -0,243 | 1,414 | 1,265 | -0,557 | -0,089 |
| pfactstr_def | -1,239 | 0,645 | -0,396 | -0,033 | -0,465 | -0,944 |
| factstr_def | -1,442 | 1,190 | 0,582 | 0,583 | -0,271 | -0,674 |
| r24_def | -4,919 | 1,300 | 0,651 | 0,765 | -0,583 | -0,146 |
| imp_3_def | -0,174 | -0,349 | -0,764 | 0,870 | -0,066 | 0,357 |
| imp_4_def | -0,174 | 1,137 | 1,528 | 0,590 | -0,835 | 0,060 |
| instit_def | -0,918 | -0,708 | -0,057 | 1,404 | -0,105 | -0,220 |
| infitem_1_def | -1,420 | -1,428 | -1,068 | -1,064 | 0,002 | -1,168 |
| infitem_2_def | -0,321 | 0,084 | 0,283 | -0,327 | 1,389 | -0,151 |
| infitem_3_def | 1,298 | -0,053 | 1,504 | 0,464 | -0,685 | 0,045 |
| infitem_4_def | 0,443 | 1,397 | -0,719 | 0,927 | -0,706 | 1,274 |
| mm_11_5_mean | 2,000 | 1,250 | 1,571 | 1,907 | 1,867 | 1,882 |
| mm_11_6_mean | 1,815 | 1,250 | 1,714 | 2,051 | 1,867 | 2,013 |
| mm_11_1_mean | 1,778 | 1,500 | 1,857 | 2,153 | 2,011 | 2,134 |
| mm_11_2_mean | 2,346 | 1,250 | 2,143 | 2,262 | 2,456 | 2,402 |
| mm_11_3_mean | 2,370 | 1,250 | 2,429 | 2,234 | 2,322 | 2,457 |
| mm_10_8_mean | 2,741 | 2,000 | 2,571 | 3,078 | 2,767 | 2,756 |

*Source: own elaboration on TRUE dataset*

*Note: With the exception of DMP and AU, the less is the value, the strong is the item*