MINELLI E., REBORA G., TURRI M.

HOW EVALUATION PRACTICE CONTRIBUTES TO THE ENHANCEMENT OF UNIVERSITY RESEARCH. AN ANALYSIS OF THE RECENT ITALIAN EXPERIENCE IN THE EUROPEAN CONTEXT

1. Introduction[1]

The three year research evaluation exercise (VTR) promoted by the Committee for Evaluation of Research (CIVR) between 2001 and 2003 published its results at the beginning of 2006. In Italy this is the first national evaluation experience for research products with the participation of all universities and numerous other research institutions. The exercise is also unique in Italian public institutions in that it directly compares performance in institutions that receive State funding.

Italian universities have thus once again, with even greater impact than before, played the role of an experimental laboratory in the field of evaluation (Rebora, 1999 – Minelli, Rebora, Turri, 2002). This role helps to understand the potentials, the contribution, the limits and criticality of the use of evaluation practices that many people throughout Europe consider to be the basic tool for improving the Higher Education System.

The aim of this study is an initial direct analysis of the main features of the CIVR exercise taking into consideration the entire process, from its conception to its

realisation, in order to weigh up the strong and the weak points, bearing in mind the state of the art of research evaluation in Europe.

2 The three year research evaluation exercise, CIVR 2001-2003

The CIVR was set up by legislative decree 204 in 1998 to promote research evaluation by supporting quality and the enhanced use of national technological and scientific research. The seven members of the committee are appointed by a decree of the President of the Council of Ministers. These members, who include foreigners, have to possess proven qualifications and experience, must be chosen from among a wide range of disciplines and methodologies and may ask for leave if they are civil servants.

The exercise, known as VTR, regards research activities in the three year period 2001-2003. In December 2003 the ministerial decree no. 2206 set the assessment procedure in motion. The universities fulfilled their duties in 2004 and the results were published in the first semester of 2006.

In order to carry out the VTR, the CIVR sets up area committees or panels. Each panel is responsible for one of the 20 scientific-disciplinary areas, 14 CUN areas and 6 special areas[2], on which the evaluation exercise is articulated. The members of the panels are appointed annually by the Ministry of Education, University and Research following proposals from the CIVR and according to suitable criteria of selection.

Altogether there are 151 panellists: 54 per cent from Italian universities, 12 per cent from Italian research institutions, 9 per cent from Italian industries, 25 per cent from foreign universities and institutions. Each panel elects a president from among its members who convenes meetings and ensures that activities are carried out correctly.

The panellists are bound by a code of conduct by which they must:

- perform their duties autonomously of the universities that are being assessed;
- follow the procedures and indications of the CIVR;
- guarantee confidentiality of procedures.

The evaluation exercise aims to evaluate the research activities in State and legally-recognised Italian universities, public research institutions (DPCM 593/1993), ENEA and ASI and other public and private institutions that request the evaluation exercise.

This paper focuses on the 77 universities involved in the VTR and not on research institutions. This is mainly to avoid complication over names and terms, even though the performance in the research institutions is the same as the universities.

The evaluation procedure adopts a peer review system based on assessment of the merits of certain research products indicated by the universities[3]. The exercise (CIVR, 2004a, 2004b and 2004c) is articulated in three consecutive stages under the auspices of the universities, panels and CIVR respectively. The evaluation procedure requires the research institutions to select and send the CIVR a number of research products that corresponds to 25 per cent of the total number of permanent academic staff (researchers, associate professors and professors). Research products include chapters in books, articles published in scientific journals, patents, projects, compositions, drawings and designs, performances, shows and exhibitions, manufactures and works of art. Works that are purely editorial or exclusively for teaching purposes are not considered research products.

The universities themselves select the research products and the evaluation nucleus must certify that the selected products exist. The CIVR has, however, laid down guidelines for selecting the products (also through the informative system it has

adopted). In addition, each evaluation nucleus has to prepare a comprehensive report on the research activity.

During the evaluation procedure the president of each panel gives every panellist a maximum of 150 research products to assess. The panellists do not personally carry out the assessment but assign it to external referees proposed by the panellists themselves and then appointed collectively by the panel. The informative system then informs all the panel members on how the products for assessment have been allocated. On average each referee is given 5 research products and has to make an assessment of merit for each one.

The panellists have to ensure that the evaluation process of the products is carried out correctly and promptly until the referees give the rating. The CIVR and the panels make sure the rating stays anonymous. 6661 referees were involved and included:

- 59 per cent from Italian universities;
- 22 per cent from foreign institutions;
- 17 per cent from Italian research institutions;
- 2 per cent from Italian firms.

Each product is assessed by at least two referees who formulate a descriptive evaluation according to the following criteria:

- quality, which expresses the opinion of peers on the positioning of the product in terms of scientific excellence in the scale of ratings laid down by the international scientific community;
- importance, which expresses the added value for progress in the sector, science in general and also for the ensuing social benefits;

- originality of innovation, which expresses the contribution to progress in the sector in question;
- international standing and international competitive potential, which expresses the position in the international scenario in terms of importance, competitiveness, editorial circulation and the recognition by the scientific community;
- occupational-economic impact (as far as the application of the results is concerned) which expresses the economic, social, cultural and also potential effects of the product.

Each referee is required to synthesise his opinion in a comprehensive assessment of the product according to one of the four levels shown in Table 1.

Table 1

The presidents of the panels check whether identical products exist elsewhere and ensure that assessment is homogeneous.

After evaluation by the referees, each panel, divided into sub-panels if necessary, prepares a final report in three parts:

1. Consensus report, where the referees' assessment of the individual research products is critically re-examined and a single assessment is given (excellent, good, acceptable, limited).
2. Ranking list, where the universities are listed according to their rating. Each research product is given a numerical appraisal depending on the rating obtained (excellent: 1; good: 0,8; acceptable: 0,6; limited: 0,2). The university's rating in a specific area equals the mean of the score obtained by its products selected in

that area. The ranking list gives each university's performance in the area, grouping the results according to the number of products selected (mega structures, over 74 products – large structures between 25 and 74 – medium sized structures between 10 and 24 and small structures, less than 10 products). Within each grouping the universities are first ranked according to their rating, second according to the percentage of excellent products and third on the basis of the degree of ownership of the excellent products.

3. Final report on the area, where the methodology and organisation of the work of the panel (including the setting up of any sub panels) are synthesised, the strong and weak points of the area are identified, the elements of applicative valorisation are further investigated and possible specific actions for improvement are suggested.

Lastly, the CIVR prepares a final report based on the panel reports[4]. This report takes into account the data transmitted by the universities and the reports of each evaluation nucleus. It is articulated on:

- assessment of merit expressing each University's scientific productivity and ability to manage the development and spin-off of research;
- evaluation of the area and global merit of the national research system, that results from the integration of the elements of analysis in the hands of the CIVR.

3 First results of the evaluation exercise and their uses

The use of the results and, above all, the consequences on the funding of the structures are extremely important in terms of the role played by the research evaluation system.

The experience of the RAE in England where the panels express a synthetic appraisal with a scale from 1 to 5 for each research institution is a model for Europe. The final assessment is made without the panels ever visiting the university and is used with different criteria by the four main public-funding bodies to decide on the amount of funds to be allocated to research. For example, the HEFCE fund, unlike the RAE 2001, did not give funds to institutions with a rating of 1 and 2 whereas an institution with a rating of 5 received approximately 4 times more funds (for the same amount of research) than an institution with a rating of 3.

However, in Europe there are also very different concepts of evaluation. One only has to think of the Dutch system which has existed since 1993 and was significantly modified in 2003. Here, the three-year cycles of self assessment are integrated every six years with exercises carried out by external panels. There are no links between results and university funding whereas great emphasis is placed on self-assessment and in general on the initiatives of the individual universities which organise the evaluation activities of the panels and ensure that the results are published. The national authorities, on the other hand, limit themselves to the role of meta-evaluation guaranteeing the overall correctness of the system.

Peer review is the methodological basis shared by the English RAE, the Dutch system and the VTR of the CIVR but the impacts are completely different.

In the Italian exercise the uses of assessment have still not been made absolutely clear. The official documents regarding the exercise state that (CIVR, 2004b): “when

allocating public funds to research activities, priority will be given to those structures that have taken part in the evaluation exercise".

However, at present the main source of the State university funding, the FFO, which makes the annual transfer of the necessary funds for running the universities, allocates a larger sum on the basis of well consolidated experience. The remaining sum is allocated according to criteria specially formulated by the National Committee for Assessment of the University System (CNVSU) which stipulate that 30 per cent of these funds will depend on the results of scientific research activities. The Ministry of Education, University and Research for the 2006 exercise[5] is oriented towards using the results of the VTR when allocating these funds.

The first results of the VTR were presented during a press conference held at the ministry of Education, University and Research at the end of January 2006 and posted on the CIVR website.

At the same time, the results of the assessments formulated by the area panels were given. These included:

- a final report for each area;
- the rating of each university;
- the ranking list of the universities subdivided into classes according to the number of selected products.

The Rector of each university was then confidentially informed of the evaluation given to each research product.

To date (April 2006), the use of evaluation results appears to be confined to diffusion and publication of information since programmes on how they will affect the allocation of funds are not yet operative.

The posting of the data on the website and their presentation at the press conference highlight the trend that the CIVR intends to follow, despite the fact that it has no say in research funding. Because of the novelty of the evaluation exercise for the university system, communication is also extremely important and is linked to the natural dialectics between universities, national governments bodies and society.

The preliminary phases of the VTR involved academics who were asked by their universities to set up the area committees for selecting products and to act as referees in the evaluation process. So far, the interest shown in the results of the exercise is a good sign in that it highlights the universities' intention to improve the quality of scientific production and the research process in general. The comparative presentation of results also encourages the universities and departments to make improvements. The ability of evaluation results to stimulate improvement will be verified over the next few months and years.

In any case the widespread publicity has led to great expectations regarding the pre-announced use of the VTR results for the allocation of funds. This could affect both the university system whose funds are allocated by the FFO and systems inside the universities that allocate funds to the various departments.

4 Light and shade of an innovative experiment

The CIVR exercise must be given credit for presenting (for the first time in Italy) an overall picture of the results of research activities. These results are shown in Table II. In the case of Italian public institutions the CIVR made the unprecedented decision to

post on its website the rating obtained in the 20 areas by all the universities and then placed the universities' performance in a ranking list. This is something completely new compared to the lack of courage often shown by many evaluation systems, including those in universities (Turri, 2003; Minelli, Rebora, Turri, 2005) whose results give no indication whatsoever of merit. Moreover the fact of using peer review, based on the assessment of research products by high-level experts, means that evaluation is now a legitimate part of the university tradition and has consequently been accepted by the universities with very little opposition. The VTR has succeeded in introducing sound elements of judgment in the Italian university environment which is traditionally founded on collegiate harmony that seeks for consensus.

In the case of VTR it has been fundamental not to contradict the principles at the basis of university culture but to formalise evaluation practices that already existed in the university tradition through an exercise articulated on participation in the selection of research products.

In fact, although there were no national evaluation experiences in the past, researchers have always submitted their products to evaluation procedures. One only has to think of the referee procedures for publication of articles in scientific journals, the competitiveness to obtain research funds or appointment to an academic post after comparative evaluation.

Basically the CIVR procedure has shifted evaluation from that of a personal initiative to a national exercise. Even if this shift at first has no direct consequences on funding and the power system, it greatly affects the culture of the university organisation by laying the basis for a change in individual behaviour and operative

mechanisms. It has activated a series of dynamic processes that offer openings for a whole series of potential change agents.

When considering the natural elements that make up an evaluation system for complex activities, that is to say concepts, methods, bodies and uses (see Figure 1), the promoters of the exercise have made simple, safe choices, minimising the risks of delegitimisation and making it possible to speed up the process.

Figure 1

At the basis of the system were the bodies that were set up and the operative methodologies. Once the CIVR was constituted the people in charge opted for panels and referees following in the steps of the European tradition. Peer review was the guarantee of a system that was widely accepted. Less attention was paid to operative procedures and in any case they were conditioned by the main purpose which was to act quickly and achieve a result that could be presented at a national level.

The idea or concept of evaluation may be criticised for its limited elaboration or excessive simplicity. It has been decided not to dwell on the subject and play safe rather than go into infinitely long discussions on the concept of scientific quality. A general concept has been accepted with reliance on the panels' evaluative competence: "scientific quality will be what the panel considers it to be". Similarly, they have avoided directly facing the critical question of funding, since, from this point of view, the general guidelines expressed by Ministry are sufficient.

Those in charge of the CIVR have made great efforts to encourage widespread direct communication with members of the different scientific communities involved in each stage of the setting-up and management of the exercise. Right from the start, the

VTR and CIVR have shaken off the label of being “a mysterious object” and have made themselves known to all important interlocutors.

In the change process an important role is played by player and agents so it was obvious that the central bodies and panellists were not sufficient. Thus, people in charge of research activities throughout the country: directors of departments, institutions, research centres and deans have become active agents. Evaluation was to be seen as a dynamic lever for stimulating their own specific structures, creating new openings, reorienting groups and people, starting all over again if necessary and reallocating funds. Some actors understood the rules of the game and its potentials and acted as a result but others understood very little and remained inert and passive. Even the selection of research products indicated by the evaluation exercise has meant that decision-making is oriented to certain products rather than others. In the long term this type of behaviour means that everyone is responsible for his own personal behaviour, decision-making capability in the university departments is enhanced and shifty behaviour is overcome.

The evaluation exercise has led many universities to question their ability to carry out quality research and to make a comparison with other universities. This is the first important step towards strategies of valorisation and differentiation between universities. These strategies foresee that the universities and departments have greater say in management with the definition of policies for improving the quality of research. In addition, the international character of the exercise is fully aligned with the requisites of the process for the integration of education and university research in Europe and satisfies a widely-felt viewpoint.

Overall, the university and research world has responded well to the VTR and CIVR, accepted positively the challenge of evaluation and has made it a resource for innovation and change.

In this perspective of change, technical imperfections are not so important although in the medium term they make room for the improper effects and unexpected undesired consequences that reference theories take into consideration (Smith, 1995 – Power, 1997 - Rebora, 1999 - Van Thiel and Leeuw, 2002 - Turri, 2005). The following is a summary:

1. The fact that the current evaluation exercise involves a small number of products means that the VTR does not assess the overall scientific production in Italian universities but only assesses the universities' ability to produce a certain number of research outputs that are recognised as being of international excellence. This is a legitimate methodological choice but one must be fully aware of it when interpreting the results. It is not possible or appropriate to expect that the exercise will give indications on the performance of the university staff, nor on the ability to produce volumes of research activities that are in line with the funds used. The exercise only refers to excellence and must be considered in this light. Consequently, any considerations on the results of the exercise must bear in mind that the reports show the peaks of excellence in the national system but are of no use in other matters such as the productivity of the university staff during the period that is being examined. On the contrary, the relationship between the results of the VTR and funding should take into consideration:
   - the proportion between permanent staff and selected research products, which should include more products in the evaluation exercise and ensure that there is

a closer link between active researchers in the research structures undergoing evaluation and the number of selected products;

- the number of products that the university can present in each area;
- the degree of ownership of the selected products since some have more than 20 authors of whom only one is actually from the university presenting the product;
- the number of products presented by a single author, as there are currently no limits. This brings about obvious distortions since the proportion between products that can be presented and university personnel is constant.

2. Another drawback regards the criteria for drawing up the ranking list which puts universities of the same size in order of the absolute rating they have achieved. If the rating is not well balanced there is a risk that the results will be distorted. If we take a practical example the situation is the following (Table IV): in area 10 (Philological-literary sciences, antiquities and arts) in the section referring to large structures the University of Venice and the State University of Milan have a similar number of researchers but selected a different number of products. However following the VTR criteria, the University of Venice legitimately reduced the number of products for evaluation from 49 (that corresponded to 25 per cent of the researchers in the area) to 29. This greatly increased the incidence of the total of selected excellent products and therefore had a positive effect on its rating. Milan, on the other hand, selected 48 research products corresponding exactly to the researchers in the area but although it had a larger absolute number of excellent products was given a lower rating which put it 9 places behind Venice in the ranking list. Although at present there is no proof that the universities use particular strategies when submitting products in order to maximise the outcome of the

exercise, it is essential that any future evaluation exercises pay more attention to the prevention of possible distortions.

Table 2

3. The way the ranking list is structured is also questionable. Although the decision of ranking of university results is officially approved in the document presented at the beginning of the procedure, the criteria on which the ranking was to be based have been completely ignored[6]. Moreover, the sub-division of the structures according to the number of selected products seems to be arbitrary in the case of a product that is assessed as excellent when the proportion between the number of researchers and selectable products is fixed. The only justification appears to be the need to increase the number of prestigious positions available. The solution adopted by the CIVR increases the available positions on the podium so up to 12 structures in each area can claim that they have reached one of the top three positions. In addition, this criteria favours universities in the mega-structure category because as there are very few they can easily reach the top positions.
4. Another drawback is the difficulty in making comparisons between the research areas due to the fact that the panels use different criteria. The structure of the rating system is extremely unhomogeneous and this falsifies an overall consideration of the ranking obtained in different areas. The difficulty in making a comparison hampers the potentially valuable initiatives for comparing one university with another.
5. The special areas distort the picture. Whereas the number of researchers in the 14 CUN areas indicates how many products can be selected, in the special areas the

university's desire to select products leads them to collocate researchers fictitiously in a particular area. This mechanism confuses the exercise by making it possible to cover up the different ratio between the number of personnel in one area and the number of selected products. It would have been far better either to put the special CUN areas on the same level as the other 14 thus making it possible to appoint personnel directly, or consider them as part of the 14 traditional areas and give the task of assessment to specially set up sub-panels.

6. If the institutions that have been evaluated want to make a study of the present three year period in order to plan improvements in the period to come, they will be unable to do so, because the time for making changes is limited to one year. It would be more appropriate to increase the amount of time for collecting products, so that all those involved could see the outcome and decide whether to make any changes. In practice, the way in which the exercise is articulated over time risks weakening its chances of achieving more direct results.
7. The last weak point, but certainly not the least important, is that the exercise pays too little attention to transparency and the complete autonomy of assessment procedures. The efficient and determined line of action of the CIVR has shortcomings in terms of transparency:
   - In theory the procedure for proposing possible candidates for the panels, who are essential in the evaluation exercise, is to be an open public one. However, in practice the academic community was kept in the dark and was only given information on the panel when it was already set up with no details of how the panellists were chosen. Moreover, the CIVR document stated that there were to be between 5 and 9 panellists but in some cases this number rose to 17. It would

have been preferable to increase the number of foreign experts (at present limited to 25 per cent), as over 75 per cent of the selected products is in English and to recognise some sort of approval in the choice of panellists by the scientific community.

- In the case of the 6661 referees who were chosen to give a rating (just 22 per cent were foreigners), the Committee decided to keep their names secret on the pretext of wanting to safeguard the anonymity of the evaluation. On the one hand, this decision may be endorsed but on the other it means that the panels are endowed with great responsibility so there must be absolute correctness and transparency when they are set up.
- The overall results of the exercise were given widespread publicity whereas the detailed results for each university (i.e. the rating of the individual research products) were only given to the Rector and, what is more, were synthesised. Even the authors were not given the chance to know their own rating directly.
- In the same way that the criteria for making the ranking list and the number of panellists changed during the exercise so did the evaluation regulations laid down in the official CIVR documents. The universities involved in the VTR were in no way informed nor was any reason given for the changes.

The process also reveals large shady areas in terms of autonomy. In fact, many of the members of the CIVR and the area panels have posts (some of which are even very important) in the universities that are being assessed. In addition, they have also submitted research products (often in large numbers) to the evaluation exercise in which they have personally had a say in the evaluation criteria and have chosen and monitored the referees.

5. Conclusions

In order to understand the overall impact of the exercise on the Italian universities much will depend on whether it is repeated in future as the term “three-year research evaluation” implies. Repetition of the evaluation exercise and improvements in the weak areas would have important consequences on the universities, otherwise the great ferment caused by the VTR is destined to fade away in a short time.

The relationship between the evaluation impact and time factors is the central point in Jeliazkova and Westerheijdens’ paper (2002) and was taken up by Turri (2005) who puts forward the theory that the impact varies according to which stage in the life-cycle the evaluation process is in (introduction-consolidation-full development-maturity). If the VTR exercise were repeated, universities would find themselves in a completely new situation. They would be forced to pass from one situation where they had no idea of the result because the exercise was new to one where they were aware of the cyclical repetition of evaluation and would thus be spurred on to activate strategies and behaviour enhancing the quality of their research.

To conclude, two considerations are worthy of note in the light of European experiences. The first regards the absolute methodological coherence of the VTR with European exercises through the adoption of an evaluation mechanism based on peer review. Peer review is the key element in the evaluation process and is responsible for the overall positive opinion diffused about it. The emphasis that Italian universities place on the evaluation procedures hinges on peer review which also acts as a link

between the Italian experience and the main evaluation experiences throughout Europe[7].

Thus, by using peer review the CIVR procedure makes a fundamental contribution non only to the university evaluation scenario but also to public administration in general. The VTR confirms the issues highlighted by literature i.e. the evaluation practice cannot be limited to a mechanical exercise, especially in non-standardisable environments (Noordegraaf and Abma, 2003; Popper and Wilson, 2003; Baker and Hayes, 2004). Automatic mechanisms, statistical surveys and indicators will never substitute a critical judgment (which finds its natural home in human mediation and elaboration) but may be helpful when making one. Evaluation is therefore not an activity that can be entrusted to automatic sterile mechanisms or to standard procedures but is an act of intelligence which requires one to take on critical responsibilities (Power, 2003, Rebora, 2003). So, more attention must be paid to professional accountability that is more capable of operating in highly autonomous operative contexts (Pomzek, 2000; Huisman and Currie, 2004) and to the importance of organisational culture in evaluation (Kunda, 1992; Turri, 2005). Even if the VTR is promoted after an overall comparison with European experiences, there is nonetheless a warm call to overcome the present limitations of the exercise.

The more advanced evaluation exercises show that there are marked phenomena of distortion with the repetition and consolidation of the exercises. In the English system where there are strong links with funding, many opportunistic or in any case improper effects and strategies have been singled out by literature (Elton, 2000; Talib and Steele, 2000; Morgan, 2004; Sharp and Coleman, 2005; Taper and Salter, 2003; Turri, 2005).

On the part of the organisations that are being assessed there is a strong tendency to develop perverse learning behaviour, once they have understood the evaluation mechanisms that are finalised to using any means possible to maximise results (Smith, 1995; Power, 1997; Van Thiel and Leeuw, 2002; Rebora, 2003; Turri, 2005). The stimulus to act in this way increases exponentially when there are connections between evaluation reports and the mechanisms for funding research activities.

As shown in this article, the present structure of the VTR CIVR exercise is extremely fragile. The novelty of the first edition may have prevented the large scale diffusion of improper phenomena but this situation must be resolved before any future exercises take place since there is the risk that the impact of the weak points will increase in the next exercise and compromise the whole outcome.

Table 3

Bibliography

Baker C.R. and Hayes R.(2004), 'Reflecting form over substance: the case of Enron Corp.', *Critical Perspectives on Accounting*, **15**, 767-785.

Borgonovi, E. (2002), *Principi e sistemi aziendali per le amministrazioni pubbliche*, Milano: Egea.

CIVR (2004a), '*Linee Guida per la Valutazione della Ricerca'*. Available on www.civr.it.

CIVR (2004b), '*Valutazione triennale della ricerca (VTR). Bando di partecipazione all'esercizio 2001-2003*'. Available on www.civr.it.

CIVR (2004c), '*Adempimenti dei Comitati (Panel) di Area*',. Available on www.civr.it.

CIVR (2004d), '*Segnalazione di esperti per la costituzione dei Comitati di Area (Panel) per l'esercizio VTR 2001-2003'*. Available on www.civr.it.

CIVR (2004e), '*Linee guida per i comitati (Panel) di Area per l'esercizio 2001-2003'*. Available on www.civr.it.

CIVR (2006a), '*VTR 2001-2003. Risultati delle valutazione dei panel di area'*. Available on www.civr.it.

CIVR (2006b), '*VTR 2001-2003*. Available on www.civr.it

Commission of The European Communities (2000), 'Communication from the commission to the council, the European parliament, the economic and social committee and the committee of the regions. Towards a European research area'. Available on www.europa.eu.int/comm/research/era/index_en.html

Elton, L. (2000), 'The UK Research Assessment Exercise: Unintended Consequences', *Higher Educational Quarterly*, **54** (3), 274 – 283.

Gareth, R. (2003), Review of research assessment Joint funding bodies'. Available on: http://www.ra-review.ac.uk/

Hood, C. (1991), *'A public management for all seasons'*, Public Administration, **69**, 3-19.

Jeliazkova, M. and Westerheijden, D. (2002), 'Systemic adaptation to a changing environment: Towards a next generation of quality assurance models', *Higher education*, **44**, 433–448.

Kunda, G. (1992), Engineering *Culture: Control and Commitment in a High-Tech Corporation (Labor & Social Change)*, Philadelphia: Temple University Press

Minelli E., Rebora G. and Turri M., (2005), '*The risk of failure of controls: the exam of two Italian public sectors*', paper presented at the 3rd Conference on performance measurement and management control improving organizations and society, Nizza, 22-23 September 2005.

Minelli, E., Rebora, G. and Turri, M. (2002), *Il valore dell'università*, Milano: Guerini and Associati.

Ministero dell'Istruzione, dell'Università e della Ricerca, Comitato Nazionale per la Valutazione del Sistema Universitario (2005), Il modello per la ripartizione del Fondo di Finanziamento Ordinario (FFO) all'interno del sistema universitario: riflessioni a valle dell'applicazione sperimentale prevista dal D.M. 28 luglio 2004. Doc. 4/05. Available on: http://www.cnvsu.it

Morgan, K.J. (2004),*'The research assessment exercise in English universities, 2001'*, Higher Education, **48**, 461-482.

Noordegraaf, M. and Abma, T. (2003), 'Management by measurement? Public management practices amidst ambiguity', *Public Administration*, **81** (4), 853-871

Pollit, C. and Bouckeaert, G. (2000), *Public Management Reform*, Boston: Oxford University Press.

Power, M.K.(1997), *The audit society*, Oxford: Oxford University Press.

Power, M.K.(2003), 'Auditing and the production of legitimacy', *Accounting, Organizations and Society*, **28**, 379-395.

Propper, C. and Wilson, D. (2003), The use and Usefulness of performance Measures in the Public Sector, CMPO Working Papers Series N. 03/073. Available on: http://www.ra-review.ac.uk/

Rebora G., Minelli E. and Turri M., '*The evaluation as central element of the reform process?*'. paper accepted, presented and discussed at Workshop on the process of reform of the university across Europe EIASM, University of Siena, 24/26 may 2004.

Rebora, G. (1999), *La valutazione dei risultati delle amministrazioni pubbliche*, Milano: Guerini and Associati.

Rebora, G. (2003), 'La valutazione degli atenei: alla ricerca della qualità', *Azienda pubblica*, 16 (5-6), 411-436

Romzek, B.S. (2000), 'Dynamics of public accountability in an era of reform', *International Review of Administrative Sciences*, **66** (1), 21-44.

Ruffini, R. (2004), *Fondamenti di economia delle aziende e delle amministrazioni pubbliche*, Roma: DeAgostini Professionale.

Sharp, S. and Coleman, S. (2005), 'Rating in the Research Assessment Exercise 2001 – the Patterns of University Status and *Panel* Membership', *Higher Education Quarterly*, **59** (2), 153-171.

Smith, P. (1995), 'On the unintended consequences of publishing performance indicators in the public sector', *International Journal of Public Administration*, **18**, 277-310.

Talib, D. and Steele, A. (2000), 'The Research Assessment Exercise: strategies and trade-offs', *Higher Education Quarterly*, **54** (1), 68-87.

Tapper, T. (2003), 'Interpreting the process of Change in Higher Education: the case of Research Assessment Exercises', *Higher Education Quarterly*, **57** (1), 4-23.

Tapper, T. and Salter, B. (2003), 'Interpreting the Process of change in Higher Education: the case of the Research Assessment Exercises', *Higher Education Quarterly*, **57** (1), 4-23.

Tapper, T. and Salter, B. (2004), 'Governance of Higher Education in Britain: the significance of the research assessment exercises for the funding council model', *Higher Education Quarterly*, **57** (1), 4-3.

Taut, S. and Brauns, D. (2003), 'Resistance to Evaluation: A Psychological Perspective', *Evaluation*, **9** (3), 247-264.

Turri M. (2005), *La valutazione dell'università. Un'analisi dell'impatto istituzionale e organizzativo*, Milano, Guerini

Turri, M. (2002), 'Le esperienze di valutazione nelle università italiane: un'analisi critica, in Minelli, E., Rebora, G. and Turri, M., *Il valore dell'università*. Milano: Guerini e Associati, pp. 65-96

Turri, M. (2003), 'Lo stato dell'arte, limiti ed evoluzione nella valutazione a livello di sistema e di singolo ateneo', *Azienda pubblica*, XVI (5-6), 439-462.

UNESCO (2005), *Towards Knowledge Societies*. Available on www.unesco.org/publications

VanThiel, S. and Leeuw, F.L. (2002), 'The performance paradox in the public sector, *Public performance and management review*', **25** (3), 267-281.

VNSU, NOW, KNAW (2003), *Standard Evaluation Protocol 2003-2009 for Public Research Organisations*. Available on www.qanu.nl

Von Tunzelmann, N. and Kraemer, M. (2003), Changes in research assessment practices in other countries since 1999: final report. Available on www.qanu.nl

Notes

[1] Grateful thanks are due to Prof. Pieter De Meijer, former Rector of the University of Amsterdam and currently President of the Evaluation Nucleus of the University of Padova and Roma Three, for the stimulating debates on VTR.

[2] The Areas correspond to 14 scientific-disciplinary areas defined by the National University Committee (CUN) and 6 special Areas chosen by CIVR on the basis of their importance in Italy and their coherence with the aims of the National Research Programme (PNR) and the E.U. research and development programmes (Science and technologies for an information and communication society, Science and Technologies for quality and safety of food, Sciences and technologies of nano/microsystems, Areo-spatial science and technologies, Sciences and technologies for sustainable development and governance: financial, social, energetic and environmental aspects, Sciences and technologies for the evaluation and enhancement of cultural heritage).

[3] As well as the evaluation of research products in universities and research institutions, the official documents foresee the prospective and retrospective evaluation of research products that come within the PNR (National Research Programme) and are the subject of specific funding. However, the relevant panels have not yet been set up and so the process is not operative.

[4] Up to April 2006 this report has not been issued.

[5] This transpires from the scheme of the decree for the allocation of the FFO 2006 transmitted by the MIUR to the CRUI.

[6] The reported data correspond to what the panels have effectively done. However , the official documents lay down that the universities be collocated in the ranking lists (according to the sector of competence of the panels) in the merit ranking shown below.

- A: At least 50 per cent of the products is assessed as excellent and the remaining as good

- B: At least 30 per cent of the products is assessed as excellent and the remaining as good
- C: At least 50 per cent of the products is assessed as excellent or good and the remaining as acceptable
- D: At least 30 per cent of the products is assessed as excellent or good and less than 50 per cent as limited
- E: At least 20 per cent of the products is assessed as excellent or good and less than 50 per cent as limited
- F: Less than 20 per cent of the products is assessed as excellent or good and less than 50 per cent as limited
- NV (Not assessable): More than 50 per cent of the products is assessed as

In the ranking list the universities were also to be ranked according to:

- the number of products selected by the structure for each specific area
- the degree of mean ownership of the selected products.

[7] When considering the changes that have taken place in the RAE since its introduction in 1986, Tapper (2003) acknowledges the fact that the true identity of the exercise is to be found in peer review.