Huisman J – Rebora G. – Turri M.

THE EFFECTS OF QUALITY ASSURANCE IN UNIVERSITIES: EMPIRICAL EVIDENCE FROM THREE CASES

Abstract

The paper focuses on the impact of evaluation in the university organisation using a systemic approach with particular attention to the effects of evaluation at an institutional and organisational level. The impact of evaluation is divided into three areas: organisational learning, resource development and power management. Through the study of three European universities the impact of evaluation has been empirically recorded.

In conclusion the important role of quality culture and organisational bodies on the determination of the impact of evaluation is investigated. The analysis pinpointed certain factors that are fundamental for understanding the effects of evaluation. These include the importance of peer review, difference between teaching and research, significance of time factors and the risk of the degeneration of evaluation.

1. Introduction

In Europe increased attention has been paid to the consequences and quality of evaluation in the past few years. One result of the Bologna Process is the diffusion of evaluation in European countries (Schwarz and Westerheijden, 2004). Where national evaluation systems have developed it is particularly necessary to go beyond a study of methods and uses and to start discussing the effects of evaluation.

The institutionalisation of evaluation through evaluation exercises and changes in systems and methods, as for example in teaching in the UK (Brown, 2004) shows the need to understand its effects and consequences. The issue is interesting in Italy which is in the early stages of the diffusion of evaluation and is keen to know the consequences of assessment procedures when choosing a system.

Questions like "How effective are these measures?" (Huisman and Currie, 2004) regarding the contribution of evaluation systems to the improvement of quality are becoming more frequent.

However, there is a real gap in the knowledge and understanding of the effects of evaluation. After analysing the evaluation experience in seventeen different countries Brennan and Shah (2000a, p.9) assert that "[…] there is a growing literature on approaches and methods of quality assessment at both institutional and national levels, very little of it addresses the effects of quality assessment on the educational and organisational process in higher education ". Referring to the situation in Britain, Underwood (2000, p.87) holds: " there is a lack of empirical evidence about exercise […] in terms of its effectiveness". Understanding the impact of evaluation is vital for its development. The time when evaluation was blindly accepted without ascertaining its real impact seems to be coming to an end. Many people are starting to question the system and ask "what do audits produce?" (Power, 2003). The answers are often not reassuring: "Most impact studies reinforce the view that quality is about compliance and accountability and has contributed little to any effective transformation to make it more appropriate" (Harvey and Newton, 2004). There is a widespread feeling that evaluation is more closely connected to the need to justify its existence than to its real contribution to activities (Power, 1997 and 2003). Some authors go so far as to say that it drives academic activities to even further standardisation. According to Hayes and Winyard (2002) quality assurance practices lead to the creation of modern forms of bureaucracy where the old certainties foreseen by long-since abolished rules and regulations are substituted by the search for objective evaluative mechanisms that produce legitimacy. It is precisely by examining the true impact of evaluation in the universities that the question can be answered. The issue shifts from an idealistic level, influenced by a priori opinions and values to the real state of affairs.

The present study aims to analyse the institutional and organisational impact of evaluation in universities paying particular attention to its impact on teaching and research. The analysis takes in to account one university in each country rather than the country as a whole.

Only by examining a real situation in this way is it possible to understand the true outcome and impact of evaluation.

This paper aims to:

- put forward a theoretical framework for analysing evaluation systems
- present a framework for reviewing the evaluation impact in universities
- analyse the impact in three different universities: one in the U.K, one in the Netherlands and one in Italy
- show evidence and comment on determinants of the impact of evaluation

The study is exploratory and no definite assumptions will be made. The aim is to investigate not confirm the situation and make some observations on which future studies can be based.

Section 2 discusses the theoretical framework adopted and section 3 deals with methods and the gathering of empirical data in the universities of Bath (United Kingdom), Amsterdam (Netherlands) and Trento (Italy). Sections 4 and 5 deal with the actual evaluation system in the three universities paying particular attention to the impact of evaluation in each. In section 6 evidence and reflections on determinants of evaluation impact are put forward.

In this paper the word "impact" is used to indicate the overall consequences or outcome of evaluation on institutions. The single consequence of evaluation is defined as "effect" and the sum of all these effects gives the impact of evaluation.

2. Theoretical framework

In order to achieve a critical examination of the evaluation system in universities, a consolidated analytical scheme, based on a systemic approach has been adopted (Rebora, 1999; Minelli, Rebora, Turri, 2002; Rebora, 2003; Turri, 2003; Turri, 2005). People involved in the evaluation process too often limit the concept to just one aspect and lose sight of the others. Evaluation and evaluation methods do not coincide, nor can evaluation be considered as a simple methodological technique. The evaluation process does not only correspond to an official formulation either but also to the way in which it is carried out.

A complete understanding of evaluation requires awareness of its complexity and its various components.

The analysis is based on systemic examination of the following five factors:

1. the evaluation idea or concept, that is the explanation or official statement regarding evaluation. This corresponds to the official concept derived from formal written or oral declarations and reflects only the ideal publicly declared purpose, contents and process, which probably differ from the theory-in-action, i.e. the concept revealed by the real practice of evaluation.

2. evaluation methods. These are the specific instruments and techniques that give substance to assessment exercise.

3. the bodies responsible for evaluation, how they are made up and their relationship to the institutional and organisational structure. They not only have to carry out their duties according to the methods devised for evaluation (for this purpose special units can be created) but also accept responsibility for evaluation, direct evaluative procedures and answer for the outcome.

4. the real use made of evaluation output. Evaluation results are basically used in three ways:
    o cognitive: to inform and diffuse knowledge of the evaluated activities
    o as a stimulus for improving the quality of activities and motivating people
    o to give awards, benefits, sanctions and also ensure that behaviour is correct at the individual, collective and organisational level
5. the coherence of the system as a whole, that is to say the degree of connection and homogeneity between the above-mentioned factors. This is not predefined but is a question of finding a balance between the different factors involved in evaluation

These are five essential elements for understanding the structure of evaluation systems after which its impact on the university institution will be analysed[*].

The proposed scheme points up and shows the impact by dividing it into the three categories that were developed in the studies on organisational change (Argyris and Schon, 1978 - Pfeffer, 1981 - Mintzberg, 1983 - Tichy, 1983 - Argyris, 1985 - Beer and Nohria, 2000), that is to say the impact on organisational learning, development of resources and power management.

To complete the examination of evaluation impact two other categories must be taken into account: undesired consequences which, not being planned, do not comply with the institutional aims of evaluation and inertia, which hinders and reduces the effects of evaluation. Studies on public management in fact widely acknowledge that the adoption of evaluation systems has to take into account the organisation, the social relations that develop within it and individual behaviour and resistance (Taut and Browns, 2003 - Propper and Wilson, 2003 - Argyris, 1985 - Taut and Brauns, 2003 - Skinner, 2004- Royal Statistical Society, 2005). Power (1997), adopting some positions of the neoinstitutionalist school (Powell and Di Maggio, 1991), notes how audit activities are transformed into a codified ritual for producing legitimacy and inflexibly require that techniques are to be adopted. However they end up by creating phenomena of disconnection (the division between auditing routines and the true auditing activities) or colonisation (the substitution of institutional objectives with evaluation criteria).

So the study of the evaluation impact is based on:

- Organisational learning which includes the capacity to further knowledge, renew aims and adopt innovative behaviour and schemes for action (Argyris and Schön, 1978). University organisation is regulated by schemes i.e. a set of unwritten rules, ways of looking at things, behaviour codes, procedures and routines that deeply influence people's behaviour, to the point of guiding or adjusting it.

---

[*] The authors make a clear distinction between the "use" and "impact" of assessment. "Use" refers to the real tangible use of evaluation output inside or outside the university whereas "impact" refers to the consequences of the whole evaluation exercise, not just its use, on the university organisation.

- Development of resources, which regards the extent to which evaluation affects the prestige of the university and the technical, financial and human issues necessary for management.

- Power management, which are systems for the control and governance of power in the university's institutional and organisational functioning.

- Undesired consequences linked to the growing awareness that evaluation does not only bring about effects that are in line with institutional objectives but can also create undesired effects.

- Inertia, which refers to those elements that slow down, divert and hinder the creation and diffusion of the evaluation impact thus preventing change (Argyris, 1985; Taut and Brauns, 2003; Skinner, 2004).

Fig.1 – **Theoretical framework**
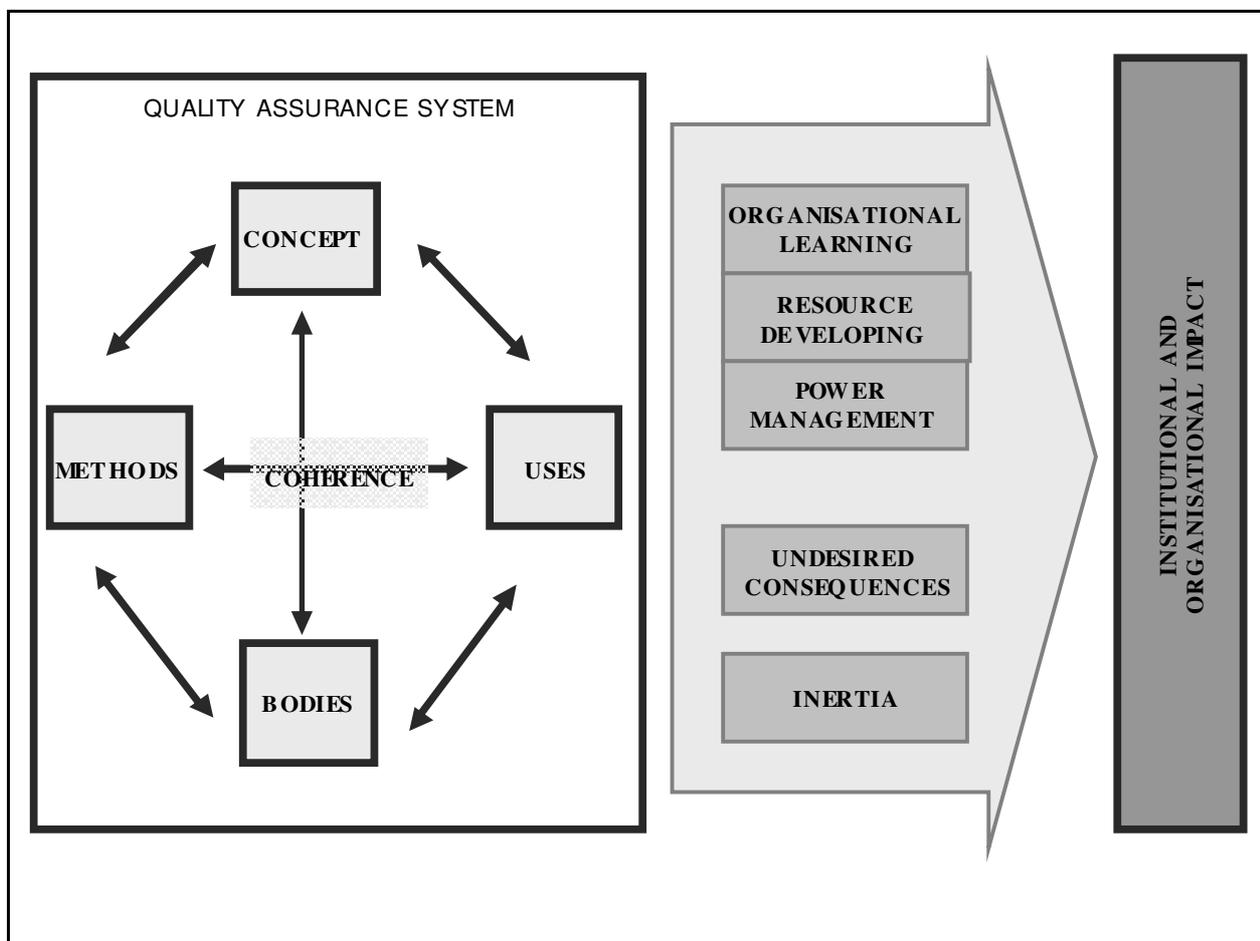
3. Methodology

The impact of evaluation in three European universities has been analysed in order to make proposals that will be confirmed by future studies.

The first case study is a medium-sized state university in the south-west of England – the *University of Bath*, founded in 1966 which has 12.000 students and about 3.000 employees of whom 600 are academics. There are 4 faculties and 2 schools with 17 departments in all. The annual budget is about 159 million euros.

The second case study is the prestigious old Dutch University – the *University of Amsterdam*, which has 7 faculties including the faculty of Medicine with its University Hospital. There are 22.000 students, 2.080 academics and 1.800 administrative employees. The annual budget is 448 million euros.

The third case is the *University of Trento*, a medium-sized state university in northern Italy founded in 1966. The University has 460 professors and researchers, 470 administrative employees and 14.500 students. There are 6 faculties and 12 research departments. The annual budget is 96 million euros.

The Netherlands and United Kingdom are the first European countries to have a well–developed evaluation system in Higher Education. Evaluation practices are long term and well–established in the institutional context of the national system of Higher Education evaluation which laid down its rules and established evaluation committees over 15 years ago. The national evaluation system in Italy is new and not well–consolidated. However, between 1998 and 2004 Trento University not only met national criteria but developed a set of autonomous evaluation practices showing that it is certainly avant-garde in this field.

Thus a comparison is made between two firmly–established situations and one in a university that is seriously tackling the initial stages of the evaluation process. Data were gathered directly in the universities at various intervals between November 2003 and July 2005.

The study aims specifically at understanding the practice of evaluation and its consequences over the past five years. Primary sources of information included semi–structured interviews with people in key positions in the University. These comprised vice-chancellors, pro vice-chancellors, deans, heads of research departments, academics, members of the evaluation board and general and administrative managers. Secondary sources included documents from internal archives, annual reports, other internal reports, research reports and minutes of evaluation board meetings. In all, 18 interviews were conducted in Bath ,18 in Amsterdam and 26 in Trento.

The study of the evaluation system in the three universities was based on the analytical scheme described in section 2.

This paper aims to pay particular attention to the study of the impact of evaluation. The empirical data gathered in the universities highlighted different effects in the three categories affected by evaluation impact:

- Organisational learning. Effects on: strategies and policies, culture, change in professional practices, increase in knowledge of university practices, attention to evaluation
- Development of resources. Effects on: technical infrastructure, reputation, university funds, allocation of economic funds within the university, human resources management
- Power systems. Effects on: institutional relationships, organisational structure, management operating systems, the evaluation system and bodies in charge of evaluation, variation in decision-making capabilities and leadership

In addition there were certain undesired consequences and inertia factors:

- Undesired consequences: Increased paperwork, overriding of institutional objectives, standardisation and flattening of processes, behaviour aimed at maximising evaluation results
- Inertia: difficult use of evaluation by the organisational structure, defensive attitude among academics

In the three universities the impact of evaluation in teaching and research has been considered separately underlining the different effects of evaluation and undesired consequences and inertia factors when they were present. A synoptic examination of the intensity of impact has permitted the grading of the impact, undesired consequences and inertia according to ordinary or strong intensity. It should be pointed out that this classification is strictly linked to the three cases in question.

4. Three European Universities: analysis of evaluation systems

Bath University

The concept of evaluation makes a clear–cut distinction between the structure of evaluation systems for teaching and for research. The former is the sum of the national and quality assurance system with an articulate view that is well–disposed to organisational needs and without sanctions. The latter system takes into account the formulation of the national Research Assessment Exercise (RAE) with a rigid system focused on international excellence, oriented towards comparison with other universities and closely linked to financial resources.

There is a marked difference between teaching and research methods. In teaching the most widespread methods are procedural (i.e. the study of the respect of procedures and prearranged requisites) and occasionally peer review. Different evaluation exercises such as national audit, internal quality assurance, self-evaluation and accreditation by external bodies are used. In research, the methods are defined within the framework of the RAE and centre on the review of scientific publications.

The bodies in charge of teaching evaluation, side by side with external bodies appointed under national evaluation procedures, show a marked level of institutionalisation that directly involves the academic structure. In research, however, the exercise is organised and supervised by external disciplinary panels who are chosen at national level on the basis of their own particular prestige.

Bath University makes great use of teaching evaluation, which is however focused on certain aspects and all in all not homogeneous. On the contrary, research evaluation, which is closely linked to the allocation of resources, is highly effective and centred on accountability.

University of Amsterdam

In Amsterdam, as in Bath, there is a dichotomy between the evaluation systems in teaching and research. The teaching system combines phases of evaluation within the university, which are aimed at improvement and open to organisational needs, with external phases that provide a synthesis and permit comparisons between universities. This harmonious synthesis makes it possible to deal with different needs. In the field of research, evaluation takes into account the formulation of the national system and is focused on excellence and oriented towards benchmarks. This is a rigid concept which, unlike the practice in England, has no direct link with the funding system.

Evaluation of teaching involves different methods. The various phases (self-evaluation, self-assessment reports, assessment inspections and meta-evaluation by the Ministry) permit the gathering and processing of a great deal of information. Research evaluation is based on a well-developed peer review method which summarises different kinds of information. The output is concise and guarantees a noticeable capacity of synthesis.

In the Dutch university two different bodies are responsible for teaching evaluation: the review board and the Ministerial Inspectorate and are both external, influential and independent. The bodies for the evaluation of research are national, highly authoritative and independent.

The uses of evaluation too are different. In teaching the results are used in a well–balanced way for many different purposes whereas research evaluation results are mainly directed towards accountability and offering guarantees.

Trento University

In Trento, unlike in the other two universities, evaluation in teaching and research is not distinct. Although it is affected by the national system it is, on the whole, an internal system. The concept is extensive and articulated but the approach is ambiguous and there are signs of disorderly, even if rampant evolution.

There are a large number of evaluation methods and particular attention is paid to analyses with quantitative indicators and indexes. A peer review exercise that makes direct reference to RAE criteria is also worthy of note. Generally speaking, unlike in the other two universities, the assessment exercise is not well–integrated as the different assessment techniques do not have a common logic.

There are two evaluation bodies: the University Evaluation Nucleus and the Commission for Scientific Research whose duties are ambiguously sub–divided. Both are internal bodies or in any case their members are appointed by the university. As the bodies are not well – established they are directly influenced by the choices of their chiefs.

Finally, the uses of evaluation are different in teaching and research as in the other universities. Whereas in Bath and Amsterdam they are the results of two totally separate evaluation systems, in Trento the uses are different despite the fact that there is only one system. In teaching the uses are purely cognitive and the results of the many evaluation methods are purely informational. Research evaluation there, on the other hand, is put to many uses by different people involved in university governance and affects the allocation of resources.

Degree of coherence

As mentioned above, the universities of Bath and Amsterdam have two evaluation subsystems, one for teaching and one for research whereas Trento has a single system with distinct uses.

The main differences in the systems not only concern their structure but also the different degree of coherence of each one.

In Bath, teaching evaluation has some trouble in putting official ideas into practice and this leads to difficulties that affect its uses. The system is extensive but not very coherent and too often focuses on individual aspects without summarising them. However, the research assessment system is overall extremely coherent, greatly influenced by the national assessment exercise, based on conceptual choices and strict, well–oriented methods that result in widespread, totally aligned uses.

The well-consolidated situation in the University of Amsterdam is based on two independent extremely coherent subsystems: these layered evaluation systems are the fruit of many changes and improvements and the synthesis of different needs and requirements.

Lastly, Trento is actively trying out evaluation systems but experimental factors are obvious and are more often the result of personal initiatives than a consolidated recognised system. Thus, the system is articulated but contradictory. It is widely used in some areas and being insignificant in others.

Fig. 2 – Principal characteristics of evaluation in the three universities

	Bath University		University of Amsterdam		Trento University	
	Teaching	Research	Teaching	Research	Teaching	Research
Concept	Articulated	External, rigid and focused	Sound and balanced	Strict and concise	Extensive and articulated but sometimes ambiguous	
Methods	Articulated and often procedural	Peer review (RAE)	Articulated, with peer review	Peer review	Fair number of evaluation methods	
Bodies	More bodies and strong institutionalisation	External panels	Two, external	One, external	Two bodies but not clearly divided	
Uses	Concentrated and not homogeneous	Outstanding and linked to distribution of funds	Considerable and well – balanced way	Considerable, directed towards accountability	Only cognitive	Considerable
Coherence	Limited	Strong	Strong	Strong	Weak, result of personal initiatives	

5. Three European Universities: the impact of evaluation

The institutional and organisational impact of evaluation is now examined.

The impact of evaluation on teaching

The impact of evaluation on teaching is different in each university as shown in Fig. 3.

Fig. 3 – Impact of evaluation on teaching

IMPACT OF EVALUATION - TEACHING				
CATEGORIES	**ELEMENTS**	**BATH**	**AMSTERDAM**	**TRENTO**
ORGANISA-TIONAL LEARNING	Effect on strategies and policies		XX	
	Effect on culture		XX	XX
	Change in professional practices	X	XX	
	Increase in knowledge of university practices	XX	XX	XX
	More attention to evaluation	XX	XX	X
RESOURCE DEVELOPMENT	Effect on technical infrastructure			
	Reputation	X	XX	X
	Effects on university funds		X	
	Influence on allocation of economic funds within university			
	Effects on human resources management		X	
POWER SYSTEM	Effects on institutional relationships		X	
	Effects on organisational structure		XX	
	Effects on management operating system	XX	X	
	Effect on evaluation system and on bodies in charge of evaluation	XX	XX	XX
	Variation in decision-making capabilities and leadership		XX	
UNDESIRED CONSEQUENCES	Increased paperwork	XX	X	X
	Overriding institutional objectives			
	Standardisation and flattening of operating processes	XX		
	Behaviour aimed at maximising evaluation results	X	X	
FACTORS OF INERTIA	Difficult use by organisational structure			XX
	Defensive attitude among academics			X
X = ordinary level of intensity – XX = strong level of intensity				

In *Bath* it is limited and centres on strengthening of the evaluation structure, the attention paid to it and the diffusion of information on university practices and activities. The evaluation system is extremely well-established and as a result of the widespread diffusion of quality assurance influences the teaching process and even occasionally modifies professional practice among professors. The impact is however confined to the management operating sphere since organisation, strategies, institutional and organisational relations are not affected by the assessment exercise.

On referring to teaching assessment a key figure in Bath university's governing body stated:"*it is certainly not one of our main worries even because we have no problems*". Basically, there is a situation of overall high quality teaching programmes that have no bearing on the funding system.

The general effect seems to be that of strengthening the evaluation system as a combination of instrumental procedures regulated by the pro-vice chancellor for teaching activities and the university's quality assurance committee.

One of the people interviewed stated: "*it is the quality system that is affected, not what is actually taught. There are no effects on teaching contents. Quality assurance regards its packaging*". Attention is therefore focused on evaluation, the widespread diffusion of information, standardisation of the structure of programmes and guaranteeing that standard requisites are met.

A worrying factor is the amount of administrative work and bureaucratic requirements involved as well as increasing standardisation of operating processes and a levelling of routine procedures. There is also a slight tendency to improve student satisfaction in the sense that professors are more congenial and less strict. The deep-rooted institutionalisation of the evaluation system in organisation at Bath obviously means that there is no noticeable inertia.

In *Amsterdam* the impact of evaluation on teaching is more strongly felt than in the other two universities. In organisational learning it has an enormous effect on the university's strategies and policies, culture and professional practice together with increased attention to evaluation and the diffusion of information on teaching. The impact of evaluation on the development of resources centres on the reputation of the university though of course negative evaluation could adversely affect state–funding. There is practically no effect on human resources management. The power system is greatly affected by evaluation, particularly at an organisational level, in decision–making skills and reinforcement of evaluation bodies. Slight effects are apparent in institutional relations and management operating systems. According to the interviewed professors, the assessment of teaching quality has led to the increased importance of teaching compared to research during the last few years. One of them affirmed: "*now everyone realises that didactics isn't the university's Cinderella*". There are also moderate undesired effects such as an increase in administrative duties and the spread of opportunistic behaviour among professors through maximisation of satisfaction among students and excessive reduction of the study load. There were no problems of resistance to evaluation practice nor inertia factors.

The evaluation impact in the *University of Trento* is in organisational culture (which is beginning to accept the system and come to grips with it), in increased knowledge of university practices and in the greater attention paid to evaluation practices and reinforcement of the bodies responsible for them. There were no effects in the area of the development of resources, but a slight effect concerning reputation. Undesired consequences were also insignificant and only involved a slight increase in administrative work whereas inertia factors were marked, which hindered the diffusion of the effects of evaluation in teaching. The organisational structure (the faculties) had great

difficulty in making use of evaluation output. This is a serious problem that limits the effects of evaluation to a cultural and cognitive level. The causes are to be found in the structure of the evaluation system used at Trento which only considers cognitive aspects and methodology that has given priority to information using sophisticated evaluation instruments to the detriment of synthesis and the usefulness of evaluative outputs. The deans who were interviewed declared that they do not have effective managerial instruments for influencing the teaching practice. "We can only speak to teachers whose rating is low and hope they will improve their performance in the future", says one of them.

In conclusion, a synoptic analysis of the three universities shows that the greatest impact of evaluation is in the organisational learning area and the lowest is in the area of development of resources.

Evaluation output is not linked to internal procedures for allocating economic resources or national funding in any of the universities.

The impact of teaching evaluation is greatest in Amsterdam University where evaluation (the result of a balance between internal and external processes) has far-reaching effects on organisational learning, the power system and reputation. The impact is not so strong in the universities of Bath and Trento for totally opposite reasons.

- In Bath evaluation is a long standing tradition and has undergone many changes over the years. It is extremely well-established and well-developed and there are no inertia factors. However the impact is not linked to the contents of teaching but to its organisation and administration bringing about standardisation and a flattening of the mechanisms for the management of teaching.

- In Trento the evaluation of teaching is in its early days and is thus hindered by inertia factors. The limited impact relates to the consolidation of evaluation, diffusion of knowledge and cultural aspects.

Impact of evaluation on research.

The impact of evaluation on research varies in the three universities.

At *Bath University* the well–established system of research assessment is the key to interpreting the impact of evaluation. University strategies are greatly influenced by the RAE whose criteria interweave with those of the university. The impact on professional practices is also strong whereas the effects on university culture and the diffusion of information on research in the university is less so since evaluation of research is now fully assimilated. The significance of the evaluation impact

on research in Bath University can be inferred from the importance given to the development of resources which are affected by research evaluation. The availability of resources from national funds is dependent on the results of the RAE just as the internal allocation of financial, human and technical resources also depends on the RAE rating. There are considerable effects on careers, remuneration of staff and university reputation. The important effects in the power system influence institutional relations, organisational structure, management operating systems and an increase in decision–making skills and leadership on the part of the heads of the university and departments. One interviewee clearly stated, "*The university has a strategy of portfolio for research according to the R.A.E. There are privileged departments, the 5 or 5*, others have to improve quickly or are eliminated*" and a Head of Department states "*For us the only impact is survival*", arguing that, in the case of a lower score in the next evaluation exercise (R.A.E. 2008), the department would probably be closed down. There are no factors of inertia but there are some undesired effects ranging from the increase of behaviour aimed at making the evaluation results better to the phenomenon of "colonisation" (Power, 1997) i.e. the subordination of institutional objectives to evaluative criteria. Finally, there is a tendency to level off and standardise research which risks conforming to the particular area that is most congenial for satisfying international excellence criteria. One interviewee made the following comment on the undesired consequences of the RAE"*we give priority to publications, we do less innovative and more sure research. We need international publications in certain journals and go all out to achieve that goal*";

Fig. 4 – Impact of evaluation on research

IMPACT OF EVALUATION - RESEARCH				
CATEGORIES	ELEMENTS	BATH	AMSTERDAM	TRENTO
ORGANISA-TIONAL LEARNING	Effect on strategies and policies	XX	X	X
	Effect on culture	X		XX
	Change in professional practices	XX	XX	XX
	Increase in knowledge of university practices	X		XX
	More attention to evaluation	XX		XX
RESOURCES DEVELOPMENT	Effect on technical infrastructure	X		
	Reputation	XX	X	
	Effects on university funds	XX		
	Influence on allocation of economic funds within university	XX	XX	X
	Effects on human resources management	XX	X	X
POWER SYSTEM	Effects on institutional relationships	XX		
	Effects on organisational structure	XX	XX	XX
	Effects on management operating system	XX		
	Effect on evaluation system and on bodies in charge of evaluation			XX
	Variation in decision-making capabilities and leadership		XX	XX
UNDESIRED CONSEQUENCES	Increased paperwork			
	Overriding institutional objectives	XX	X	
	Standardisation and flattening of operating processes	X		
	Behaviour aimed at maximising evaluation results	XX	X	X
FACTORS OF INERTIA	Difficult use by organisational structure			
	Defensive attitude among academics			
X = ordinary level of intensity – XX = strong level of intensity				

The impact of evaluation on strategies and above all professional practices in the *University of Amsterdam* is notable. In the words of the President of the University: "*Not words but ratings. There are thousands of universities in the world. There's no room here for research activities with ratings of 2 or 3. We want 4 and 5*". Evaluation and the diffusion of information on research ensure that no other effects are assimilated into the area of organisational learning. There are no links between the rating of evaluation and the system for financing university research whereas the internal mechanism for allocating funds to departments is greatly affected by evaluation. One of the people interviewed stated "*If I have to make any cuts, I don't cut each research program by 10% but only the one that isn't good and I don't harm the others. Yes, if funds are reduced, it's very useful for the Head of Department to have an objective evaluation. He doesn't need to justify his own personal choices but bases his decisions on evaluation*". The impact on reputation and the

management of human resources is slight. The impact on the power system that concerns organisational structure and the reinforcement of decision–making skills and leadership in the departments is strong. Undesired consequences include behaviour that aims at maximising evaluative criteria. There are however no inertia factors.

In the *University of Trento* the impact of evaluation on organisational learning is marked and widespread. Evaluation affects organisational culture, professional practices and leads to wider diffusion of information on research and greater attention to the evaluation system. The effect on the development of resources is less obvious, however, and only concerns to a small extent the internal mechanisms for the distribution of funds. The impact on the power system is centred on the organisational structure, where, amongst other things, the bodies responsible for the evaluation process have become stronger. The evaluation results have made a great contribution to reinforcing and consolidating leadership and decision-making mechanisms as evaluation makes it possible to face up to existing problems and overcome otherwise insuperable hurdles. As one Head of Department says: "*It's a bit like saying that you know that your son or daughter has problems at school. It's one thing to know it but its another to be told by a teacher*". There is very little undesired behaviour in Trento and it is confined to maximising the results of evaluation among researchers. No factors of inertia exist.

An overall comparison of the evaluation impact in research at the three universities shows that there are no factors of inertia and confirms that examination of scientific production using peer review (common to all three universities) does not meet with strong resistance. The three universities have no other factors in common. At Trento the internal evaluation system is innovative and mainly affects cultural aspects, organisational learning and a strengthening of the system. In the universities of Bath and Amsterdam, however, the evaluation systems are well-established and the effects show the consequences of a certain tradition and habit towards evaluation practices: the result is that the cultural impact in both universities is negligible. However there are great differences in the evaluation impact in research in Bath and Amsterdam universities.

- In Bath University the link between the results of the Research Assessment Exercise and financing for the departments means that the effects of assessment are of fundamental importance in the area of resource development. The importance of the assessment exercise and its outcome is such that its evaluative criteria may override the institutional objectives of the university.

- In the University of Amsterdam, on the other hand, where there are no links between the results of evaluation and the national financing of research, the impact tends to be less

strong particularly in the area of resource development. Undesired consequences are limited to colonisation and satisfying evaluation criteria as much as possible.

6. Reflections on determinants for the impact of evaluation

A comparison between the evaluation systems used in the three universities shows that:

- evaluation has a great impact on the institutional and organisational running of the universities. Decisions on evaluation cannot therefore be limited to technical and specialist regulations but must also take into account their importance in terms of strategy;

- the scheme used for recording the impact of evaluation appears to be a useful instruments for gathering, classifying and comparing the effects of evaluation.

- there is a relationship between the characteristics of the evaluation systems and evaluation impact. The structure of the evaluation system, the concept of evaluation, the methodology used, the characteristics of the evaluation bodies and making use of the reports influence the organisational and institutional impact of evaluation in the universities;

- the specific characteristics of the universities where assessment is carried out also affect the institutional and organisational impact of evaluation. Thus, no deterministic relationship exists between the structure and impact of the evaluation system unless the organisational and institutional characteristics of each university are taken into account.

Overall, the impact of evaluation is not automatic but is only comprehensible as a function of the characteristics of the evaluation systems and the universities that use them.

Data collected in the three universities show how players in the evaluation system tend to emphasise the importance of the instruments, procedures and stages of evaluation in which they have been directly involved or are more in agreement with, even comparing them with other evaluation exercises in which they have not been involved. The latter fact is confirmed by Newton (2000) who, when referring to the situation in England, states that academics do not passively accept management systems and quality assurance but, on the contrary, are involved in "interpretative work" to decide between acceptance or opposition to their own advantage. The link between evaluation, internal quality culture and behaviour of organisational actors is fundamental for determining evaluation impact.

The degree of involvement and approval of organisational actors determines whether evaluation will have an impact or not. The more the academics are involved in the evaluation process, follow its development, appreciate its contribution and share in it, the greater the impact will be. Generally speaking, the importance given by actors to the awareness, use and promotion of evaluation means

that by actively taking part in the assessment exercise they can make it dynamic and thus greatly increase the likelihood of overcoming resistance and factors of inertia. Literature recognises the importance of this concept which is often known as "*evaluation property*" (Hodson and Thomas, 2003 – Rebora, 2003). The situation in the three universities shows how sharing and taking part in the evaluative process is essential for the success of evaluation.

These observations are not complete unless the cultural background of the organisational actors is taken into account since the organisational impact of evaluation systems is strongly influenced by this factor. Culture can be an obstacle when a priori positions or deeply-rooted professional values interfere with the use of evaluation results. On the other hand, when an assimilated evaluative proposal is shared by organisational environments there are many evaluation effects. In practice the importance of cultural aspects in the evaluation of academic activities is crucial for determining factors of inertia and resistance.

Examination of the effects of evaluation shows that the more evaluation is assimilated by the organisation and cultural needs are taken into account the fewer elements of inertia and effects in organisational learning there are. Conversely there is greater impact on the power system and resource development.

If evaluation is alien to organisational culture then its diffusion will be extremely slow and its impact will have to be considered in relation to organisational learning rather than to the development of resources and the power system. On the contrary, if practice and evaluative criteria are closely linked to university culture there will be no factors of inertia or organisational learning impact but rather widespread use and impact of evaluation on resource development and in the power system.

Taking an empirical approach with an exploratory aim it is possible to pinpoint certain recurrent factors that are fundamental for understanding the effects of evaluation in the universities. A strategic analysis of evaluation processes in the light of quality culture emphasises the following points:

- importance of peer review
- difference between teaching and research
- importance of time factors
- risk of degeneration of evaluation.

Importance of peer review

The acceptance of evaluation by organisers is an important pre-condition but it is fundamental that the methods used either reflect or respect the culture and values of the people involved in it. This explains the importance of peer review methods in the three universities that allocate the task of evaluation to experts chosen on the basis of their own authority in the particular field that is being assessed. In many cases the community singles them out or in any case expresses a preference for them. The evaluative output is thus consistent with academic values, is consequently faced with fewer obstacles and produces a positive impact more easily. Other types and methods of evaluation are much more at risk of being rejected by academics as they are considered alien to the principles upon which the organisation is founded. Self-referentiality is a well-known practice in universities and evaluation risks coming up against this with very little possibility of winning. However, if evaluation is accepted by the organisation, then by working its way into the structure with the loosely coupled systems that are typical of all universities it becomes a reference point for decision-making processes. In fact when assessment reports are available they slot into the gaps in the mechanism of professional bureaucracy helping to create a large of important organisational effects that the organisational actors can influence and handle.

Thus:

- evaluation data stimulate government bodies to make decisions and give them elements for justifying or motivating their choices;
- in a professional context (Mintzberg, 1983) characterised by great discretionary power and lack of regulations, the assessment exercise lays down fixed points even under a time profile (expiry dates, periods, deadlines) that the organisation uses as a reference point and stimulus;
- evaluation can help to defeat the fence-sitting attitude of academic staff (no participation in characteristic processes except through generational substitution) by introducing stimuli which come into line with professional behaviour and practice.

Differences between teaching and research

The empirical analyses carried out have studied the impact of evaluation on teaching and research separately. In term of evaluation, teaching and research are two distinct issues. It is not a coincidence that in literature they are rarely discussed in the same paper but are considered individually. Some authors even go so far as to say that evaluation of research is less complicated than teaching (Tavenas, 2004). Research evaluation is a normal part of much of the researcher's work. Especially in the more advanced university systems the acquisition of research funds, publication of results or promotion all depend on assessment of the researcher's work. From a

cultural point of view evaluation is therefore not such a new concept in research as in teaching where it immediately calls up the concept of lack of trust, accountability and reform of university professors' traditional self-referentiality. Evaluation of teaching takes on threatening connotations reinforced by the professors' difficulty in foreseeing and assessing its outcome. There is in fact no method that immediately indicates the value of the teaching performance. Once again, research is at a clear advantage since published papers are a good indication of the scientific activities that they refer to. Roughly speaking, one can say that understanding the value of a researcher's paper means understanding the scientific value of his or her work. Evaluation does not surprise a researcher who knows how to assess his own papers and above all knows what evaluation method his own scientific community uses.

Thus, the habitual use and ease in recording results in research are what make evaluation and its impact different from teaching. There is however another difference between the two areas: in teaching the diffusion of information is irregular and variable whereas in research it is more widespread and easier to interpret. This does not depend on the setting in which they take place but on the intrinsic nature of the two activities:

- Teaching and in particular the relationship between professors and students, is in itself transitory, indefinite and difficult to assess. For various reasons university students express their personal choices on the basis of totally different factors from those that really indicate the quality of the options available. Often some data may be missing or information is scattered and difficult to summarise. There is a great lack of homogeneity between the information in the hands of the students and that of their professors and the university in general. This lack is felt before going to university but also throughout all the time spent studying there. Students, the world at large and even many people involved in teaching have difficulty in understanding the true quality of teaching. On the contrary, there is a strong tendency to judge it in terms that have very little to do with it (social, political, cultural trends etc.).

- Research is a field where the people involved are well-informed and have the necessary instruments to appreciate the quality of research. In any university an academic recognises the value of his colleagues' research work, the institute concerned and the main exponents of the subject.

In teaching and research the discrepancy in the diffusion of information explains the different role and, above all, the different impact of assessment in the universities. In the teaching field evaluation attempts, more or less knowingly, to reduce the discrepancy in information between the participants and the teaching process. This involves introducing assessment exercises which in some way satisfy

the growing needs of accountability which have been prompted by society's eve increasing lack of trust in universities (Hoecht, 2004).

Research evaluation is already a tradition so setting in motion assessment exercises reinforces and strengthens processes that already exist.

To conclude, teaching and research evaluation show marked differences in terms of organisational impact. The most appropriate solution seems to be the one adopted by the Universities of Bath and Amsterdam which have independent evaluation systems. It is however worth inquiring whether this is the only solution, what advantages it offers and whether it is opposed to the widespread, deeply-rooted academic idea of unity between research and teaching. The question is: if one hopes that in our new society with its European outlook the old concept of unity between research and teaching in Higher Education will not completely disappear, is it then sensible to use an evaluation system based on a clear distinction between these two sectors (Rebora, 2003; Felt, 2004)?

Importance of time factors

In the empirical data recorded on the evaluation process there is a strong sharing of the incremental logic, according to which until the evaluation exercise reaches a certain consistency permitting the consolidation of evaluation procedures, evaluation only has a slight cognitive effect that is not connected to specific action by the university. Only the continuation of assessment and sustained use will produce a significant impact on the university

The variable time factor has often been decisive in accounting for certain effects existing in one case and not in an other. Time is important in evaluation and in particular affects inertia and undesired consequences. Inertia factors are more strongly felt in the early stages of evaluation and decrease with the consolidation of evaluation. On the other hand as evaluation becomes stronger the undesiderable consequences that were non-existent in the early stage intensify and take root in the organisation.

Overall, it is possible to agree with and extend the conclusions of Jeliazkova and Westerheijden (2002) who hold that evaluation systems evolve over time. As studies show, time is important for understanding not only the dynamics of evaluation but also its impact.

Risk of degeneration of evaluation.

In literature various authors have contested the capacity of evaluation to have positive results on university problems and attribute it with the phenomena of standardisation and dequalification (Hayes and Winyard, 2002).

Many papers have discussed the difficulties and unacceptable effects of evaluation in the university setting (among others: Oliver, 1991; Parker, Jary, 1995; Rhoades, Sporn, 2002; Talib, Steel, 2000; Power, 2002; Power, 2003; Czarniawska, Gene II, 2002; Morley, 2003; Brown, 2004). Empirical data resulting from this study show that there are many positive desirable effects as well as with some inappropriate or undesirable factors resulting from the evaluation exercise.

The impasse that exists between supporters and denigrators of evaluation is at the centre of a study by Hoecht (2004) who, before illustrating the adverse effects of the introduction of evaluation in Higher Education in the U.K. affirms: "*the former writers argue that quality management colonises our academic life-world but they tend to ignore or deny their own self-interest in maintaining a traditional university system that lacks accountability and transparency*".

Our empirical analysis further supports Hoecht's argument according to which evaluation in itself is neutral and makes it possible to achieve important and knowingly oriented organisational impact as well as spark off processes of organisational change in the sense of improvement. Degenerative phenomena are risks that can be dealt with as long as opportune countermeasures are activated and the exercise is conducted with the necessary awareness and attention on the part of government and management bodies.

References

Argyris, C. (1985) *Strategy, Change and Defensives Routines*. Boston: Pitman
Argyris, C., Schon, D. (1978) "Organizational Learning", Reading, MA: Addison-Wesley
Beer, M., Nohria, N. (2000) "Cracking the code of change", Harvard Business Review (May-June): 133-141
Boer, H., de (2002), *Trust, the essence of governance?* In Amaral, A. et al. (2002), *Governing higher Education: national Perspectives on Institutional Governance*. Dordrecht: Kluwer Academic Publishers
Boer, H., de and Denters, B. (1999), *Analysis of institutions of university governance: a classification scheme applied to postwar changes in Dutch higher education* in Jongbloed, B. et al. (1999), *From the Eye of storm; higher education's changing Institutions*. Dordrecht: Kluwer Academic Publishers
Boer, H., de and Huisman, J. (1999), *The new Public Management in Dutch Universities* in Braun, D. and Merrien, F. (1999), *Towards a New Model of Governance Universities? A comparative View*. London: Jessica Kingsley
Brennan, J. and Shah, T. (2000a), *Managing Quality in Higher education. An International perspective on institutional assessment and Change*. Buckingham: Open University Press
Brennan, J. and Shah, T. (2000b), *Quality assessment and institutional change: experiences from 14 countries*. Higher Education, 40
Brown, R. (2000), *The new UK Quality framework*. Higher Education Quarterly, Vol. *54, n.4*
Brown, R. (2004), *Quality assurance in higher education. The UK experience since 1992*. London: RoutledgeFalmer
Czarniawska, B. and Genell, K. (2002), *Gone shopping? Universities on their way to the market*. Scandinavian Journal of Management, 18, pp455-474
Elton, L. (2000), *The UK Research Assessment Exercise: Unintended Consequences*. Higher Educational Quarterly, Vol. 54, N. 3
Felt, U. (2004), *University Autonomy in Europe: Shifting Paradigms in University Research*, in *University Autonomy: Shifting Paradigms in University Research*. Proceedings of the Seminar of the Magna Charta Observatory, 15 September 2003, Bologna: Bononia University Press, 2004
Hare, P. (2003), *The United Kingdom's Research Assessment Exercise: Impact on the institutions, Departments, Individuals*. Higher Education Management and Policy, Vol. 15, N. 2
Harvey, L., Newton, J. (2004), *Transforming Quality Evaluation*. Quality in Higher Education, Vol.10, N. 2
Hayes, D. and Winyard, R. (2002), *The mcdonaldization of higher education*. Westport: Bergin & Garvey
Henkel, M. (1998), *Evaluation in Higher Education: conceptual and epistemological foundations*. European Journal of Education, Vol. 33, N. 3
Henkel, M. and Little, B. (1999), *Changing relationship between higher education and the state* London: Jessica Kingley Publisher
Hodson, P. and Thomas, H. (2003), *Quality assurance in higher education: fit for the new millennium or simply year 2000 compliant?* Higher Education, Vol. 45, pp 375–387
Hoecht, A. (2004), *Accountability, Trust and Professional Autonomy in UK Higher Education Quality Management*. In atti del Workshop on the process of reform of the university across Europe EIASM (Cd-Rom), Siena, May 24-26, 2004
Huisman, J. and Currie, J. (2004), *Accountability in higher education: bridge over troubled water?* Higher Education, Vol. 48, pp. 529-551
Jeliazkova, M. and Westerheijden, D. (2002), *Systemic adaptation to a changing environment: Towards a next generation of quality assurance models*. Higher education, 44, pp. 433–448
Minelli, E., Rebora, G. and Turri, M. (2002), *Il valore dell'università*. Milano: Guerini e Associati.
Minelli, E., Rebora, G. and Turri, M. (2005), *University and evaluation: a case study of the organizational impact*. Liuc Papers, n. 166, Serie Economia e Istituzioni 17, suppl. a marzo 2005 (Working paper)
Mintzberg, H. (1983) "Structures in fives: defining effective organizations", Englewood Cliffs, Prentice-Hall
Mintzberg, H. (1983), *Power In and Around Organizations*. Englewood Cliffs: Prentice-Hall
Morley, L. (2003), *Quality and Power in Higher Education*. Philadelphia: SRHE & Open University Press
Newby, P. (1999), *Culture and quality in higher education*. Higher education policy, 12
Newton, J. (2000), *Feeding the Beast or Improving Quality?: academics' perceptions of quality assurance and quality monitoring* Quality in Higher education policy, Vol. 6, N. 2
Oliver, C. (1991), Strategic responses to institutional processes. Academy of management review, Vol.16, N.1, pp. 145-179
Parker, M. and Jary, D. (1995), The *McUniversity: Organization, Management and Academic Subjectivity*. Organization, Vol.2., No.2, pp 319-338.
Pfeffer, J. (1981) "Power in organizations", Marshfield, Pitman Pub.
Pollit, C. (2003), *The essential public manager*. Maidenhead: Open U.P.
Pollit, C. (2004), *Performance information for democracy – the missing link?* EES Sixth Conference. Berlin, 30th September-2nd October 2004
Powell, W. W., & DiMaggio, P. J. (Eds.) (1991) "The new institutionalism in organizational analysis", Chicago, University

of Chicago Press
Power, M. (1997), The audit society. Oxford: Oxford University press
Power, M. (2003), *Auditing and the production of legitimacy.* Accounting, Organizations and Society, Vol. 28, pp. 379–394
Power, M. (2004), *Counting, control and calculation: reflections on measuring and management.* Human relations, Vol. 57, N.6, pp. 765-783
Propper, C. e Wilson, D. (2003) "The use and Usefulness of performance Measures in the Public Sector", CMPO Working Papers Series N. 03/073
Randall, J. (2002), *Quality Assurance: Meeting the Needs of the User.* Higher Education Quarterly, Vol. 56, N.2, pp. 188-203
Rebora, G. (1999), *La valutazione dei risultati delle amministrazioni pubbliche.* Milano: Guerini e Associati
Rebora, G. (2003), *La valutazione degli atenei: alla ricerca della qualità.* Azienda pubblica Anno XVI, N.5-6
Rhoades, G. and Sporn, B. (2002), *Quality assurance in Europe and the U.S.: professional and political economic framing of higher education policy.* Higher Education, Vol. 43, pp. 355- 390.
Royal Statistic Society, (2005) "Performance indicators: good, bad, and ugly", http://www.rss.org.uk/main.asp?page=1222 http://www.rss.org.uk/main.asp?page=1222
Schwarz, S. and Westerheijden, D. (2004) Accreditation and Evaluation in the European Higher Education Area. Dordrecht: Kluwer Academic Plubliscers
Sharp, S. and House, M. (1995), *The quality of teaching and learning in Higher Education: evaluating the evidence.* Higher Education Quarterly, Vol. 49, N. 4, pp. 301-315
Skinner, D. (2004) Primary and Secondary Bariers to the evaluation of Change. Evidence form two public sector organizations. *Evaluation*, Vol. 10, n.2., pp. 135-154
Stensaker, B. (1999*), External quality auditing in Sweden: are departments affected?* Higher Education Quarterly. Vol. 53, N. 4, pp. 353-368
Talib, D. and Steele, A. (2000), *The Research Assessment Exercise: strategies and trade-offs.* Higher Education Quarterly, Vol. 54, N. 1, pp. 68-87
Tapper, T. and Salter, B. *Interpreting the Process of change in Higher Education: the case of the Research Assessment Exercises.* Higher Education Quarterly, Vol. 57, N. 1, pp. 4-23
Taut, S. and Brauns, D. (2003) *Résistance to evaluation. A psychological perspective.* Evaluation. Vol.9. N. 3. pp. 247-264
Tavenas, F. (2004), *Quality assurance: a reference system for indicators and evaluation procedures.* Brussel: EUA:. (doc. disponibile www.eua.be)
Taylor, J. (2001), *The impact of Performance Indicators on the Work of university Academics: Evidence from Australian Universities.* Higher Education Quarterly, Vol 55, N. 1, pp. 42-61
Tichy, N. (1983) "Managing Strategic Change: Technical, Political and Cultural Dynamics", New York, John Wiley & Sons
Turri, M. (2003), *Lo stato dell'arte, limiti ed evoluzione nella valutazione a livello di sistema e di singolo ateneo.* Azienda pubblica, Anno XVI, N. 5-6
Turri, M. (2005), La valutazione dell'università. Un'analisi dell'impatto istituzionale e organizzativo. Milano: Guerini e Associati
Underwood, S. (2000), *Assessing the Quality of Quality assessment: the inspection of teaching and learning in British universities.* Journal of Education for Teaching, Vol. 26, N. 1
Weick, K.E., (1976), *Educational organisation as loosely coupled systems.* Administrative Science Quarterly, 21, pp. 1-19