Knowledge Assessment in the Agora

Frode Frederiksen, Finn Hansson
& Søren Wenneberg

WP 14/2001

December 2001

MPP Working Paper No. 14/2001 ©
December 2001
ISBN: 87-91181-01-1
ISSN: 1396-2817

Department of Management, Politics and Philosophy
Copenhagen Business School
Blaagaardsgade 23B
DK-2200 Copenhagen N
Denmark
Phone: +45 38 15 36 30
Fax: +45 38 15 36 35
E-mail: as.lpf@cbs.dk

Knowledge assessment in the Agora

Frode Frederiksen
Finn Hansson
Søren Barlebo Wenneberg[1]

all from

Department of Management, Politics and Philosophy
Copenhagen Business School
Blaagaardsgade 23 B
2200 Kbh. N.
Copenhagen
Denmark

Abstract:
The internal scientific evaluation of knowledge claims done by peer review is todayhas for the last 20 years been supplemented by new methods of knowledge assessment. These methods are usually taken from other sectors such as the public administration and the corporate sector. Because of their different backgrounds compared to the peer review method, it is necessary to critically examine them in relation to assessment of quality. But this also calls for a new understanding of science and knowledge. The article starts by briefly explaining this new understanding, before going into more details on the new methods of knowledge assessment and the perspectives behind them.

Keywords:
The agora, The Triple Helix, research evaluation, research quality, a new understanding of science, universities, industry.

Introduction

Science is under change. The classic understanding of the purpose of science as production of knowledge for its own sake is still very dominant, but a new understanding focusing on the application of science is under rapid development. Concepts such as "mode 2 knowledge production" (Gibbons et al. 1994), "The Triple Helix" (Etzkowitz and Leydesdorff 1997) and "the entrepreneurial university" (Etzkowitz 2000) are all pointing in the same direction, namely that science no longer only produces knowledge for its own sake, but also for the society. The new book "Re-thinking science" (Nowotny et al. 2001) also has the connection between science and society as its central focus. The most important concept in the book is "the Agora" and it accentuates precisely that the societal interest in science is so vigorous that the distance between science and society is diminishing rapidly. Instead, a discussion place – an Agora – is being established. This place is a melting pot of all kinds of discussions about knowledge.

[1] Paper presented at the 5th ESA-conference, Helsinki, August 28th-September 1st in the SSTNET session.
Contact by email: Frode Frederiksen, ff.lpf@cbs.dk - Finn Hansson, fh.lpf@cbs.dk - Søren Wenneberg, wenne@cbs.dk .

In fact the Agora is not a new physical location, but a common name for places, networks and institutions with influence on knowledge. That means that the Agora contains many of the well-known "institutions": journals, scientific conferences, universities, research councils, ministries, parliaments, media, public opinions, social movements etc. The important change is that one cannot anticipate where the essential discussions and final decisions take place. Hopefully, these problems will be an object for future studies of the Agora. It is important to emphasize the difference between the classic concept of the market and the concept of the Agora. The Agora is much broader and encompasses not only the traditional market for products with quantities and prices (which often is implicit when talking about market), but also the political market for negotiations, the scientific market for reputation. (Foss Hansen, 1988)

The orientation of science toward society also demands a social robustness of the scientifically produced knowledge in order to be used in the Agora (Nowotny 1999; Nowotny et. al. 2001). The demand for societal robustness is a rather new dimension in quality assurance in science. In this paper, we will look into how it changes and influences the traditional procedures and systems for evaluation of scientific knowledge. We will, especially, look at the changes in how scientific knowledge is being evaluated. How is this very important activity in science perceived in the market and discussion places where both knowledge producers and knowledge users meet? In the traditional Mertonian view on research evaluation, science is a closed and autonomous system that in a sovereign way decides on questions regarding the quality of the scientific work and results. The functional CUDOS norms and the peer review guarantee that the knowledge produced is trustworthy (Merton 1968; Ziman 2000)). However, it is exactly this conception of science that is under change. Science is in numerous ways involved in society and is influenced by its societal stakeholders - also regarding how the knowledge claims should be assessed. The aim here is to go into some of these new perceptions of research evaluation – its goals and methods. We will, in particular, examine some of the forms of evaluation that today are initiated from political side (e.g. from public authorities and agencies) and the forms imported from industrial research.

The paper is structured in the following way. Firstly, we examine the new views on science and knowledge in detail. Secondly, we look upon the new perspectives on research evaluation and their background. Thirdly, we study some of the new evaluation methods originating from industry. Lastly, it is argued that we should change our view on research and especially research management to use these new assessment methods in an effective way.

A new understanding of knowledge and science

The traditional understanding of science and scientific knowledge has at least three defining characteristics: 1) Knowledge is defined as "justified true belief". 2) The goal of science is to produce this kind of knowledge. 3) Scientific work is done by "lonely seekers of truth". This could be called "the legend of science" (Ziman 2000). It is precisely this view that is under change. Not only because the new emerging developments in society in general (e.g. the trend towards the knowledge society where production of new knowledge is paramount for the economy), but also as a consequence of developments inside science itself. Let us shortly go through the three characteristics and the changes they are exposed to one by one.

An idea of what truth is, is together with the idea of objectivity the fundament of the view on science coming from the philosophy of science tradition. The "correspondence theory of truth" is the most well-known idea of what truth is. This idea also influences the way justification is perceived. The scientific method shall ensure that your belief is a true representation of the field under study. Consequently, knowledge is true representations of reality produced through the right scientific methods.

The problem with this definition of knowledge is that philosophers have never been able to produce a convincing account of what *the* right method that secures true representations is. Exactly, this problem is the point of departure for the changed conception of knowledge. One of the fundamental problems is how you measure representations against reality itself. Because it is exactly our best and most advanced scientific beliefs that control how we perceive reality. We do not have "a bird eye view" on reality that can tell us how reality is mirrored in our scientific knowledge.

As a consequence of these problems, another understanding of knowledge has been proposed. Pragmatist philosophers (e.g. James, Dewey and Rorty) see knowledge as something that makes action possible. In a popular way, the idea expressed is something like: When you know what you can do, you know what you know. They opposed the possibility of making representations of reality. Instead, they propose an understanding where knowledge is established through its use. Knowledge is a way of world making.

The pragmatist (and constructivist) view on knowledge has also consequences for the goal of science. Science can no longer produce knowledge that is true. The problem is twofold. The first problem is that science cannot decide with certainty what truth is. The second problem is that knowledge is not produced in a vacuum, but only through concrete application. You have to use it before you know what you know. Therefore, knowledge application is not a boring residual activity done after knowledge is produced. It is produced and used in the same instant. As a consequence of this, science should not only strive for the production of true knowledge, but also organise itself as an institution in such a way that its stakeholders consider it as trustworthy as possible and therefore accept and use the knowledge. This is the new goal of science.

This way science is not a republic of lonely seekers of truth. Instead, science is perceived as a social institution based on co-operation and teamwork. New transdisciplinary problems and the need for huge experiments and apparatus necessitate this development. The traditional history of science as a succession of great discoveries done by geniuses must be rewritten. Instead, we should focus on science as an evolving institution based on co-operation. An important theoretical perspective offering itself here is to focus on the evolution of science as a development of trust producing institutions (Bordum and Wenneberg 2001) – for example the evolution of the witness function that makes scientists believe in each others' scientific experiments (Holton 1993).

These three changed perceptions of science are not the only ones. But they all point in the same direction – science is no longer seen as a special kind of divine cognitive activity done by 'geniuses'. Instead, science is "just" another social knowledge-producing institution, one among many in modern society. But it is *the* institution that – because of its historically origin – is most advanced in looking independently, critically and reflexively on its own institutional knowledge base and therefore until recently the most trustworthy one.

As a result of an internal development inside science this new understanding is now not only something that research politicians promote. Consequently, this new understanding also has to be used as a basis for the design and organisation of science as an institution. Just like the old model of science has been the blueprint for the organisation of science for at least the last 50 years. This also implies that science studies and research policy should begin to address the micro-level of science – instead of only promoting these changes on a macro level. And the micro level is about how to implement these new concepts and understandings through management, organisation and micro politics. The following discussions on research evaluation should be seen in this light.

New perspectives on knowledge assessment

Evaluations have always been an integrated part of the research system. When researchers apply for an academic position, for publishing a scientific article, for research grants, academic degrees, prices etc., an evaluation of their research performance is a natural element in those processes. The researchers themselves often claim that they are the profession group, which most frequently is evaluated. But the last 20 (or so) years, we have seen a development of new types of evaluations, both concerning scale (amount) and scope (purpose). As mentioned by OECD, research evaluation has emerged as an "rapid growth industry" (OECD 1997 p. 5). This development can be seen – among other factors – as a result of the new understanding of science and society (a shift from a mode-1 society to a mode-2 society). When research or knowledge production is developed in a "context of application" (Gibbons et al.1994) and in a "context of implication" (Nowotny et al. 2001), it is expected to contribute to various goals (scientifically and practically); and when knowledge and research are produced by different institutions, it is necessary to develop and use new kinds of evaluations. New ways of producing trustworthy information of the knowledge produced. To help us understand the changes on the surface, we will have a look at the changes in the function and the meaning of the evaluations. These changes in research evaluation have many sources – from general societal development and from experiences of evaluations in other areas. The important changes in perspectives can be observed in several dimensions. Let us look at some of them.

The first 'answer' to the problem of creating societal trustworthy or robust information on scientific quality under changing social conditions came in the early 1980ties. As a part of the general reforms of the public sector starting, often known as New Public Management, the 'value for money' or auditing of the public sponsored research became a part of this development. The major line in this development was the creation of a large number of quantitative systems or techniques to be used for counting almost everything. "Everything we touch turns into numbers" (Rescher 1995:82). This development has been captured by Michael Power (1997) in the concept of 'the audit society' describing how the visible traditional forms of control in organisations are replaced by new internal systems of control and monitoring. In this process, trust is based on information produced by the control systems and not by some inside information only accessible to the community of scientists. The result is a marked tendency to base trust in the use of quantitative information.

In the light of New Public Management, one of the purposes of evaluation is, as well as in nearly all other public areas, to control the expenditure or at least to create an impression that the expenditures is under control by the public managers. The latter is called symbolic evaluations. Evaluations have developed into a kind of ritual behaviour, which is not to discussion any longer. They do certainly

have effects, but not in a traditional rational manner of measuring official purposes or goals (Dahler Larsen 2000).

Another basic – but not official – purpose for evaluations is (just as the earlier discussions around utilisation of applied research in general (Weiss 1979)) to use them as a kind of political ammunition in a decision making context in order to legitimise decisions (Albæk 1996).

Another trend in a lot of public institutions – after the first wave of external evaluations – has been a greate emphasis on internal evaluations as a means in organisational processes and development. In that way, evaluations are used as a tool for learning in organisations (Foss Hansen and Borum 1999, 2000).

All these above-mentioned changes in the understanding of why we evaluate knowledge – from rational explanations to symbolic functions – are not the only changes in the understanding of research evaluation. In the following paragraphs, we will look into some of the other changes and developments.

Firstly, a great part of evaluations is *initiated* by bodies and institutions outside the universities. It is typically from different parts of the research policy system, covering a wide spectrum from parliaments to advisory boards and research councils. Several purposes of evaluations can exist side by side, and the design of evaluations is usually made outside the research system. For many evaluations, the shift in initiative is the only modification, but because such evaluations are initiated in a political landscape, they often result in remarkable vibrations in the various research groups.

Secondly, the *users of evaluations* are now not only researchers, but more and more non-researchers, either belonging to the research system or representing other parts of the society. The profile of the group of users depends on the different purposes of the evaluations, for instance if it is research planning in general or a specific grant allocation. Furthermore, while the purpose can change during the various phases of an evaluation, it is not always clear at the beginning of an evaluation whom the users will be. It is often a result of the political dimension of the evaluations.

Thirdly, the *evaluators* are not any longer only other researchers. The traditional characteristics of research evaluation is that it is made by peers (Hansson 2001), and the peers could almost freely be selected among the qualified peers by the science community. In this new development of research evaluations, the researchers can be evaluated by persons who are not peers in the traditional meaning (they must cover a very broad area or they are experts in other fields). Or they could be experts, but chosen by people outside academia, or they may be generalists and not scientific experts at all. Evaluators have emerged as a new profession.

And – fourthly – it is characteristic for the new development of evaluations that they have both seen a *change in focus and a large degree of variation of foci*. The most important ones are:

- Evaluations in the classical perspective have had the focus on *the substance in the scientific contributions* (validity, originality, novelty etc.). The new perspective in a mode-2 context has the focus on *practical relevance and usefulness* for the society. It can be a technical, environmental or commercial success on the market or in the public sphere in general.

- The focus is concerning *different levels of research systems*. In the mode-1 context, the predominant focus was on the work of individual researchers. It is still important, but is now supplemented by a row of other levels. Evaluations can comprehend larger research groups, laboratories and institutions such as universities, they can focus on an entire scientific discipline, government programmes and funding agencies or evaluation methodologies can be applied to a country's entire research base (OECD 1997 p.6).

- Another remarkable change is the *time-dimension* for evaluations. Do they have an ex ante or an ex post focus. Many evaluations have had a tendency to reward "past heroes", and this is not a guarantee for the best opportunities for future research. To assess future possibilities, the focus can be on research education, equipment, strategies, network, organising and management of research.

These changes are not a matter of a shift – once and for all – from one perspective to another, but the new foci will exist in different degrees in various evaluations. Modifications in just one of the dimensions are often – by the researchers – perceived as a rather comprehensive change.

To sum up: the society does not any longer trust a one-dimensional evaluation approach conducted in an isolated system with one-dimensional quality criteria. The development and the increase of the numbers of research evaluations can be seen as an attempt to create public trust in the research – it is the symbolic dimension as we mentioned above.

A major problem when quality control moves into the agora is that scientists have not been used to the feedback and influence from other parts of society.

> *"Caught on the defensive, they blame contextualization, and the agora in which it emerges, for the rise of anti-science sentiments, for the subversive influence of social scientists and other 'relativists'. They fear that irrationality will break through the fragile crust of scientification*." (Nowotny et al. 2001 p. 207)

Consequently, one of the main challenges of research evaluation is to re-establish trustworthy relations to research in general, but also in an enlargement of these trustworthy relations in order to cover/reach the various groups who participate in the agora: scientists, politicians, the media, social movements, non-governmental organisations, business people (and lay citizens to include all). The first step in restoring the trustworthy relations regarding knowledge assessment is to explicate the huge variety of processes that has knowledge assessment as an important part. This is the purpose of the last part of the paper.

New types of assessment methods

The following figure is an attempt to present a picture of the complicated and many-sidedness of what is understood as research evaluation or assessment today.

THE MANY DIMENSIONS OF RESEARCH EVALUATION: PROFESSIONAL EVALUATORS BETWEEN PEERS AND LAY PERSONS

Evaluator ➔ **Evaluation Goal➔** **The evaluated unit ↓(Degree of organisational complexity)**	Peer **Quality**	Professional evaluator **Quality /relevance/ efficiency**	Lay person (NGO)/politicians (Agora) **Social relevance-quality**
Individual	Peer review of articles for publication, appointments, awards	Research management, control with researchers productivity, human resource management	Participation in public debates, appearance in public affairs, popular publications
Institute/Department	Modified peer review Rating/quality assurance Peer advisory board	Research Management: - marketing - resource managing - benchmarking	User studies, interest groups in the field, training of young researchers, production of new knowledge
Institution	Modified peer review Rating	Contract steering- Negotiation of conditions (contracts) Resource allocation Benchmarking TQM	Overall user/interest group studies, evaluation of quality, employers of candidates (business), users of new knowledge (business, politics)
Cross-institutional Evaluations (institutions, disciplines, areas)	Informed peer review Rating/ quality assurance	Research Policy formulation Evaluation and prioritation of areas, Accreditation; EQUIS, ISO9000	Social consequences, Public hearings, consensus conferences, Public or NGO advisory boards
Territorial/-nationstate, national Policies	Rating/Expert-statements, visions, scenarios, proposals	Welfare policies Evaluation and prioritation of areas, cost-benefit studies, forecasting	- Political debates, public discussion of visions, public or NGO advisory boards

The figure originally stems from a presentation on a REMAP seminar on management and research evaluation on the Copenhagen Business School, may 2001, " The dilemmas of Research Evaluation" , by Finn Hansson, LPF and Birte Holst Jørgensen, RISØ.

The figure does not seek to produce neither a complete and systematic picture nor overview. The idea is to show how the growing organisational complexity and the professionalization of evaluators produces a large number of new types of research evaluation and assessment and the role or place of institutions and organisations in the Agora.

It is a differentiated and complex picture of attempts to solve the basic problem: how to produce trustworthy or social robust knowledge on research quality and prospects to a wider public than the traditional research community. Trust in research evaluation was once based on the norms of the

scientific community itself (Merton) and the classic peer review system continues today to be important in many connections and situations of knowledge assessment, especially when it comes to publication. Beside the classic system of knowledge control (the left column), we have the general public discussion on science and assessment methods normally used here (the right column). In between, we have a large number of different and mixed attempts to solve the problem of producing trustworthy knowledge on the value of research to specific goals.

Assessment or evaluation of scientific knowledge has always been about how to establish trust in knowledge. The problem we are facing today is how to establish trust in scientific knowledge claims in modern societies and what kind of procedures and institutions are needed when. Or to use the concepts introduced by from Nowotny, Gibbons and Scott 2001, how to establish 'social robust knowledge' in or through the Agora.
In the next section, we will present four different kinds of knowledge assessment to exemplify the differences in the three columns of the figure and to show the variation and the width of research evaluation today and in the years to come. The four examples are selected in order to illustrate the growing interrelation between science and society which also manifests itself in a process of social and political influence and/or changes in the science assessment processes**.**

Four examples of research evaluation

1) The *peer review* in science is based on the idea of having the quality of a scientific article evaluated by some of one's peers and it has a story of more than 200 years. In this period, the system of peer reviews developed to an important instrument in evaluation of research quality - unquestioned as an instrument for quality measurement in research in the 20th century, when industry and state required a great deal of natural and social science knowledge (Merton & Zuckerman 1971). The very complicated goal of selecting 'new' knowledge and knowledge of 'high quality' was performed mainly through the peer review system. The peer review system is closely related to a disciplinary structure in science, mainly because the peers have to be specialists or experts in the field they are evaluating. The system is based on qualitative evaluations and by the very nature of the whole arrangement, the evaluations they render are subjective and therefore in the end open to discussion. This took place to an extent that the peer review system over a period acquired a reputation of being able to select the best quality in research in a way close to that of the legal system. But peer reviews are not easy to administer in a world where the demands for control and assessment of research are growing and therefore demand peer reviews in a large scale and as an administrative tool for a number of new applications.

The most important prerequisites for a good peer review are motivation, competence and independence among the reviews (Kostoff 1999, Hemlin 1996). But this is not all there is to it. The cost problems are important too both in the form of direct costs but also and maybe most importantly in the amount of time consumed by peer reviews (Kostoff 1999). Beside the administrative problems and demands, new knowledge is difficult to evaluate in a world of competing knowledge paradigms and a growing integration of knowledge production in new organisational settings outside the universities (Mode 2 Science (Gibbons et. al.1994), Triple Helix (Etzkowitz and Leydesdorff 2000)). Considering these problems, the classic version of the peer review system has been under scrutiny as the major evaluation model in the field of research.

2) The *modified peer review model* is an example of an attempt to modernise the classic peer review model. It has been developed over a number of years in Denmark. Evaluation of university research

was a political demand by the Danish government in the late 1980's. This was partly due to several OECD studies and publications on the necessity of developing a national state research policy and partly to the overall implementation of the general concepts of auditing and quality management from NPM. The evaluation resulted in the implementation of a model more or less as the result of compromises between the central government and the universities. The 'compromise' was the "modified peer review model", an open model for evaluating research, where evaluation criteria, evaluators and specific procedures are open to negotiation. At the Copenhagen Business School, it took the form of an assessment of research quality and potentials at the various institutes – based on a selection of publications by the institute presented to and evaluated by external peers combined with site visits (Foss Hansen and Borum 1999, 2000).

After a few years of experience with evaluations conducted after top-down decisions from the Ministry of Education, universities took over this new modified form of evaluation and used the results more and more as an internal instrument for organisational development. According to Foss Hansen and Borum (1999), the newest or youngest universities in Denmark have been remarkably involved in initiating this bottom-up evaluation model, probably in order to legitimate "their status and position in the academic realm." (ibid. p. 311). But as it often happens with new methods in evaluation after a couple of years, this modified model became more or less institutionalised or standardised as the internal organisational standard for research evaluation. An open and less formalised model as the "modified peer review model" ends up in entangled procedures and standardisation as time goes by. The modified peer review model is based on a traditional standard, the peer review, and it is primarily a standard for organising institutional evaluations. But a combination of pressures for cost-reduction and development of routines as the natural result of a bureaucratically organised activity will after a while result in a more formalised and standardised 'modification' of the modified peer review model. For instance, the costly peer reviews tend to be replaced by quantitative techniques, especially bibliometrics (Kostoff 1999).

3) The development toward *standardisation of evaluation procedures* at the universities and public research organisations has introduced a number of new systems and imported systems for quality control from private business. One of the most visible developments throughout Europe is the work to implement an accrediting system for universities, the European Quality Improvement System. "EQUIS is an international system of strategic audit and accreditation designed by Europeans for the assessment of institutions in widely different national contexts. Although, it is inspired by the particular needs of the European situation characterised by extreme cultural diversity within a large geographical area, EQUIS is not limited to European problems in its scope. The standards are those of effective education for international management and apply to schools in any cultural environment whether in Europe or outside Europe."[2]

This type of quality assessment is based on a combination of site visits and presentations of a number of key statistics on the institution and its services evaluated by an expert panel. All information is then collected and evaluated by the accrediting organisation which is in charge of issuing the accrediting diplomas to organisations and institutions.

[2] quoted from the cbs-web. The EQUIS was "launched in 1997 by the EFMD which has championed the issue of quality in management development for many years and has brought together the key operators in Europe in a common initiative. As Europe's largest network association in the field of management development, with some 400 members from academia, business, public service and consultancy in 40 countries of Europe and the world, EFMD plays a central role in shaping the European approach to management education".

Benchmarking is a slightly different approach to improve the quality in organisations. Originally coming from private business, benchmarking is a tool to improve the function of a system by analysing the best practice by other companies in order to determine standards of performance. Next step is then to implement procedures in order to reach these standards. Total Quality Management, TQM, is another system for quality improvement coming from the private sector. TQM can be described as the application of quality systems on all levels in an organisation in order to continuously improve performance. TQM is really a number of different techniques and ideas used to focus the organisation on specified goals and focus on product quality. In a review of the impact of the TQM movement on R&D, Chatterij and Davidson (2001) conclude that "the greater participation of R&D in the vital business processes of the company, and its adoption of the more business-oriented management practices have been two of the most natural but powerful outcomes of the TQM movement."

4) The upcoming of *Mode 2 science and the Agora* has crossed the traditional borders defined by the 'legend of science' (Ziman 2000). How do the changes on the micro and macro level of knowledge assessment at the universities and research organisations function and how do they influence scientific knowledge assessment in the public sphere? And how does the public sphere assessment of scientific knowledge influence the internal scientific procedures[3]? These very complicated questions cannot be answered here, but the problem can be illustrated by a case story in order to demonstrate the methods used to assess scientific knowledge in public debates[4].

The multinational company Monsanto had for a period up to the summer of 1999 argued for permission to market GMO and had presented a large number of peer reviewed scientific evidence on the revolutionary effects of GMO in different basic foods and on its safety, the EU science advisory system had discussed risk and contamination limits trying to assess the long term environmental effects but they did not have neither resources nor access to control the knowledge based on laboratory experiments presented by Monsanto. In June 1999, several EU nations issued a joint declaration stating their intention to 'take steps to insure that any approval for the sale or release of genetically modified organisms (in the public discussion named GMOs) in Europe will be suspended.' This declaration was the basis of what has come to be known as the 'de facto moratorium' on genetically modified organisms in Europe. The background of this decision was a long public debate on the possible risks in relation to GMO. The debate was heavily influenced by the uses of scientific knowledge by proponents and opponents of gene technology in the context of European agricultural policy.

In the debate on GMO in EU, proponents of biotechnology deployed scientific knowledge in order to solve the problem of public and political resistance and uncertainty using arguments such as 'The best scientists in the world have shown these means to be safe.' This argument was normally combined with (often quite passionate) appeals to the potential benefits of biotechnology, for example: '...researchers are working not to boost farm incomes, but to save the lives of their people.'

The opponents, the Europeans in general and in a later stage several European governments exhibit widespread scepticism with respect to the food they buy. This was very clearly demonstrated in a Eurobarometer survey (1998) about the attitudes of consumers towards food safety: "In general,

[3] The problem to be discussed here is slightly different to the discussion that has been going for some years under the label 'the public understanding of science' (Wynne 1996, Yearley 2000).
[4] The case is based on Hansson and Horst 2000.

there is a relatively low level of confidence in producers, whether because of the information they provide about the product or, more importantly, because of the fact that they sell products that are profitable, but not safe."[5] An earlier survey also showed that there is no simple relationship between a public being sceptical and a public being more or less knowledgeable. For instance, while the Danish public is relatively knowledgeable, it is at the same time rather sceptical compared to other European publics. A study of the Danish media debate on the GMO discussion in Europe showed, that the skepticism or lack of trust in the scientific knowledge claims from the proponents was based on a reasonable level of factual knowledge (Hansson and Horst 2000).

The opponents' conception of scientific rationality was fundamentally different from that of the proponents. In the eyes of the opponents, the researchers have only a limited ability to predict and control, and therefore only limited rationality (Beck 1999). This limited rationality was seen as being in glaring contrast with the wide-ranging consequences that gene technologies could bring. The opponents, thus, did not share the proponents' optimistic perception of the inherent rationality of technological progress. On the contrary, the rationale of the opponents' arguments was that the risks associated with a global technological experiment were far too great to be justified by the advantages of GMOs.
It is worth noting that the rationale behind the creation of scientific knowledge was not often put up for discussion.

The lines of division between opponents and proponents did not run between accepting and rejecting scientific rationality, but between providing answers to two rather different questions related to the assessment of the knowledge in question. First, who appeared as a trustworthy representative of scientific rationality? And second, how broad was the sphere of application for scientific rationality?

What does this case tell us about knowledge claims in complex situation with many different stakeholders?

The public discussion and confrontation between experts and counter-experts over the interpretation of scientific knowledge and its consequences, the increasing use of alternative expert opinion and surveys in public debate by social movements and organizations are signs of the softening of structure of authority within science and of a lack of social and political trust in scientific knowledge.
The many participants in the process - the multinational corporations, scientists, scientist-adminstrators (eu), politicians (national and eu), social movements/grass root organizations, the press (media) can be understood as separate and interacting actors in the Agora. (Nowotny et.al 2001) .
But the case illustrates some of the structural difficulties in how to establish social robust knowledge. It is not only a question of organization and involvement of different parties or stakeholders. The knowledge in question in this case was not really social robust knowledge, or Mode 2 knowledge, the knowledge claims in question was about the limits or borders for scientific rationality on GMO knowledge. The question of contested knowledge or constitution of scientific knowledge did not play any significant role in the case. The questions on borders of knowledge or 'our unawareness to know' (Beck 1999) did not surface in this case.

[5] Eurobarometer 49 - Food safety, 3.9.1998

The trust in and within science is today normally based on institutions that are connected to the idea of an autonomous and disinterested science. If science shall engage in the agora and face all the financial interests and power games and at the same time keep the public trust, it demands a change of perspective on the new methods and a cautious implementation of them.

The above four examples exemplify the diversity and complications that many new assessment methods raise. Diversification and complications that the stakeholders need to be aware of and to explicitly address in order to gain the full value of the assessments. Otherwise, the assessment efforts can be counter productive and result in less trust in knowledge. But this new and explicit understanding of the different assessment methods rest on some preconditions.

Preconditions and conclusion

The new evaluation methods offer themselves to the scientific community as part of the answer to the question of establishing 'social robust knowledge'. However, both the practical utilisation of these methods and the economic and political interests that usually follow them call for some caution. The involvement of science in the agora is untrodden territory. The trust in and within science is today based on institutions that are connected to the idea of an autonomous and disinterested science (Wenneberg 2001). If science shall engage in the agora and face all the financial interests and power games and at the same time keep the public trust, it demands a change of perspective on the new methods and a cautious implementation of them.

Evaluations are not just a practical tool for implementing "the organised scepticism" of science and securing the truth of knowledge. Evaluations can also be seen as an important technique in organisation development processes and therefore a management tool. In fact, management is becoming increasingly important in research. Effective resource management is necessary if the trust of the paying stakeholders of science shall be contained. Management as a way of protecting the researchers of the power and money struggles that are a defining characteristic of the agora is also important if the trust in the scientific results is to be retained. Management should, therefore, not be seen as a threat to the researchers' autonomy and disinterestedness, but as a "taking care of the whole" that sees both the interests of the stakeholders in the agora and the researchers' desire for autonomy as legitimate. The primary research management task is to balance these two very different concerns (Erno-Kjolhede 2001).

If science's involvement in the agora shall produce "the socially robust knowledge" hoped for (Nowotny et. al. 2001), it is also of paramount importance that the entanglement between science and society does not produce a huge "organised irresponsibility". A clear and visible management of this meeting can also minimise this risk. To sum up, we think, as stated in the introduction that it is of utmost importance that we begin to study management and organisation of the micro-processes of the agora. A recent study of the various concepts of quality control between mode-1 and mode-2 research can be seen in Fujigaki & Leydesdorff (2000).

In this paper, as also stated in the beginning, we have focused our analysis on how science and research organisations have reacted or will be forced to react in order to demonstrate social relevance or robustness of scientific knowledge claims. We have not taken up the discussion of the many and very complicated problems that arise when these knowledge claims leave the systems of science and research and are taken up by other systems and organisations – not the least the media. These latent problems also have to be addressed by future research.

Literature:

Albæk, E. (1996). Why All This Evaluation? Theoretical Notes and Empirical Observations on the Function and Growth of Evaluation, with Denmark as an illustrative Case., *Canadian Journal of Program Evaluation*, vol. 11, (2), 1-34.

Beck, U. (1999). *World Risk Society.* Polity Press, Cambridge.

Boath, D. D. & Bodnarzcyk, M. (1995). Measures of Effectiveness. in G. W. Roberts, ed. *Quality Planning, Control, and Improvement in Research and Development.*(323-352) Marcel Dekker,

Bordum, A. & Wenneberg, S. B. (2001). Tillid som forskningstema. En indledning. in A. Bordum & S. Wenneberg, eds. *Det handler om tillid.*(7-10) Samfundslitteratur, Copenhagen

Chatterji, D. & Davidson, J. M. (2001). Examining TQM's legacies For R&D, *Research - Technology Management* (Janauary-February), 10-12.

Dahler-Larsen, P. (2000). Surviving the Routinization of Evaluation, *Administration & Society*, vol. 32, (1), 70-93.

Ernø-Kjølhede, E., Husted, K., Mønsted, M., & Wenneberg, S. (2001). Managing university research in the triple helix, *Science and Public Policy*, vol. 28, (1), 49-55.

Etzkowitz, H. & Stevens, A. J. (1995). Inching toward industrial policy.the university's role in government initiatives to assist small, innovative companies in the U.S, *Science Studies*, vol. 8, (2), 13-31.

Etzkowitz, H. & Leydesdorff, L. (2000). The dynamics of innovation: from national systems and "Mode 2" to a triple Helix of university-industry-government relations, *Research Policy*, vol. 29,109-123.

Foss Hansen, H. & Borum, F. (1999). The construction and standardization of evaluation. The case of the Danish University Sector., *Evaluation*, vol. 5, (3), 303-329.

Foss Hansen, H. & Borum, F. (2000). The Local Construction and Entactment of Standards for Research Evaluation. The case of the Copenhagen Business School, *Evaluation*, vol. 6, (3), 281-299.

Fujigaki, Y. & Leydesdorff, L. (2000). Quality control and validation boudaries in a triple helix of university-indutry-government: "Mode 2" and the future of university research, *Social Science Information*, vol. 39, (4), 635-655.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S. S. P., & Trow, M. (1994). *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies.* Sage Pubications, London

Hansson, F. (2001). Kan man have tillid til peer reviews? Evaluering af universitetsforskningen ved brug af peer reviews, *Vest*, vol. 14, (2), 7-30.

Hansson, F. & Horst, M. (2000). To trump the passions. The use of scientific knowledge in the introduction of new genetechnologies in Europe . Paper presented at *The Endless Transition,* Conference april 2000*,* Rio de Janeiro.

Hemlin, S. (1996). Research on Research Evaluation, *Social Epistemology*, vol. 10, (2), 209-250.

Holton, G. (1993). *Science and Anti-Science.* Harvard University Press, Cambridge.

INRA (Europe) European Coordination Office. Eurobarometer 49 - Food Safety. http://www.eu.int/comm/dg24/library/surveys/eb49_en.html [Eurobarometer 49], 1-2. 1998. European Commission DG24.

Kostoff, R. N. *Handbook of Research Impact Assessment*. Office of Naval Research [Seventh Edition] 1999. http://www.dtic.mil/dtic/kostoff/Handweb.html.

Merton, R. K. (1968). The Matthew Effect in Science. The reward and communication systems of science are considered., Science, vol. 159, (3810), 56-63.

Nowotny, Helga (1999). The Need for Socially Robust Knowledge. TA-Datenbank-Nachrichten 3/4 Vol. 8.

Nowotny, H., Gibbons, M., & Scott, P. (2001). *Re-thinking science.Knowledge and the Public in an Age of Uncertainty.* Polity Press, Oxford.

OECD (1997). *The Evaluation of Scientific Research: Selected Experiences.* OECD/GD (97) 194, OECD, Paris.

Power, M. (1997). *The Audit Society. Rituals of Verification.* Oxford University Press, Oxford.

Weiss, C. H. (1979). The many meanings of Research utilization, *Public Administration Review* (September/October), 426-431.

Wenneberg, S. B. (2001). Tillid til og i videnskaben.den brugte videnskabs kendetegn. in A. Bordum & S. Wenneberg, eds. *Det handler om tillid.*(172-182) Samfundslitteratur, Copenhagen

Wynne, B. (1996). Misunderstood misunderstandings: social identities and public uptake of science. in B. Wynne & A. Irwin, eds. *Misunderstanding science? The public reconstruction of science and technology.* (19-46) Cambridge University Press, Cambridge

Yearley, S. (2000). What does science mean in the "public understanding of science. in D. Meinolf & C. von Grote, eds. *Between understanding and trust.*(217-236) Harwood Academic, Amsterdam

Ziman, J. (2000). *Real Science. What it is, and what it means.* Cambridge University Press, Cambridge.