

## **Assessing Excellence in Social Sciences and Humanities: Concepts and Methods**

Prof. Peter Scott  
Kingston University

Kingston University  
River House  
53-57 High Street  
Kingston-upon-Thames  
KT1 1LQ  
+44 20 8547 7001  
[p.scott@kingston.ac.uk](mailto:p.scott@kingston.ac.uk)

*This paper was prepared within a Special Support Action “Research Collaboration in the Social and Human Sciences between Europe, Russia, other CIS countries and China” – Global SSH. The Action is supported by the European Community under the Sixth Framework Programme for research, technological development and demonstration (FP6) Thematic Area 'Citizens and governance in a Knowledge-based society' Call FP6-2004-CITIZENS-6 -- 8.3.3 – Promoting international research and policy cooperation in social sciences and humanities, contract nr. 028997.*

*Additional support to the Action is provided by INTAS under INTAS Thematic Call in Social and Human Sciences 2005, contract nr. 05-1000006-8446. More information on the Action is at [www.globalsocialscience.org](http://www.globalsocialscience.org).*

## Introduction

This short paper explores the concepts and methods of assessing excellence in the social sciences and the humanities. Its predominant focus is on research. However, an important characteristic both of mass higher education systems and of more distributed research (or ‘knowledge production’) systems is that once discrete activities, such as ‘teaching’ (or ‘higher education’) and ‘research’, have become increasingly fuzzy. As a result this paper will also consider the social sciences and humanities as wider intellectual domains.

There is a growing tension between the two strands of this paper – the conceptual and the methodological:

1. At a conceptual level the possibility of defining ‘excellence’ is becoming more problematical as a result of two phenomena:
  - i) The first is that the core identities of the social sciences and the humanities have become more diffuse – partly as a result of the internal evolution of these disciplines; and partly as a result of changes in the external environment.
    - a) The former process is familiar enough. As disciplines evolve, sub-disciplines emerge, establish their own independent identities and create their own institutional structures (university departments, research institutes, doctoral programmes and so on); they also coalesce into broader, and sometimes more interdisciplinary, fields which may be more weakly institutionalised. Even when the formal taxonomy of disciplines remains comparatively unchanged, a constant process of intellectual renewal is at work. The force that is shaping and reshaping disciplines in these ways is assumed to be the constantly shifting research agenda.
    - b) The latter process, the impact of the external environment, is also (apparently) straightforward. Social change has led to the emergence of much more complex social forms – and consequently raised the educational threshold for effective participation in contemporary society. At the same time the evolution of the labour market has raised the initial skills threshold required to participate in the workforce. Together they have created new demands for graduates (a ‘pull’ factor) while the expansion of higher education has led to the replacement of formerly ‘lower-level’ occupations by graduate professions (a ‘push’ factor).

Once it seemed to be possible to separate these two sets of influences – the research-led intellectual dynamics that shaped the ‘interior’ life of the social sciences and humanities; and the market-led ‘exterior’ forces that stimulated the demand for new academic programmes and qualifications. As a result the social sciences (to a greater extent than

the humanities) were seen as both ‘academic, i.e. research-led, and as ‘professional’, i.e. servicing the workforce needs of an increasingly complex society. Furthermore it was sometimes argued that these two aspects were not only conceptually distinct but could also be distinguished in organisational terms (for example, in so-called ‘binary systems’ composed of traditional universities providing academic, or scientific, education and higher professional schools offering vocational programmes). In fact the relationship between ‘academic’ and ‘professional’ strands was always highly iterative – because the research agenda, certainly in the social sciences, responded to new social and economic phenomena which, in turn, had often been helped to emerge as a result of earlier research.

- ii) The second phenomenon that has made it more difficult to define ‘excellence’ in the social sciences and humanities is closely linked to the first. It is the emergence of more distributed research systems (or, more accurately, knowledge production – or innovation, even creativity, systems). These new systems (although ‘systems’ may imply a degree of deliberate organisation which is not always present) have a number of characteristics:
  - a) One is the multiplication of sites where serious research is undertaken – both in a spatial and quantitative sense because advances in information and communication technologies (and other, more cultural, trends related to globalisation) have made it easier to mobilise large research groups; but also in a qualitative sense because these groups are more heterogeneous and include ‘non-academic’ members (for example, in Ministries, ‘think tanks, management consultancies and so on). The boundaries between the academic (or research) system and other knowledge production systems have become more porous;
  - b) A second characteristic is that the roles of the various research actors have become more fluid. It is no longer always possible to distinguish between the ‘producers’ and the ‘users’ of research because their relationship is often highly reflexive (and the key actors are often those who mediate between these two groups);
  - c) A third characteristic is that the cognitive core, both epistemological and methodological, of many disciplines has been broadened, opened up, even invaded. This has happened in two ways. The first is the multiplication of research sites and fluidity of research actors have tended to undermine the normal processes of socialisation and institutionalisation which have created the existing taxonomy of academic disciplines; instead they have encouraged the development of multi- (even trans-) disciplinary cultures, if not structures. The second is that

‘fringe’ or ‘alternative’ knowledge traditions have invaded, and even been incorporated in, the cognitive core;

- d) A fourth characteristic is that traditional forms of quality control, or assessing excellence, are no longer sufficient. At the very least they need to be supplemented by other methods – relating, perhaps, to social impact, economic utility or cultural significance. It may even be that they need to be superseded entirely. ‘Excellence’ needs to be unpacked under these new conditions.

2. But, if at a conceptual level, it has become more difficult to define ‘excellence’, at a methodological (and operational) level the pressure to set targets, raise standards, undertake formal evaluations and assessments and audit outcomes has increased. There are several reasons for this growing pressure.

- i) One is the growth of the so-called ‘audit society’, a phenomenon with which social scientists are familiar as a research topic. This is one aspect of the transformation of the 20<sup>th</sup>-century welfare state into the 21<sup>st</sup>-century regulatory state. As Governments have struggled (or, under the influence of neo-liberal ideology, been unwilling) to maintain the elaborate public service infrastructure developed after 1945 and, in particular, after 1960 (or, more accurately, as shifting demographic patterns have made it much more expensive to maintain this infrastructure), they have ceased to define themselves as the providers of services but instead have taken on new roles as the purchasers and/or regulators of these services. The extent to which this shift from the welfare state to the regulatory state has been substantial as opposed to rhetorical has been much debated. National statistics suggest that the key shifts have been within the public sector rather than between the public and private sectors. However, another feature of the regulatory state is that it impacts as much, or even more, on the private sector – partly because of the proliferation of regulations on matters such as health and safety, the environment and equal opportunities; and partly because of the advance of privatisation which has produced a large ‘grey area’ between the public and private sectors (to such an extent, indeed, that distinguishing between the ‘public’ and the ‘private’ domains in contemporary society has become increasingly problematical. ‘Audit’ has become a dominant discourse (and also practice) in the regulatory state.

- ii) Another reason is the emergence of the so-called ‘knowledge society’. The more than ‘knowledge’ is defined as key (and primary) economic resource (and also an equally significant social and cultural resource), the more societies (and Governments) seek to ensure they are securing value for money from their direct or indirect investment in ‘knowledge’ institutions. Just as students, and graduates, are encouraged to consider the ‘rate of return’ on their higher education, so at a higher level similar considerations are given increasing weight in

contemporary political discourse. It does not matter that ‘knowledge’ is often poorly defined – ranging from data, images and brands through to highly skilled graduates and university research. As with the ‘audit society’ a discourse has developed around the ‘knowledge society’ which encourages politicians and other policy makers to regard higher education and research in more immediately instrumental terms – which, in turn, stimulates the demand for assessment and measurement. The ‘knowledge society’ has also become a kind of battle-ground, an arena in which nations (and regions and cities) struggle for competitive advantage – which has also stimulated the demand for ‘excellence’;

- iii) A third reason, of course, is the desire to combat the complexities and transgressions that have been produced by the development of mass higher education systems, in which the production of highly skilled graduates and professional workers is only one goal among many, and the emergence of more widely distributed knowledge production and innovation systems, of which (conventional) research is only one component. In other words, the drive to measure, evaluate, assess and audit is a reaction against those changes in the academic and research systems which have made measurement, evaluation, assessment and audit more problematical. Linked to this is an intriguing shift from self-policing systems and institutions, characterised by high levels of mutual trust (often between elite groups), to self-organising systems and institutions, which present a greater challenge (even threat) because they are difficult to define and unpredictable. The desire to identify ‘excellence’ is part of a wider desire to categorise, neither of which can easily be reconciled with the dynamics of disciplinary evolution.

This emerging tension between concepts and methods, therefore, must inform any discussion of how best to assess excellence in the social sciences and humanities. Of course, the degree of tension varies between disciplines. Typically research (and scholarship) in the humanities may have been less influenced by the development of more widely distributed knowledge production systems (although it can be argued that it has always been an intellectually reflexive and socially embedded activity); but the humanities have played a key role in the massification of higher education. In contrast the impact of massification on some of the social sciences has been less – at any rate, in the past two decades – but social science research has been more deeply implicated in so-called ‘Mode 2’ knowledge production. However, generalisations of this kind are dangerous. Perhaps it is enough simply to be aware of the tension between, on the one hand, the growing difficulty of defining ‘excellence’ in the more open academic and research systems that now prevail and, on the other, the drive to measure, evaluate, assess, audit – and rank.

The rest of this paper is divided into three sections:

- Social sciences and humanities in context: a brief review of the major intellectual currents and logistical trends which have shaped the social sciences and humanities.

- Unpacking ‘excellence’: an analysis of the major elements within an up-to-date and acceptable definition, or conception, of high quality.
- Changing methods: a discussion of the most important methods which can be used to assess ‘excellence’ across these disciplines.

## **Social science and humanities in context**

Any attempt to summarise in a few hundred words the key intellectual trends in the social sciences and humanities is certain to be unsatisfactory. Only slightly less satisfactory will be an attempt to summarise the key logistical trends that have influenced the development of these disciplines over the past 30 years. But, however unsatisfactory the results, both attempts may be necessary in order to understand shifting conceptions of ‘excellence’ in social sciences and humanities – and, in particular, to understand the wider context in which such conceptions have been framed.

### *Intellectual trends*

Some of the key trends, especially in the social sciences but also apparent in the development of many humanities disciplines, reflect wider socio-economic, political and cultural changes.

- One has been the rise (and now the fall?) of the post-war welfare state which fed the growth of the social sciences – in terms of the supply of students and the formation of new professional fields, but also as an extended research domain and even as a normative system. As a young journalist I remember being told by a right-wing politician in Britain: ‘sociology, social work, socialism – they’re all the same thing, aren’t they?’ Closely linked to (rather than being, as is sometimes supposed, opposed to) the welfare state has been the development of new economic structures and employment patterns which replaced the regular contours of industrial society. A good example is the emergence of the creative and cultural industries (design, film and TV, ‘higher’ forms of tourism and so on) which have transformed the status of the humanities in the modern university (for the better – and worse);
- Another fundamental change has been the reconstitution (partly as a result of the development of the welfare state) of traditional 19<sup>th</sup> and early 20<sup>th</sup>-century ‘class society’ (I say ‘reconstitution’ because social inequalities have not disappeared); a bourgeois (or middle-class) culture has become pervasive in many so-called advanced societies. So-called individualisation is one aspect of this change; but so too is the rise of a mass consumption society with its new cultural and visual landscapes;
- A third change is the emergence of new ‘social movements’ led by the women’s movement in the 1960s and 1970s but now espoused by many other social groups which have often been self- rather than structurally defined (experientially shaped identities rather than objectively observed categories).

The impact of these new 'social movements' has been reinforced, and magnified, by the growth of multi-culturalism which has produced new layers of social complexity.

- A fourth change, of course, is the development of mass higher education systems which has produced very substantial minorities, even majorities, of highly educated people in most developed countries (and developing countries are not far behind). Contrary to the hopes and expectations of many politicians it was the social sciences and humanities, not the natural and applied sciences, which have been key drivers of massification. At times the emergence of mass higher education has seemed like a gigantic social science experiment (although not one that mainstream social scientists have been enthusiastic about analysing).

These changes (and others) have led to important shifts in the intellectual agenda of many social sciences (and some humanities) disciplines. Of course, generalisations, although easy to make, are difficult to sustain. But four significant (but, not necessarily contemporaneous, shifts can perhaps be identified which can be observed in at least a majority of the social sciences.

- i) The first is (was?) the waning of ideology – not so much in terms of the 'end of ideology' as first celebrated in the 1950s and 1960s but more as an effect of the professionalisation and institutionalisation of social science research (as a result of which the role of the 'intellectual' has been diminished and the status of the 'expert' – and her/his empirical methods of inquiry – elevated). Paradoxically the waning of 'ideology' may have been counter-balanced by the rise of 'theory', although the role of theory in the social sciences (and still more the humanities) has been kept in check in some countries and disciplines by this expert-empirical research culture and in others by the fecundity, and volatility, of 'theories';
- ii) A second has been a waning of interest in 'structures' and a waxing of interest in 'cultures' – or, perhaps more accurately, the re-description of structural factors in terms of cultural change. Even when the language of 'systems' and 'structures' continues to be used, it is now more likely to be used in softer, more contingent and more pliable ways; it has become another 'discourse'. The influence of anthropology on its fellow social sciences is an example of this shift at a more fundamental level, while the popularity of cultural studies is an example of the same at a more popular level;
- iii) The third, and more recent, shift has been (back?) from more empirical methods (and positivistic values borrowed from the natural sciences?). One way in which this has been experienced is the (comparative) decline in quantitative research in many of the social sciences and the rise of more qualitative studies (to such an extent that recently in England quantitative social science has been added to the list of 'strategic and vulnerable subjects'). Another way is the (alleged) spread of highly relativistic post-modern ideas; (which has probably been exaggerated, although it is revealing that the 'paradigm shift – normal science' ideas of Thomas Kuhn

have been embraced much more enthusiastically by social scientists than by his fellow natural scientists);

- iv) A fourth trend has been the complex inter-weaving of subject (and even topic) specialisation, as pronounced in the social sciences and humanities as in other disciplines, and inter-disciplinary imperatives. The boundaries between the social sciences and humanities disciplines have perhaps been more porous than those between the natural sciences (maybe because they are defended by the complex and expensive infrastructures of experimental science). But the characteristics of social science research subjects have also encouraged eclectic approaches. To some extent the humanities have been ‘captured’ by the social sciences in their more ‘cultural’, qualitative and inter-disciplinary guises (although in some countries it has been the other way round, notably France because of history’s pivotal position as the ‘queen of the social sciences’).

#### *Logistical and organisational trends*

Three important logistical and/or organisational changes can be observed – the impact of the massification of higher education (which has already been mentioned); and the influence of new research environments and cultures.

- i) The first has already been mentioned – the growth of mass higher education systems. It was always anticipated that the social sciences were likely to play an important role in massification (partly because of their affinities – both were aspects of the opening up of social and intellectual life in the second half of the 20<sup>th</sup> century – and partly through an accident of timing – the social sciences ‘came of age’ at the time when access to higher education was being rapidly expanded). But fewer people had anticipated that the humanities too would become ‘mass’ subjects. Because of this transformation of the higher education (or teaching) base of the social sciences and humanities, their research base has also been substantially modified – positively, in terms of the development of new research topics and (more fundamentally) knowledge traditions; but, more negatively perhaps, because it has become more difficult to define excellence in this more open environment. In contrast the research base in science, engineering and technology (SET) has been affected less by the massification of higher education – despite the more linear dynamics that characterise research in these disciplines. As a result notions of excellence have been less contested;
- ii) The second important change has also been mentioned – the trend away from large-scale quantitative research enquiry and towards smaller-scale and more qualitative projects. This reflects not only a trend within the ‘inner life’ of many of the social sciences (and humanities, although quantification has always played a subordinate role within scholarship in the humanities; the fate of so-called ‘cliometricians’ is a good example of this ambivalence towards quantification). Nor does it simply reflect a drift towards a wider definition of research, both research actors and research practice, which has already been discussed in the first section of this paper

and which, arguably, has had a greater impact in the social sciences (and the humanities?) than in SET disciplines. This trend also reflects a change in the economy of social science research. Fewer large research grants are available, and the competition for all forms of research funding has increased. In an important sense this is a logistical as much as an intellectual, or scientific, shift. But its effect, once again, has probably been to make it more difficult to evaluate the quality of research outputs which have become more eclectic (and even more incommensurable);

- iii) The third big organisational change has been the shift towards the 'market'. This is apparent not only in the competition for research funding (and for research 'profile' through successful publication strategies) and in the variation among the resource bases enjoyed by different disciplines (which has been perhaps over-dependent on student choice). Neither is a new phenomenon, although both have been legitimated (and intensified?) as the result of the espousal of 'market' policies in higher education and research. This more competitive environment has had a substantial impact of the behaviour of researchers – and, in particular, on their patterns of publication and strategies for dissemination. Of course, it can be argued that competition drives up standards (and it is probably true that it punctures complacency). But it also encourages patterns of behaviour such as the development, and vigorous propagation, of 'brands' – for individual researchers, research groups and institutes, institutions and even national systems. At its worst such behaviour can lead to plagiarism and falsification of research data. Even if these dangers are avoided, there remains a risk of premature publication, and exaggerated 'selling', of research findings. Such behaviour makes it more, rather than less, difficult to measure excellence – despite the proliferation of 'rankings' of all types.

### **Unpacking excellence**

The changes in the wider social, and higher education and research, systems described in the first part of this paper have combined with the more specific changes in the intellectual agendas social science and humanities research (and its logistical and organisational environment) to produce a situation in which definitions of excellence have become more problematical – but at the very time when the pressure to identify and reward excellence has intensified. In the third part of this brief paper I will try to 'unpack excellence' by identifying its constituent elements.

There are four major elements in any definition of excellence in research – outputs (mainly publications in the social sciences and humanities); methodologies (in both a detailed and specific sense within particular disciplines but also including ethical considerations); sustainability (the reproduction of research capacity); and relevance or applicability.

#### *Outputs*

The first element of excellence is the quality of research outputs – books, articles and other publications. Two questions arise; the first is how these outputs should be defined (the 'what') and the second is how their quality should be assessed (the

'how). Neither is as straightforward as it appears. The 'how', methods for assessing quality, will be discussed in the next section of this paper so at this stage attention will be concentrated on the 'what', which outputs should count. In a traditional 'Mode 1' context the answer seems to be simple – academic books (rather than 'popular' books or textbooks designed predominantly for use in teaching) and articles in refereed journals. But even in this restricted domain difficulties have arisen. For example, the distinction between an academic monograph and a more popular work in, say, history is not always clear; much may depend on the prestige of the author and/or of the publisher (although little attention is typically paid to the influence of the market imperatives that are as strong in academic publishing as in other parts of the industry). In some social sciences the distinction between broadly conceived and synoptic surveys of major topics and textbooks designed for teaching purposes is also not easy to draw. However, in general terms, it has been possible to accept that such judgements can be left to the relevant academic community (even if this may tend to discriminate against more 'adventurous' types of research output).

Within the more open knowledge production (and scholarly publication) systems which now prevail new difficulties have emerged. One, which will be discussed in the next section of this paper, is that academic communities have become more diffuse which has tended to increase the contestability of their authority. But a second difficulty is the definition of potential research 'outputs' is also being stretched. For example, the growing emphasis on the so-called 'creative and cultural industries', a major focus of development in most advanced economies (and post-industrial societies), has tended to make more 'popular' forms of scholarly production more significant, especially in the humanities. Or, to take another example, the proliferation of policy and evaluation studies in some social sciences has tended to place greater emphasis on new kinds of 'outputs' such as consultancy reports. In the past the view may have been taken that, in order to count as research, such reports had to be translated in academic monographs or referred articles (rather as keynote addresses, even at academic conferences, had to undergo a similar process of translation). But today important forms of intellectual production may never undergo such a process of translation. Teams of evaluators / commentators / researchers move on to their next projects because they have no time, or inclination, to translate their findings back into appropriately academic contexts – yet new knowledge (and often valuable knowledge) has been created.

There are two issues here:

- i) The first issue is the stretching of appropriate forms of academic output, what counts as research. This stretching has three components – actors (the number and type of 'producer' have both expanded); content (for example, can 'popular' books, policy studies, evaluation reports and the rest be counted – even if they have not undergone any process of academic translation); and form (because even in more traditional research domains new forms of publication are emerging – for example, open-source or pre-prints – largely as a result of advances in information and communication technologies. It is interesting to reflect on the extent to which current definitions of research 'outputs' are contingent in particular technologies of scholarly communication);

- ii) The second issue is to what extent it is possible to define a research process as in itself an ‘output’ or must that label be reserved for products in amore conventional sense. This issue can be approached at two levels. The first level is pragmatic and organisational. For example, when do pre-prints of research findings, circulated with the deliberate intention of provoking critique (which probably requires these findings to be modified) become firm ‘outputs’? Although this difficulty is not entirely new, the dynamics of research production now make it possible for findings to be left ‘open’ for much longer (and possibly for ever). The second level is more fundamental. Previously it may have been possible to agree that research findings, or more generally new knowledge, were sufficiently ‘complete’ to be presented in a ‘final’ form (in other words the ‘process’ had been translated into a ‘product’). But the acceleration of research has combined with the indeterminacy of knowledge (and new communication technologies) to undermine these processes of ‘closure’.

### *Methodologies*

The second element of excellence is methodologies. In the past this too was – relatively – unproblematic. The dominant methodologies in particular disciplines were well understood – the use of archival research in disciplines like history and of surveys and fieldwork in the more empirical social sciences, the application of recognised interpretative (or ‘theories’) in some other humanities disciplines such as literature and in the more conceptual social sciences. Also academic communities were still rooted in élite higher education systems and they not only engaged in these recognisable (and respectable?) research practices but also subscribed to broadly similar values, which encouraged a consistent approach to ethical issues; as a result these communities possessed effective self-policing systems. Therefore it was possible, despite the debates (and even ideological dissonances) with the social sciences and humanities, to agree formal definitions of ‘excellence’ on these tacit understandings – most of the time.

Today important changes are taking place which may make it more difficult to agree a stable set of definitions of ‘excellence’ in the social sciences and humanities.

- i) First, there has been a proliferation of research practices. Not only have methodologies become more diverse they have also become less commensurable. The growing number of smaller-scale qualitative studies in the social sciences, which are often undertaken by researchers who are themselves practitioners, has already been discussed. In some circumstances this active engagement has been considered a positive asset by those who espouse so-called ‘action research’. But at the same time this proliferation of research actors may also have compromised not only their objectivity (in any case the degree of attainable objectivity may have been exaggerated in culturally and socially engaged disciplines) but also their expertise (because they may have followed less well defined research careers – and, for example, received less formal training in research methods). These increasingly diverse methodologies may also be more difficult to ‘rank’ in terms of their rigour and effectiveness.

- ii) Second, there has been a parallel proliferation of interpretative frameworks – not simply in quantitative but also in qualitative terms. In the past many of these frameworks, including sharply contested and rival ‘theory’, grew out of research practice – with the possible exception of the grand meta-discourses of modernity (Marxism, Freudianism, even Keynesianism). Although they were influenced by wider social, political and cultural agendas, these frameworks grew up ‘inside’ their disciplines. Today different, and even incommensurable, knowledge traditions are encountered which are more difficult to reconcile, or develop, through a careful and sustained process of rigorous empirical enquiry.
- iii) Third, of course, following on from these first two changes there has been an expansion in both the number and type of ‘researchers’, especially in the applied social sciences but also perhaps in some creative subjects (for example, the concept of ‘research’ has been imported into disciplines such as fine art, largely as a result of its tighter association with ‘mainstream’ higher education and its subsequent professionalisation (and academicisation?). There may also be, as has already been indicated, growing confusion between various roles within the research process; ‘producer’, ‘user’ and ‘disseminator’ have all become fuzzy categories. The result is a much larger, more diffuse and less coherent research ‘community’.

These changes have impacted not only expert and discipline-specific methodologies but also on wider ethical considerations and issues of social value and accountability. It has become more difficult to rely on tacit understandings about the values and practices which are appropriate (which may help to explain why explicit ‘codes of conduct’ and ‘good practice guides’ have been developed). At the same time it has become less possible to insist on particular patterns of behaviour (because any particularities are likely to be at odds with different knowledge traditions). Together they make it more difficult to arrive at satisfactory definitions of ‘excellence’ in terms of better-or-worse methodologies.

### *Sustainability*

The third element of excellence is sustainability, which is a very broad concept. Even in the humanities where individual scholarship is still the most common form of research there is a need for an adequate research infrastructure – in terms of career opportunities, library and other resources – and also a capacity for renewal and sustainability – in terms of the higher education and doctoral training of new scholars and researchers in adequately funded institutions (and, although more speculatively, the education of the wider population – in order to appreciate, derive advantage from, and agree to support future research in these disciplines). In many of the social sciences there is a further need to develop sustainable research groups in order to provide an effective environment for academic team-working and also to manage complex research projects and relations with potential stake-holders (and the equally complex infrastructure of research grants).

Sustainability, therefore, has several dimensions.

- The first is financial. There must be sufficient funding to pay the full economic costs of research in the social sciences and humanities. This applies not simply to external research grants which must cover the full, not simply the marginal, cost of research activity; it also applies to the funding of research centres, departments, faculties and indeed whole universities. This is not always simple to guarantee in the social sciences and humanities which tend to have large numbers of undergraduate students and are regarded as ‘low cost’ disciplines;
- The second dimension of sustainability is organisational. There must be an appropriate framework of institutions that are not only adequately funded but also sufficiently autonomous to be able to develop their own research agendas and / or to create an unconstrained ‘space’ within which individual researchers and scholars can develop their own agendas;
- The third dimension is concerned with people. First, it is important to have organisational contexts and structures within which the formation of the next generation of researchers can place. This requires strong doctoral programmes which are more difficult to establish and maintain in the humanities and (some) social sciences than, for example, in the natural sciences and engineering. Second, it is equally important that proper career opportunities are offered, and robust professional structures are created, in order to motivate and retain high-quality researchers;
- The fourth dimension is normative. To create and maintain a high-quality research base in the humanities and social sciences it is necessary to nurture values within the wider higher education and research systems which recognise the importance of research in these disciplines (in terms of their validity, rigour and independence); and also within the wider political system (and, wider still, in terms of public opinion) which recognise their value to society and the economy.

The changes that are taking place in higher education and research, and in society at large, which have been discussed earlier in this paper challenges some current assumptions about the right conditions for guaranteeing excellence in terms of sustainability. The most obvious challenges are to the first, second and third dimensions of sustainability – the funding systems, organisational / institutional environment and the professional structures which help to make research in the social sciences and humanities sustainable.

- i) First, research funding is now more likely to be determined and directed by quasi-market principles. As a result it may be driven more by the dynamics of market competition (i.e. the ‘price’ at which researchers and institutions are prepared to undertake projects) rather than the sustainability of the research base (i.e. the ‘cost’ of undertaking projects). Even if costs are taken into account, there is a danger that some social sciences and humanities research may be ‘priced out’ of the market;
- ii) Second, organisations have tended to be ‘hollowed out’ and institutions to become less stable and more permeable. As a result they may be less able to

support the panoply of systems and structures required to maintain a sustainable research base (although it is also possible that research may actually flourish within more adaptable systems and structures; it is also likely that more recently established – and therefore less prestigious – social sciences may benefit from the weakening of traditional organisational systems and institutional structures);

- iii) Finally, the destabilisation – or even erosion – of ‘careers’ may make it more difficult to maintain a sustainable, and potentially excellent, research base (more perhaps in the humanities because its scholarly culture depends more on the availability of stable long-term academic employment than in those social sciences which are more user to a research-project environment). On the other hand this trend may also make it easier for new players to enter the research arena if its effect is to lower the (professional) barriers to participation.

### *Relevance*

The fourth element of excellence is relevance, a broad (and often sharply contested) notion which itself needs to be as carefully unpacked as the concept excellence itself. It has been argued that any attempt to insist that research must be relevant is an attack on ‘open science’ – and, in the end, counter-productive. In practice such a purist perspective has proved to be difficult to sustain – because it could not be denied that any research system was, to some degree, socially constructed (in the sense that both the parameters of valid research, priorities for funding, institutional constraints and the wider intellectual environment were all generated within specific social contexts, which were themselves subject to change). As a result various compromises between ‘open science’ (and ‘blue-skies’ research) and considerations of social and economic relevance had to be negotiated. In Britain this necessary compromise was once summed up in the phrase ‘timeliness and promise’ – in other words, research had to be ‘timely’ in terms of potential relevance to society but also ‘promising’ in terms of its scientific potential’ ideally both conditions should be met.

Moreover, it was accepted that relevance entered into the research process at the beginning and the end – or, more accurately, it was accepted to be a relevant consideration in determining what kinds of research should be funded; and it was also accepted that the relevance of research products should be demonstrated through the dissemination process. But relevance was not only regarded as having no place in the core process of undertaking research but also as being potentially antithetical of producing high-quality research. In other words, far from being an element in the construction of excellence, relevance was regarded as a threat.

- In the humanities relevance has always been a particularly difficult concept. Although research in the humanities might be relevant in terms of its wider dissemination (sometimes in more popular forms), relevance was not seen as a legitimate consideration in deciding which research should be funded in the first. The reasons for this were three-fold. First, much of this research was (and still is) individual scholarship rather than team-based projects, which tended to emphasise the need for free choice of research topics. Second, of course, the potential relevance of research and scholarship in the humanities was more difficult to demonstrate (before the rise of the creative and cultural

industries). Third, social influences on research and scholarly agendas were oblique and indirect – and, therefore, less easy to identify. These influences were largely mediated through the massification of higher education, which transformed its social base, on the one hand and on the other the wider intellectual climate, the famous *zeitgeist*.

- In the social sciences it was more difficult to argue that these disciplines were not socially constructed, in terms of their basic taxonomy, their fundamental values and their routine practices. They were, after all, social sciences. So it has been more difficult to resist accepting that considerations of relevance should play some role in shaping research agendas. Apart from the most theoretical social sciences it was obvious that research topics were determined from ‘outside’ to a significant degree (and even social and economic theory was generated in, and by, a wider intellectual culture as much as discipline-specific academic cultures). Equally, once again with the (doubtful?) exception of theory, it has always been clear that one of the tests of the validity, and the quality, of research has been its broader relevance. But even in the social sciences the epistemological core of social science research (and, to a lesser extent, its methodologies) was treated as a ‘secret garden’, forbidden to external stake-holders however powerful.

More recently a range of new ideas has been developed which tend to place greater emphasis on the contextualisation of knowledge, not simply as a philosophical exegesis or a recognition that all knowledge producing systems are socially embedded but as a process that shapes, in very practical ways, the ‘inner life’ of academic disciplines as well as their external environment. These ideas have been particularly applied to the social sciences. They include the assertion that some knowledge is weakly contextualised (the natural and applied sciences) and other knowledge is strongly contextualised (the social sciences and humanities); that knowledge production is much more widely distributed; that controversies, public as well as academic, play an important role in the creation of new knowledge (both in the market place and in the social, cultural and political system – an aggregate that has been labelled the *agora*); that the epistemological core is variously empty (in the sense that a research domain is largely constructed out of its social practices) or brimming over (in the sense that it must accommodate an increasing variety of knowledge traditions); that valid knowledge must no longer simply be seen as reliable (in an experimental and empirical sense) but also as ‘socially robust’; and that it is generated not only within a context of application (potentially a constraining and narrowing context) but within a context of implication (by imagining the as-yet unimaginable). There is not space in this short paper to discuss these ideas in detail. But their cumulative effect is both to emphasise the contribution of relevance to assessing excellence in social sciences and humanities research but also to cast its influence in a much more positive light.

Outputs, methodologies, sustainability and relevance – these remain the major elements in assessing excellence. But they have all be modified to some degree by the changes in the higher education and research (or knowledge production) systems which have been discussed in the introduction to this paper and also by the intellectual and logistical shifts within social sciences and humanities research which were briefly sketched in the second part of this paper. The combined impact of these ‘external’

changes and ‘internal’ shifts has been to problematise still further the notion of excellence – but not, of course, to ‘abolish’ it.

### **Changing methods**

The final part of this paper has a more pragmatic focus, on the main methods used to assess excellence in the social sciences and humanities. But it is important to recognise that the choice between different methods of assessment (or, more probably, the balance between these elements within an overall assessment package) is not simply a pragmatic one – what works best. The assessment of excellence, like all measurement tools, is not a neutral process; there are competing ‘excellences’ (to say this is not to espouse crude relativism) and the choice between which to attach the greater importance to and which the less is not simply related to the formal objectives of the assessment process but also reflects the (maybe tacit) value systems which underlie these objectives. For example, it is possible to define excellence in the social sciences and, especially perhaps, the humanities as a process – or even as a scholarly culture that must be nurtured (which, in turn, may strengthen the arguments for the autonomy of the universities which provide the space, or *habitus*, in which such a culture can survive – and, therefore, may place strict limits on assessment if it is seen as an external audit as opposed to a self-policing process). But it is equally possible to define excellence as a product – although the products that might be most highly valued by scholars and researchers themselves may be different from the products that are most desirable to industry and business, or politicians and the public. All these complexities need to be borne in mind when discussing apparently pragmatic, and technical, choices between different assessment methods.

In principle there are three such methods – peer review (although expert review is a better description); ‘metrics’ (or input measures such as external research funding); and performance indicators (which measure outputs and impact).

#### *Peer review*

Peer, or expert, review is based on the sound principle that those best able to assess the quality of research (or teaching or other aspects of higher education) are those who are themselves active researchers (or teachers) and whose standing is generally acknowledged to be high by their ‘peers’ in their own disciplines. Peer review is almost universally used to decide which scholarly monographs and journal articles should be published; expert reviewers comment on the quality of work that is submitted to editors and publishers and advise them on whether it deserves to be published. Peer review is also widely used to determine which scholars and researchers should receive research grants; once again grant proposals are assessed, and graded, by expert reviewers (although other criteria than purely scientific merit are tending to be given greater weight – for example, the relevance of research proposals to broader social or economic themes). Finally peer review is also used in some national research (and teaching quality) assessment systems. For example, in the UK Research Assessment Exercises (RAE), five or six-year assessments of the quality of research at the level of university departments which have been undertaken since the mid-1960s, the heart of the process is the grading of individual research outputs by expert panels (although less emphasis is likely to be placed on peer review in future RAEs).

There are three main critiques of peer review:

- i) The first is that it works better in some disciplines (the traditional humanities and the natural sciences) than others (some of the more novel social sciences, creative disciplines and – possibly – technology). For peer review to be reliable a number of conditions must be met. First, there must be a well defined scholarly or scientific community – and this community must be sufficiently broad to enable a variety of topics and methodologies (and discourses / ideologies?) to flourish but not so broad that it becomes incoherent. Second, there must be an agreed definition of appropriate scholarly and scientific outputs – and they must be openly available (which may not be the case in some areas of ‘near market’ research). Third, there must be sufficient agreement about other stakeholders who may be able to make a useful contribution to assessing quality (of research and teaching programmes) and who also have a legitimate stake in the quality of these programmes (and their outputs). Clearly these conditions cannot always be met;
- ii) The second is that peer review works better in some contexts than others. For example, few people would argue that it should not be the dominant determinant of which articles are published in scientific journals (although other, quasi-commercial, criteria cannot be entirely excluded – and, in the case of books, these other criteria may be more significant). But in the case of larger-scale research programmes, although peer review may still be the major determinant, wider considerations may be given greater weight. There are even greater doubts about whether peer review should be allowed to over-influence national research priorities, which is the effect of using peer review to determine the selective distribution of research funding to universities (as in the UK RAE);
- iii) The third critique of peer review is that it is an inherently conservative system – in two senses. First, it is based on the existing taxonomy of disciplines (which is inevitable because it is within these disciplines that expertise is developed and validated). But as a result interdisciplinary research (and, although less decisively, multidisciplinary teaching) may be placed at a disadvantage. This has been a particular concern in Britain (and an attempt has been made to address this concern in the forthcoming RAE by establishing a superior tier of over-arching panels to oversee the work of the more detailed expert panels). Of course, any bias against interdisciplinary research is likely to discourage the development of more open and fluid patterns of knowledge production. Second, the experts who dominate peer review systems, again inevitably, tend to be those who have been successful within the context of existing intellectual paradigms and institutional structures. Potentially, therefore, there may be a bias against novelty.

### *‘Metrics’*

The second broad type of methods for assessing excellence in the social sciences and humanities is so-called ‘metrics’. In recent years there seems to have been a drift from peer review, which appears to be too personal and subjective, to ‘metrics’, which

(superficially, at any rate) seem to be more impersonal and objective. For example, earlier this year in Britain the Government announced that future RAEs (i.e. after the current RAE in 2008) would be based on ‘metrics’ rather than peer review. However, there is a wide range of ‘metrics’ to choose from, although two general characteristics can be identified:

- First, many (probably the majority) of ‘metrics’ simply mirror the outcomes of peer review systems, although others reflect different criteria. For example, if success in securing research funding is a ‘metric’, it is better described as an indirect form of peer review (because funding bids are subject to a peer-review process). Similar considerations apply to range of other familiar ‘metrics’ such as the number of PhD students or of articles in refereed journals (or even citation indices);
- Second, many ‘metrics’ are based on measuring inputs rather than outputs. For example, however stiff the competition to secure research grants, they merely provide the means for undertaking research; they do not guarantee (and, therefore, cannot measure) successful research outcomes). Some familiar ‘metrics’ are – apparently – based on output measures – for example, the production of PhDs (although it can be argued that the production of PhDs is better defined as an input into the research system).

Many of the critiques of peer review also apply to traditional ‘metrics’. In addition as peer-review proxies they may be less accurate, because ‘metrics’ incorporate other criteria. For example, the impact of political objectives or market considerations on research programmes is likely to influence the success, or otherwise, of researchers in securing research funding. Another factor is that success rates in securing funding are also influenced by institutional affiliations with applicants from established research groups (and more prestigious institutions) being favoured – and the influence of patronage can never be entirely eradicated.

To some extent this loss of accuracy in ‘indirect’ peer review may be compensated for by greater efficiency – in other words, a reduction in the administrative burden of peer review and (possibly) in the inconsistency of outcomes from different peer review systems, although it can be argued that a multiplicity of peer review systems is necessary because they operate in different contexts and have different objectives. For example, reviewers of research grant proposals are asked to assess their potential and practicability, although they may also be asked to comment on the track-record of the proposers; they cannot be asked to assess the quality of the outputs which will be produced when the research is completed.

This highlights the difficulties with the second characteristic of traditional ‘metrics’ systems, their emphasis on input at the expense of output factors. The emphasis on inputs may tend to make any research assessment system even less dynamic. Peer review, because it focuses on outputs, makes it easier for new players to enter the game if they produce high-quality research. The comparative absence of inputs cannot invalidate the quality of these outputs; indeed, arguably, it may reinforce the achievement (as well as representing better value-added and / or value-for-money).

On the other hand, there is some correspondence between the quantity of inputs and the quality of outputs. This is more likely to be the case in the empirical – and especially the quantitative – social sciences than in the humanities where the correlation is often weak (or, more accurately, the inputs are less direct – because they are embodied in professional careers and institutional structures). But, even in the case of the more resource-intensive social sciences, new modelling and simulation techniques and the accessibility of data-sets mean that high-quality research outputs can be produced with limited inputs.

### *Performance indicators*

The third main type of method for assessing excellence in the social sciences and humanities is the development of a range of performance indicators. In one sense this is a natural extension of the trend towards relying on ‘metrics’. Performance indicators are an attempt to remedy the two main disadvantages of ‘metrics’ – their dependence on measures that are derived from – or, to put it more bluntly, are parasitic on – peer-review processes (in other words, peer-review proxies or ‘indirect’ peer review); and their emphasis on input rather than output factors. In another sense it is an attempt to substitute ‘external’ criteria, whether politically determined or market driven, for ‘internal, scientific, criteria for assessing research quality (and, more crucially, shaping research priorities) – or, at the very least, to rebalance criteria in favour of the former. The social sciences, in particular the more applied social sciences, are more likely to be affected by the development of performance indicators than the humanities – although the latter are unlikely to remain fully exempt for ever because of the increasing emphasis on their relevance to the creative and cultural industries.

The trend towards performance indicators is both a political and a scientific phenomenon:

- i) In political terms it is one element within wider efforts to develop more integrated national research, development and innovation systems. These efforts reflect the importance of such systems in knowledge-based economies. As a result the key ministries and other public agencies have tended to be in the areas of trade, industry, productivity rather than education and science. Because the scope of these systems extends far beyond even the most expansive definitions of academic, or scientific, research the alleged inadequacies of peer review of the predominant means for assessing excellence have been highlighted. Consequently there have been efforts to move ‘beyond peer review’ – by attempting to address these alleged inadequacies, by treating peer review as only one of a range of assessment tools or by relying instead on new kinds of tools;
- ii) In scientific terms the trend towards a more diverse set of performance indicators reflects the development of more open knowledge production systems which has been discussed earlier in this paper. Peer review is an effective tool for assessing the quality of research outputs, despite its inherently reductionist and conservative biases (because it is grounded in discipline-specific expertise). However, it has less effective for assessing the quality of research processes (which is where ‘metrics’) may be more useful;

and it is certainly less effective for assessing the impact of research (despite efforts to give greater weight to successful dissemination). In the more fluid world of ‘Mode 2’ knowledge production (and in the historically determined and socially constructed world of scholarship in the humanities, which has always existed?) divisions of labour are tending to break down, which makes it more difficult to devise single-purpose and uncomplicated assessment regimes.

The measurement of the ‘impact’ of research and scholarship, a preoccupation of many Governments as they recognise the challenges of globalisation, provides a good case-study. Within the university and research systems ‘impact’ has conventionally been measured in terms of publications in so-called high-impact journals or of citation indices. But this rather limited, and reputation-focused, definition of ‘impact’ is now being supplemented (or even replaced) by a much wider, and more application-oriented, definition – although measuring it remains problematical, partly because there is a wider (and more contested) range of possible indicators but partly because notions of ‘excellence’ have themselves become more multi-dimensional (or even incommensurable). Yet performance indicators, which are more flexible, adaptable and transparent, may have the potential to capture these complexities more effectively than unadorned peer review or mechanistic ‘metrics’. Or, more accurately, all three methods – peer review, ‘metrics’ and performance indicators – are now needed to assess the quality of research (or of any other academic process, such as teaching).

## **Conclusion**

The aim of this paper has been to reflect on how the assessment of excellence in the social sciences and humanities can be conceived of within the context of the modern university, of the evolution of these disciplines and of the wider evolution of knowledge production systems (primarily research but also teaching and other forms of dissemination to and interaction with society). It attempts to sketch the ‘big picture’ at the level of general concepts; consequently it has ignored much of the crucial detail (that all-important deep contextualisation which the author, together with colleagues, has discussed in other places).

In summary the key points in this paper are:

- In the social sciences and humanities, as in many other disciplines, there is an emerging tension between methodologies for assessing excellence, which are tending to become more urgent and more prescriptive, and concepts of excellence, which are becoming more open, more fluid and more contested. These more prescriptive methodologies can be interpreted as a reaction to these more open definitions of excellence, but they also reflect deeper changes in political discourse and the political system;
- Both the social sciences and humanities are dynamic disciplines – subject both to far-reaching ‘internal’ changes in terms of their intellectual agendas and research methodologies but also ‘external’ changes which impact on their development (for example, the massification of higher education and the shifting logistics of research). Because of their ethos and history the social sciences and humanities may be better able to manage the tension between

more urgent and prescriptive methodologies for assessing excellence (the ‘audit culture’) and more open and fluid definitions of excellence than some other disciplines;

- It is important to ‘unpack excellence’. Its four major elements are research and scholarly outputs (but in the humanities especially too great an emphasis on products as opposed to process may not be helpful); methodologies (of increasing relevance in a ‘Mode 2’ knowledge production environment in which methodologies have proliferated and become more heterogeneous – but also more central to the research process); sustainability (whether in terms of capacity building or of sustaining scholarly cultures in a broader sense); and relevance and dissemination (the interface with society – so crucial in the social sciences);
- Three main types of methods are available for assessing excellence in the social sciences and humanities – peer review (still, perhaps, like democracy a bad system but better than any of the others!); ‘metrics’ (which may be more appropriate in an ‘audit society’ because, apparently, they are easier to quantify – but which often build on, and aggregate, antecedent peer review judgements); and performance indicators (which by potentially introducing other – less academic? – assessment criteria may appear more intrusive and threatening – but which may also be more comprehensive, more transparent and more adaptable).

On the basis of these conclusions, if generally accepted, it would be possible to discuss the merits of more detailed mechanisms for assessing excellence in the social sciences and humanities. That is, perhaps, the next task.