

Essay

Clashing Concepts and Methods: Assessing Excellence in the Humanities and Social Sciences

Peter Scott

UCL Institute of Education, 20 Bedford Way, London WC1H 0AL, UK; E-Mail: p.scott@ioe.ac.uk;
Tel.: +44-20-7612-6000

Academic Editor: Cinzia Ferrini

Received: 4 November 2014 / Accepted: 3 December 2014 / Published: 4 February 2015

Abstract: The humanities are going through a period of exceptional vitality characterised by the proliferation of novel interpretative frameworks, methodologies and perspectives. Yet they—and, to a lesser extent, the social sciences—feel threatened by the rising tide of research assessment which appear to be predominantly derived from the needs and experience of the physical and natural sciences and depend on the application of standardized assessment tools. This position paper intends to contribute to the debate on the current criteria of evaluation and measures for excellence in the Humanities by casting light on their conceptual implications and methodological assumptions. It argues that decisions about their relative weight are not simply technical but also reflect underlying value systems.

Keywords: research assessment; peer review; metrics; excellence

1. Introduction

The humanities and social sciences have “never had it so good”—and also never been under such pressure. They have “never had it so good” because their creative vitality is almost unparalleled. Arguably this is one of their greatest intellectual ages; new interpretative frameworks have emerged and more rigorous modes of enquiry have been developed; disciplines have cross-fertilised each other in remarkable ways; and new social, economic and cultural connections have been made as a result of the growth of the so-called creative and “industries”.

Yet the humanities and social sciences have never been under such pressure—for number of reasons. The first is that these new, but more instrumental, connections are now regarded as providing

their primary justification [1,2]. A second reason is that new research practices are emerging that emphasise the identification overarching “themes” and the need to construct research teams (or other forms of collaboration) to address them, which may not be easy to reconcile with the more individualistic scholarly practices characteristic at any rate of the traditional humanities. A third reason is that new political vocabularies of “world-class” and “excellence” have emerged that are inevitably more difficult to define in the normatively contested domains of the humanities and (many) social sciences than in the natural, physical and mathematical sciences (and their dependent technologies), which are more empirically and experimentally grounded. A fourth reason is that universities, like so many institutions, are now subject to a “cult of measurement” and methods of review, assessment and evaluation characteristic of what has been called the “audit society” [3].

This dissonance between the rude health of the humanities and (most) social sciences on the one side, and the sense of growing pressure on these disciplines on the other side, helps to explain the proliferation of an extensive literature of “threat” and even “crisis”: calls to action to defend the humanities against the philistines (of many types). The humanities’ uneasy state of mind is revealed by these often eloquent rallying cries [4].

2. Measuring “Excellence”: The Conceptual Challenge

At a conceptual level the possibility of defining “excellence” is becoming more problematical for two main reasons. 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.

The former process is familiar enough across the academic spectrum. 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 fields, sometimes more interdisciplinary, 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 latter process, the impact of the external environment, is also (apparently) straightforward. As is well understood, 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).

In the past it appeared to be possible to distinguish between 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). But, in reality, 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 that, in turn, had often been helped to emerge as a result of earlier research.

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 [5]:

- One is the multiplication of sites where serious research is undertaken—both in a spatial and qualitative sense. In a spatial 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” or consulting firms).
- A second characteristic is that the roles of the various research actors have also 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). As a result, the boundaries between the scientific system and other systems: social, political and economic, have become highly porous.
- A third characteristic is that the cognitive core, both epistemological and methodological, of many disciplines has been invaded—in two ways. First, 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. Second, “alternative” knowledge traditions have invaded, and even been incorporated in, the cognitive core.
- 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. “Excellence” needs to be unpacked under these new conditions.

3. Measuring “Excellence”: The Political and Organisational Imperative

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 appear to be several reasons for this growing pressure.

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 20th-century welfare state into the 21st-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 [6]. 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 “intermediate” domain 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.

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 that they are securing value for money from their direct or indirect investment in “knowledge” institutions [7]. 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” that 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 battleground, an arena in which nations (and regions and cities) struggle for competitive advantage—which has also stimulated the demand for “excellence”.

A third reason, of course, is the desire to combat the complexities 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 [8]. 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. The emerging discourse of “world-class universities” is evidence of this drive towards “excellence” [9].

4. Unpacking Excellence

These changes in the wider social, and higher education and research, systems 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.

Four major elements in the current definition of excellence in research—outputs; methodologies; sustainability; and relevance or applicability—will now be examined.

4.1. 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”). The “how”, methods for assessing quality, will be discussed in the next section of this essay. In this section the focus will be 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 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).

However, in the more open knowledge production systems that now prevail, new difficulties have emerged. One, discussed in the next section of this essay, is that academic communities have become more diffuse: this has tended to increase the contestability of their authority. But a second difficulty is that 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. The proliferation of policy and evaluation studies in some social sciences has also 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.

Two issues arise as a result. The first is the stretching of appropriate forms of academic output: that is, a change in 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. Indeed, 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 would be interesting to reflect on the extent to which current definitions of research “outputs” are contingent in particular technologies of scholarly communication). The second issue is to what extent it is possible to define a research process as being in itself an “output”: must that label be reserved for products in a more conventional sense? In turn, this last issue raises two

questions. First, when do pre-prints of research findings, often circulated with the deliberate intention of provoking critique, become fixed “outputs”? The dynamics of research production now make it possible for findings to be left “open” for much longer. Second, while it may have been possible in the past to agree that research findings were sufficiently “complete” to be presented in a “final” form (in other words the “process” had been translated into a “product”), has the acceleration of research, combined with the indeterminacy of knowledge (and new communication technologies), fundamentally undermined these processes of “closure”?

4.2. Methodologies

The second element of excellence is methodologies. In the past this seemed—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.

However, important changes may have made it more difficult to agree a stable set of definitions of “excellence” in the social sciences and humanities. From my perspective, I would emphasise two major changes:

- (1) There has been a proliferation of research practices. Not only have methodologies become more diverse, they have also become less commensurable. A growing number of smaller-scale qualitative studies in the social sciences are often undertaken by researchers who are themselves practitioners. This active engagement has been considered a positive asset by those who espouse so-called “action research”. But this proliferation of research actors may also have compromised not only their objectivity (although 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).
- (2) There has been a similar proliferation of interpretative frameworks. In the past many of these frameworks, including sharply contested and rival “theories”, 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.

In my view, these changes have made it more difficult to arrive at satisfactory definitions of “excellence” in terms of better-or-worse methodologies.

4.3. Sustainability

The third element of excellence is sustainability. Even in the humanities where individual scholarship is still a 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. 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. One is clearly financial. There must be sufficient funding to pay the full economic costs of research and scholarship. A 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. A third dimension is concerned with people [10]. 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 that may be more difficult to sustain 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. A 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; and also within the wider political system which recognise their value to society and the economy.

However, changes taking place in higher education and research, and in society at large, challenge some current assumptions about the right conditions for guaranteeing excellence in terms of sustainability.

- (1) 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).
- (2) 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.
- (3) 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). However, it can also be argued this trend may make it easier for new players to enter the research arena if its effect is to lower the (professional) barriers to participation.

4.4. Relevance

The fourth element of excellence is relevance. 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. I mean 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.

As it is well known, in the humanities relevance has always been difficult to determine. 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, the potential relevance of research and scholarship in the humanities was more difficult to demonstrate (at any rate before the rise of the creative and cultural industry). Third, social influences on research and scholarly agendas were less direct. They were 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.

More recently a range of new ideas has been developed which tend to place greater emphasis on the contextualisation of knowledge, not simply as 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 following assertions [11]:

- 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 also within a context of implication (by imagining the as-yet unimaginable).

The cumulative effect of such ideas is both to emphasise the contribution of relevance to assessing excellence in social sciences and humanities research but also perhaps to cast its influence in a much more positive light.

In summary outputs, methodologies, sustainability and relevance remain the major elements in assessing excellence. But they have all be modified to some degree by wider changes in the higher education and research systems and by the intellectual and logistical shifts within social sciences and humanities research. The combined impact of these “external” changes and “internal” shifts has been to problematise still further the notion of excellence.

5. Measuring Excellence: Shifting Methods

The choice between different methods of assessment is not simply a pragmatic one—what works best. The assessment of excellence, like all measurement tools, cannot be a neutral process; there are competing “excellences” (to say this is not to espouse crude relativism). Decisions about their relative weight are not simply technical but also reflect underlying value systems.

For example, it is possible to define excellence in the social sciences and humanities as a process—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).

5.1. 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 commonly used to decide which scholarly monographs and journal articles should be published [12]. It is also used to determine which scholars and researchers should receive research grants. Finally peer review is also used in some national research assessment systems. For example, in the UK in the regular five or six-year assessments of the quality

of research at the level of university departments undertaken since the mid-1980s (originally the Research Assessment Exercise (RAE) now the Research Excellence Framework (REF)), the heart of the process is the grading of individual research outputs and their “impact” by expert panels.

However, there are four main critiques of peer review. 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 and also have a legitimate stake the quality of these programmes. Clearly these conditions cannot always be met. The second critique 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. 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). As a result, interdisciplinary research may be placed as a disadvantage. 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.

5.2. “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 [13]. To some degrees this drift has been encouraged by policy makers, who are seeking more cost-effective forms of measurement, and institutional managers, who welcome the potential “metrics” databases as instruments for managing research performance. But, this drift has also been encouraged by the academic community itself as search engines, such as the “Web of Science”, Google Scholar and “ResearchGate”, have become more popular.

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 datasets mean that high-quality research outputs can be produced with limited inputs.

5.3. Performance Indicators

The third main type of method for assessing excellence in the social sciences and humanities is the development of a wider 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 depend on—peer-review; and their emphasis on input rather than output factors. However, in another sense it is an attempt to substitute “external” criteria, whether politically determined or market driven, for “internal”, or scientific and scholarly, criteria for assessing research quality (and, more crucially, shaping research priorities).

The trend towards performance indicators can be regarded as both a political and a scientific phenomenon. 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 have tended to be in the areas of trade, industry, productivity rather than education and science. In the context of these wider, and more instrumental, agendas the inadequacies of peer review 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. In scientific terms the trend towards a more diverse set of performance indicators reflects the development of more open knowledge production systems. 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 also less effective for assessing the impact of research (despite efforts to give greater weight to successful dissemination).

It is possible to argue that as a result of the development of more open and fluid knowledge production systems (but also in the historically determined and socially constructed world of scholarship in the humanities), it is now becoming 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, the “impact” factor 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 diverse notions of “excellence” have themselves become more multi-dimensional (or even incommensurable).

It is also possible to argue that 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, in my view all three methods—peer review, “metrics” and performance indicators—are now needed to refine the assessment of the quality of research.

6. Conclusions

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 and more open and fluid definitions of excellence than some other disciplines.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Greg Hearn, Stuart Cunningham, and Diego Ordonez. “Commercialisation of Knowledge in Universities: The Case of the Creative Industries.” *Prometheus: Critical Studies in Innovation* 22 (2004): 189–200.
2. Cinzia Colapinto, and Colin Polezza. “Innovation in Creative Industries: From the Quadruple Helix to Systems Theory.” *Journal of Knowledge Economy* 9 (2012): 343–53.
3. Michael Power. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press, 1999.
4. Martha Nussbaum. *Not for Profit: Why Democracy Needs the Humanities*. Princeton: Princeton University Press, 2010.
5. Michael Gibbons, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage, 1994.
6. Giandomenico Majore. “From the Positive to the Regulatory State: Causes and Consequences of Changes in Modes of Governance.” *Journal of Public Policy* 17 (1997): 139–67.
7. Max Olsen, and Michael Peters. “Neo-Liberalism, Higher Education and the Knowledge Economy.” *Journal of Education Policy* 20 (2005): 313–45.
8. Peter Scott. *The Meanings of Mass Higher Education*. Buckingham: Open University Press, 1995.
9. Jamil Salmi. *The Challenge of Establishing World-Class Universities*. Washington: World Bank, 2009.
10. Ulrich Teichler. *The Formative Years of Scholars*. London: Portland Press, for the Academia Europaea—Wenner-Gren International Series, 2006.
11. Helga Nowotny, Peter Scott, and Michael Gibbons. *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge: Polity Press, 2001.
12. Ann Weller. *Editorial Peer Review: Its Strengths and Weaknesses*. Washington: American Association for Information Science and Technology, Monograph Series, 2001.
13. Claire Donovan. “Introduction: Future Pathways for Science Policy and Research Assessment: Metrics v. Peer Review, Quality v. Impact.” *Science and Public Policy* 34 (2007): 538–42.

© 2015 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).