DRUID society

DRUID Working Paper No. 11-05

How Journal Rankings can suppress Interdisciplinary Research - A Comparison between Innovation Studies and Business & Management

By

Ismael Rafols, Loet Leydesdorff, Alice O'Hare, Paul Nightingale and Andy Stirling

Danish Research Unit for Industrial Dynamics

www.druid.dk

How Journal Rankings can suppress Interdisciplinary Research – A Comparison between Innovation Studies and Business & Management

Ismael Rafols
SPRU- Science and Technology Policy Research
University of Sussex, Brighton, BN1 9QE, England
And
Technology Policy and Assessment Center
Georgia Institute of Technology
Atlanta, GA 30331, USA
E-mail: i.rafols@sussex.ac.uk (corresponding author)

Loet Leydesdorff
University of Amsterdam
Amsterdam School of Communication Research
1012 CX Amsterdam, The Netherlands
E-mail: loet@leydesdorff.net

Alice O'Hare, Paul Nightingale and Andy Stirling
SPRU- Science and Technology Policy Research
University of Sussex, Brighton, BN1 9QE, England

This version: 29 September 2011

Submitted to Research Policy

Updates will be available at
www.interdisciplinaryscience.net/ or
http://www.sussex.ac.uk/Users/ir28/IDR/Rafols2011-Rankings&IDR.pdf

Abstract:
This study provides new quantitative evidence on how journal rankings can disadvantage interdisciplinary research during research evaluations. Using publication data, it compares the degree of interdisciplinarity and the research performance of innovation studies units with business and management schools in the UK. Using various mappings and metrics, this study shows that: (i) innovation studies units are consistently more interdisciplinary than business and management schools; (ii) the top journals in the Association of Business Schools' rankings span a less diverse set of disciplines than lower ranked journals; (iii) this pattern results in a more favourable performance assessment of the business and management schools, which are more disciplinary-focused. Lastly, it demonstrates how a citation-based analysis challenges the ranking-based assessment. In summary, the investigation illustrates how ostensibly 'excellence-based' journal rankings have a systematic bias in favour of mono-disciplinary research. The paper concludes with a discussion of implications of these phenomena, in particular how resulting bias is likely to affect negatively the evaluation and associated financial resourcing of interdisciplinary organisations, and may encourage researchers to be more compliant with disciplinary authority.

Keywords: Interdisciplinary, Evaluation, Ranking, Innovation, Bibliometrics, REF
Jel codes: A12 ; O30

ISBN 978- 87-7873-316-0

This article received the William Page Best Paper Award in the 2011 *Atlanta Science and Innovation Policy Conference.*

1. Introduction

At a time when science is under pressure to be relevant to society (Nightingale and Scott, 2007; Hessels, 2010), interdisciplinary research (IDR) is often praised for contributing towards scientific breakthroughs (Hollingsworth and Hollingsworth, 2000), addressing societal problems (Lowe and Phillipson, 2006) and fostering innovation (Gibbons et al., 1994). Reasons given for supporting interdisciplinarity include suggestions that IDR is better at problem-solving (Page, 2007, p. 16), that it generates new research avenues by contesting established beliefs (Barry et al., 2008) and that it is a source of creativity (Heinze et al., 2009; Hemlin et al., 2004). This is argued to help rejuvenate science and contribute towards its ‘health’ (Jacobs and Frickel, 2009, p. 48).

However, IDR is also widely perceived as being at a disadvantage in research evaluations (Rinia et al., 2001a, p. 357; Nightingale and Scott, 2007, pp. 546-547). Various qualitative studies have provided evidence that peer review tends to be biased against IDR (Laudel and Origgi, 2006; Langfeldt, 2006, p. 31). But only a small number of quantitative investigations have been undertaken on this claim and they have been inconclusive (Porter and Rossini, 1985, p. 37; Rinia et al., 2001a).

Here we explore potential biases in the evaluation of IDR in the field of Innovation Studies (IS) in the UK. IS is a diverse and ambiguously-bounded area of the social sciences that studies the causes, processes and consequences of innovation (Fagerberg et al., this issue). Given its problem-oriented and interdisciplinary nature, IS research is conducted in diverse types of research units that experience a variety of institutional challenges (Clausen et al., this issue), in particular lack of fit with assessment panels.

The UK is a particularly suitable setting for this enquiry, as it has an important IS community, a comparatively homogenous higher education system, and a long history of research assessment (Collini, 2008). The UK has also repeatedly witnessed particular concerns about biases against IDR – for instance since the Boden Report (ABRC, 1990). Under current funding conditions in the UK, many IS units have been (at least partly) incorporated into Business and Management

Schools (BMS) (e.g. in Oxford, Imperial, Manchester, Cardiff and recently Sussex). BMS face acute pressures to achieve high performance in publication rankings, both for reputational purposes and because of the financial incentives associated with the research assessment procedures of the national funding council HEFCE.[1] This assessment exercise is currently referred to as the 'Research Excellence Framework' (REF) (Martin and Whitley, 2010, p. 61). UK BMS are also subject to a particularly strictly-conceived formal ranking scheme for disciplinary journals, provided by the British Association of Business Schools (ABS) (ABS, 2010).

The use of journal rankings (like those of the ABS) in research evaluations has become increasingly popular. It is seen as a means to 'objectify' research assessment and thus avoid or compensate for biases in peer review (Taylor, 2011). Yet journal-based evaluation has been severely criticised as seriously inappropriate for this role (Seglen, 1997; Oswald, 2007). Despite this, proliferation of journal ranking schemes indicate increasingly wide usage across many disciplines (both explicitly and implicitly) for a variety of quality assessment purposes, such as resourcing, recruitment, and promotion. A range of different studies have demonstrated that the journal ranks of a department's publications are by far the strongest predictor of the results obtained in the 2008 UK's Research Assessment Exercise (RAE), although journals rankings are not formally used in the evaluation (Kelly et al., 2009; Taylor, 2011, pp. 212-214). As a result university managers are making increasingly explicit use of such rankings to prepare future assessments.

In this study, three centres for IS in the UK are compared to the three leading British BMS. The choice of BMS as comparators is given by the fact that many IS centres are closely associated with BMS and, in important respects, assessed by the Business and Management panel in the REF.[2] We investigate quantitatively the relationship between the degree of interdisciplinarity and perceived performance, as shown by the ABS journal rankings. We then compare the results with more reliable citation-based indicators. In summary, the results show that ABS journal rankings favour approaches within the dominant disciplines of BMS (mainly business,

[1] Higher Education Funding Council for England.
[2] Some IS units were previously evaluated under the *Politics and International Relations* panel, indicating the difficulty of fitting them into established categories.

management, economics and finance) —and thus disadvantage the interdisciplinary IS units. Where ratings and journals ranks are highly correlated (in exercises such as previous RAEs) (Taylor, 2011), the effect we find is large enough to be expected to have a substantial negative effect on research assessment performance. To our knowledge this study is the first to demonstrate a bias against IDR on a firm quantitative basis.

The policy implications of these results will be discussed in the light of studies on the consequences of biases in assessments and journals. For example, research suggests that British economics departments have narrowed their recruitment to favour 'main-stream' economists (Harley and Lee, 1997; Lee and Harley, 1998; Lee, 2007), thus reducing the cognitive diversity of the science system's ecology. This may have lead to impoverishment in the mid or long term (Molas-Gallart and Salter, 2002; Stirling, 1998, pp. 6-36; Stirling, 2007; Martin and Whitley, 2010, pp. 64-67).

In addition to the main focus of this paper on the bias against IDR in assessment, this article also aims to make some more general contributions to advance the state of the art in methods for using bibliometric indicators in policy. First, it provides an introduction to a methodology of concepts, mathematical operationalisations and visualisations for the study of interdisciplinarity using bibliometric data. Second, it highlights that conventional measures of performance for IDR publications remain problematic, and suggests citing-side normalisation as an improved alternative. Third, it illustrates the use of multiple indicators for the study of multidimensional concepts such as interdisciplinarity or performance. In this, we follow Martin and Irvine's (1983) seminal argument that, since no simple measures exist that can fully capture the contributions made by scientists, one should use various partial indicators. Though incomplete (as well as imperfect and subject to contingency and distortion), this more 'plural and conditional' (Stirling, 2010) form of bibliometric analysis may be considered more reliable when diverse indicators converge to yield the same insights. Since plurality is more easily captured by multidimensional representations, we illustrate this point with the full set of maps (available at http://interdisciplinaryscience.net/maps/).

For the sake of brevity and focus, a number of otherwise relevant issues related to the subject raised will not be dwelt in this article. The present study does not offer any kind of assessment of the individual organisations examined – this would entail a broader evaluation than the exclusive focus on publication 'quality' used here. Second, it does not discuss the relative benefits of IDR. We note simply that IDR is highly valued by many researchers and policy-makers – which is sufficient to render important the question of whether IDR is fairly assessed. Third, we do not look into the societal impact of research. The concern here is whether there is a bias against IDR only when considering conservative, internal measures of scientific merit. Finally, we do not elaborate the details of conceptualisations and operationalisations of interdisciplinarity and performance. Instead we build on conventional indicators of performance and on published research on IDR by some of the authors to which we will direct the reader at relevant points.

2. The evaluation of interdisciplinarity research

Variously conceived notions of interdisciplinarity have become a highly valued characteristic in science policy and management (Metzger and Zare, 1999). IDR is seen as a way of sparkling creativity, supporting innovation and addressing pressing social needs (Jacobs and Frickel, 2009, p. 48). This is well-illustrated by a variety of high profile initiatives, such as the UK's Rural Economy and Land Use Programme (RELU, http://www.relu.ac.uk/, Lowe and Phillipson, 2006), the US' Integrative Graduate Education and Research Traineeship (IGERT, http://www.igert.org/, Rhoten et al., 2009), the explicit call to disciplinary transgression in the prestigious grants of the new European Research Council (ERC, 2010, p.12), or the establishment of new cross-disciplinary institutes such as the Janelia Farm of the Howard Hughes Medical Institute or the Bio-X centre at Stanford University (Cech and Rubin, 2004). This position goes hand in hand with significant increases in articles claiming to be interdisciplinary (Braun and Schubert, 2003) and with more interdisciplinary citing patterns (Porter and Rafols, 2009).

However, in parallel with this wave of declared support, IDR is, in practice, often found wanting: accused of being too risk averse; lacking disciplinary notions of quality; or not meeting policy expectations (Bruce et al., 2004, pp. 468-469). Claims over the benefits of interdisciplinarity are questioned (Jacobs and Frickel, 2009, p. 60), since they are based on limited evidence relying often on 'cherry-picked' case studies of success, that are often selected, narrated or analysed ex-post (e.g. Hollingsworth and Hollingsworth, 2000 or Heinze et al., 2009). The effects of IDR on research outcomes are difficult to prove systematically because interdisciplinarity is just one mediating aspect of the many factors that contribute to the success or relevance of research. As a result, subtle contextual differences can lead to disparate results. For example, whereas some studies have correlated IDR practices with the intensity of university-policy or university-industry interactions (Van Rijnsoever and Hessels, in press; Carayol and Thi, 2005), other studies do not find IDR to influence the success of firms founded by academic teams (Muller, 2009).

Yet, irrespective of perspective, there is agreement that IDR faces important barriers which may significantly hinder its potential contributions (Rhoten and Parker, 2006; Llerena and Mayer-Krahmer, 2004). In the first place there are difficulties in managing the coordination and integration of distributed knowledge. This is addressed by research examining various kinds of team work and collaboration (Cumming and Kiesler, 2005, 2007; Katz and Martin, 1997; Rafols, 2007).[3]

Second, there are more systemic barriers due to the institutionalisation of science along disciplinary structures (Campbell, 1969; Lowe and Phillipson, 2009). Perceived barriers include, for example: the relatively poor career prospects often experienced by interdisciplinary researchers; lower esteem from colleagues; discrimination by reviewers in proposals; and disproportionate difficulty in publishing in prestigious journals (Bruce et al., 2004, p. 464). The US National Academies (2004) *Facilitating Interdisciplinary Research* report provides a thorough review of these barriers, and suggests initiatives to lower them. Since these barriers tend to be embedded and thus implicitly 'naturalised' in institutional practices, they are generally less visible and more controversial than teamwork problems. While such hurdles to IDR are often acknowledged in policy, their mechanisms are neither well documented nor deeply

[3] An important research agenda focusing on these problems in recent years, is the so-called 'Science of Team Science' community (or SciTS, pronounced 'sahyts'), which has developed 'an amalgam of conceptual and methodological strategies aimed at understanding and enhancing the outcomes of large-scale collaborative research and training' (Stokols et al., 2008. p. S77; Börner et al., 2010).

understood (EURAB, 2004; Metzger and Zare, 1999; National Academies, 2004; Rhoten and Parker, 2006).

A widely-perceived ‘key barrier’ is the undue bias against IDR in evaluation (Rinia et al., 2001a, p. 357; Lee, 2006; Nightingale and Scott, 2007, pp. 546-547). For example, Boddington and Coe (1999, p.14) reported from a large survey (5,505 respondents) that 51% of researchers, 68% of department heads and 48% of panel members, viewed the 1996 UK’s RAE as slightly or strongly inhibiting IDR (against 24%, 15% and 19%, respectively, who viewed RAE as promoting IDR). Investigations on peer-reviewed based research evaluation support these perceptions (see e.g. a monographic issue of the journal *Research Evaluation* edited by Laudel and Origgi, 2006). In summary:

> ‘...a re-emerging awareness of interdisciplinarity as a vital form of knowledge production is accompanied by an increasing unease about what is often viewed as the ‘dubious quality’ of interdisciplinary work. Central to the controversy is the lingering challenge of assessing interdisciplinary work.’ (Boix Mansilla, 2006, p.17)

That evaluation of IDR is problematic is not a surprise. Any evaluation needs to take place using established standards. These standards can be defined within a discipline, but what standards should be used for research in between or beyond disciplinary practices? If IDR must be equally compliant with the (sometimes radically) contrasting quality criteria of more than one discipline, then it self-evidently faces an additional hurdle, compared to single discipline based research evaluated against a single set of criteria. Beyond this, peer-review has been shown to display inherently conservative and risk minimising tendencies, which ‘may disfavour unconventional and interdisciplinary research’ (Langfeldt, 2006, p. 31) and therefore favour well established fields over nascent ones (Porter and Rossini, 1985, p. 37). Of course, programmes targeting ‘high risk, high reward research’, where IDR is explicitly sought, can be an exception to this disadvantage (Balakrishnan et al., 2011). But what appears to happen generally, even in the case of multidisciplinary review panels, is that IDR ends up being assessed under the single apparently most relevant disciplinary perspective (Mallard et al., 2009, p. 22) or under the evaluator’s own favoured disciplinary criteria. This phenomenon has been dubbed ‘cognitive cronyism or particularism' (Travis and Collins, 1991).[4] As Laudel and Origgi (2006, p. 2) note:

> 'in spite of the political narratives on the need for interdisciplinarity, the criterion of quality can be turned into an instrument for suppressing interdisciplinary research because the established [disciplinary] quality standards are likely to prevail.'

Surprisingly, the clear insights and strong general impressions from qualitative studies that IDR is at a disadvantage in peer review have not been robustly substantiated by quantitative studies. Examining a sample of 257 reviews of 38 projects from 5 somewhat IDR programmes (e.g. neurobiology) of the US National Science Foundation, Porter and Rossini (1985) found a (weak, but significant) correlation between low grading and degree of interdisciplinarity ($r=0.29$, $p<0.05$). Rinia et al. (2001a) analysed the national evaluation by an international panel of 185 physics programmes in Dutch universities, but they didn't find a bias against IDR. However, they reported that IDR tends to be published in journals with a lower citation impact (Rinia et al., 2001a, p. 360; 2001b, p. 247).

In the survey study commissioned by the UK's Higher Education Funding Bodies on the RAE already mentioned, Boddington and Coe (1999, p.iii) arrived at the conclusion that 'there is no evidence that the RAE systematically discriminated against interdisciplinary research in 1996'. Interestingly, though, a closer look at their data shows that the highest RAE scores were obtained by researchers that self-reported being at the lower part of IDR spectrum (dedicating between 30-40% of their time in IDR activities), whereas the researchers reporting high IDR-involvement obtained lower scores. The effect is particularly strong for teaching-based researchers.[5] Other bibliometric studies have found that articles with an intermediate degree of interdisciplinarity are more likely to be cited than either the mono-disciplinary or the extremely interdisciplinary ones (e.g. Adams et al., 2007; Larivière and Gingras, 2010; Yegros-Yegros et al., 2010), and that in the natural sciences the

[4] 'Cognitive particularism' should not be confused with the better documented phenomenon of institutional capture by 'old boys' networks.

[5] Although the study does not report standard error or statistical significance of the results, one might assume that the results are statistically significant since they are based on a sample of 5,505 respondents.

citations per paper received by multidisciplinary journals[6] are lower than in mono-disciplinary ones (Levitt & Thelwall, 2008). However, since these studies did not make comparisons between bibliometric data and research evaluation ratings, potential biases could not be assessed.

In summary, in contrast to the many qualitative studies pointing to clear bias against IDR in evaluation (Travis and Collins, 1991; Langfeldt, 2006), the few quantitative studies on the subject have produced ambiguous and contradictory results. This overall inconclusiveness in quantitative evidence has been interpreted by some as evidence of absence of bias (Huutoniemi, 2010, p. 318). In this study we aim to help fill this gap by investigating a potential bias against IDR that would result from the use of journal rankings in research evaluation.

3. Methods and underlying conceptualisations

3.1 Methodological framework: converging partial indicators

Assessments of scientific performance and interdisciplinarity remain controversial and lack a consensus on appropriate frameworks and methodologies, even when solely based on narrow quantitative measures such as publication outputs (Bordons et al., 2004; Huutoniemi et al., 2010). This should come as no surprise, given that both performance and interdisciplinarity are multidimensional concepts, which can only be partially captured by any single indicator (Martin and Irvine, 1983; Narin and Hamilton, 1996; Sanz-Menéndez et al., 2001).

Unfortunately, as scientometrics became more widely used and institutionalised in policy and management, flaws (and associated caveats; e.g. Leydesdorff, 2008) in the use of bibliometric tools have become increasingly overlooked. Martin (1996) reminded and warned the research policy community about the lack of robustness of uni-dimensional measurement of multi-dimensional concepts such interdisciplinarity

[6] Multidisciplinary journals are defined by Levitt and Thelwall as journals that are classified into two or more subject categories. Journals satisfying this definition play a bridging role between disciplines. Notice that this is different from the popular understanding of 'multidisciplinary' journals, such as *Nature* and *Science*. These latter publish all articles from many disciplines (but mostly mono-disciplinary ones) for a wide audience.

or scientific performance. This is echoed in more general terms, in noting the greater accuracy and reliability of ‘plural and conditional’ metrics in science, under pervasive conditions of uncertainty and ambiguity (Stirling, 2010).

Here we follow the main tenets of the ‘converging partial indicators’ method and enlarge its exploratory scope by using recently developed mapping techniques which help end-users explore their own partial perspectives. This is done by providing them with a range of diverse indicators. Not only is this mode of indicator usage more robust, it is also more likely to be recognised as legitimate in circumstances where diverse perspectives converge on a similar conclusion. This is arguably the case of the findings presented here specifically on interdisciplinarity. Likewise, when different approaches lead to contradictory insights, it becomes clear how conclusions are more questionable – and may reflect the choice of indicator rather than the phenomenon under investigation. This is arguably the case of the findings presented here specifically on performance. In order for the reader to be able to engage in this exploration, the full set of 54 maps (9 for each organisation) used for the analysis is made available at http://interdisciplinaryscience.net/maps/.

3.2 The assessment of interdisciplinarity

The inherently ambiguous, plural and controversial features of prevailing understandings of interdisciplinarity have led inevitably to a lack of consensus on indicators (see Wagner et al., 2011, for a review that emphasises the plurality of perspectives). Even within bibliometrics, the operationalisation of IDR remains contentious and defies uni-dimensional descriptions (Bordons et al., 2004; Huutoniemi et al., 2009; Leydesdorff and Rafols, 2011a; Sanz-Menéndez et al., 2001). We propose to investigate interdisciplinarity from two perspectives, each of which we claim has more general applicability. First, by means of the widely-used conceptualisation of interdisciplinarity as *knowledge integration* (NAS, 2004; Porter et al., 2006), which is perceived as crucial for innovation or solving social problems. Second, by conceptualising interdisciplinarity as a form of research that lies outside or in between established practices, i.e. as *intermediation* (Leydesdorff, 2007a).

Understanding interdisciplinarity as integration suggests looking at the distribution of components (disciplines) that have been integrated under a body of research (as shown by a given output, such as a reference list). We do so here by using the concepts of diversity and coherence, as illustrated in Figure 1. A discussion on how diversity and coherence may capture knowledge integration was introduced in Rafols and Meyer (2010).[7] We proposed to explore knowledge integration in two steps. First, employing the concept of *diversity* as 'an attribute of any system whose elements may be apportioned into categories' (Stirling, 2007, p. 708). This allows exploration of the distribution of disciplines to which parts of a given body of research can be assigned. A review of the literature, reveals that many bibliometric and econometric studies of interdisciplinarity were based on (incomplete, as we will see) indicators of diversity such as Shannon entropy (Carayol and Thi, 2005; Hamilton et al., 2005; Adams et al., 2007) and Simpson diversity (equivalent to the Herfindahl index in economics, and often used in patent analysis, e.g. Youtie et al., 2008). However, knowledge integration is not only about how diverse the knowledge is, but also about making connections between the various bodies of knowledge used. This means assessing the extent to which distant disciplines in the case under study are linked– which we explore with the concept of *coherence*.

Understanding interdisciplinarity as intermediation was first proposed by Leydesdorff (2007a), building on the concept of betweenness centrality (Freeman, 1977). As illustrated in Figure 2, intermediation does not entail combining diverse bodies of knowledge, but contributing to a body of knowledge that is not in any of the dominant disciplinary territories. As in the case shown in the right hand side of Figure 2, even when diversity is low the case can be considered interdisciplinary if a large part of its components are in intermediate positions.

A comparison between Figures 1 and 2 illustrates that knowledge integration and intermediation are two distinct processes. Although these properties may overlap, they do not need to occur at the same time. Indeed in a study on multiple measures of interdisciplinarity of journals, Leydesdorff and Rafols (2011a) found that they constituted two separate statistical factors.

[7] See a more general conceptual framework developed in Liu et al. (in press).

Knowledge integration, on one hand, occurs in research that builds on many different types of expertise. This is typically the case in emergent areas that combine disparate techniques from various fields, for example in medical application of lab on a chip, which draws both on micro-fabrication and biomedical expertise (Rafols, 2007). Intermediation, on the other hand, occurs when research does not fit with dominant disciplinary structures. This is often the case for instrumental bodies of knowledge, such as microscopy or statistical techniques that have their own independent expertise, yet at the same time provide a service contribution to different disciplines (Price, 1984; Shinn and Joerges, 2002). Intermediation may also show up in what Barry et al. (2008, p. 29) called agonistic research, that emerges in opposition to the intellectual, ethical or political limits of established disciplines. Such research tends to push towards fragmentation, insularity and plurality rather than integration (Fuchsman, 2007). As a result, it is seldom captured in conventional classification categories. We therefore investigate intermediation at a lower level of aggregation than diversity and coherence.

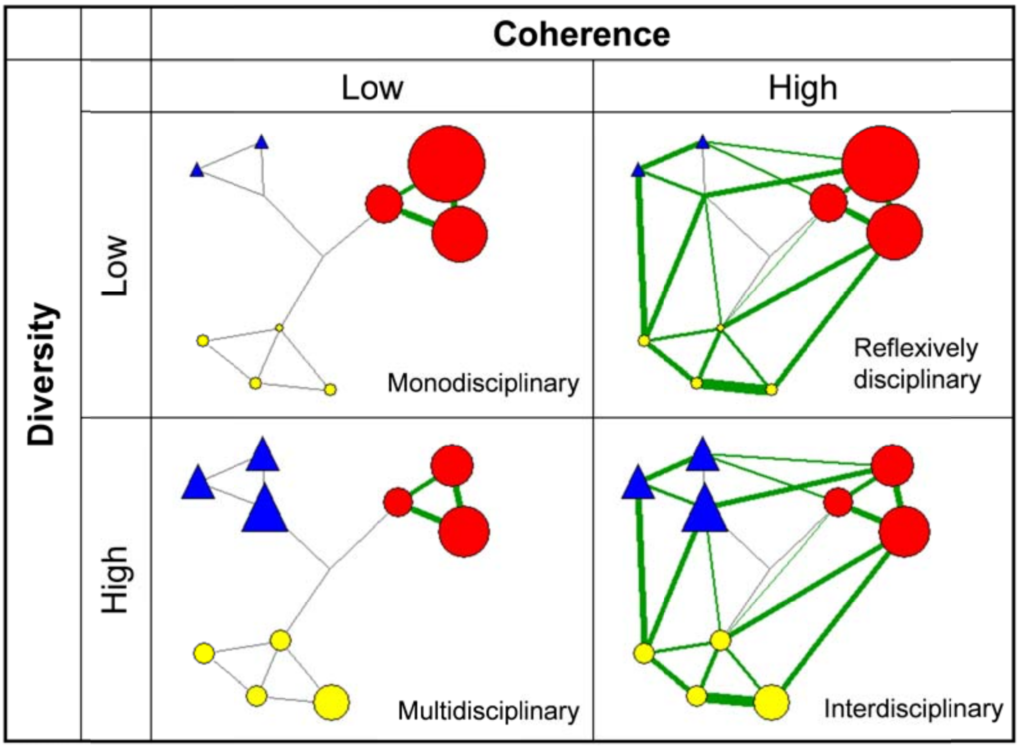


Figure 1. Conceptualisation of interdisciplinarity in terms on knowledge integration.

Each node in the networks represents a subdiscipline. Grey lines show strong similarity between subdisciplines. Same colours (shapes) illustrate clusters of subdisciplines forming a discipline. Green lines represent direct interaction between subdisciplines. The size of nodes portrays relative activity of an organisation in a given subdiscipline. Knowledge integration is achieved when an organisation is active in diverse subdisciplines and interlinks them.

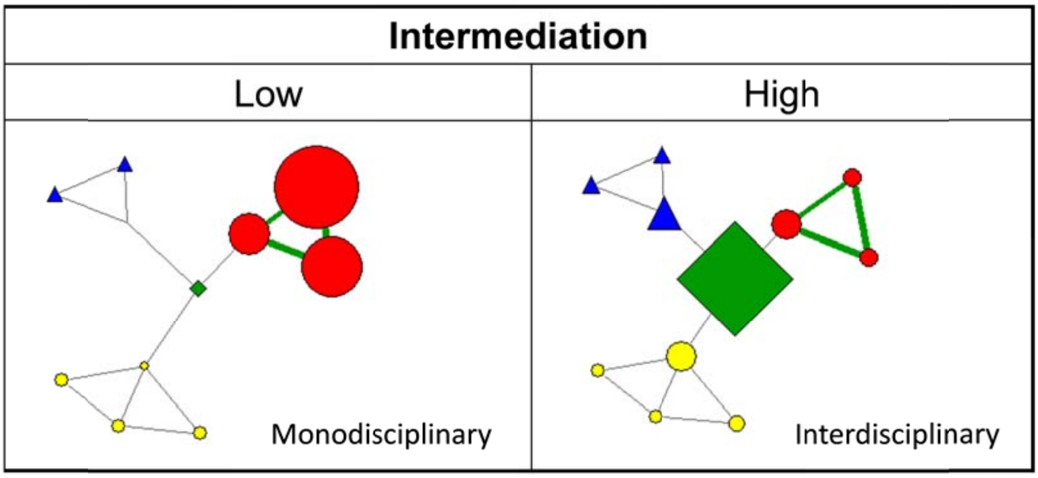


Figure 2. Conceptualisation of interdisciplinarity as intermediation.
See Figure 1 for an explanation of symbols. Intermediation is achieved when on organisation is active in subdisciplines that occupy an interstitial position, i.e. between other subdisciplines.

We now describe how the concepts of diversity, coherence and intermediation are operationalised. One advantage of using general concepts rather than ad-hoc indicators, is that it allows rigorous and plural comparison of – and choice between – different mathematical forms that are equally consistent with the processes we are seeking to capture. Hence the analysis follows the tenets of the ‘partial converging indicators’ approach (Martin, 1996). As a result, the emphasis is not simply on the incidental value of multiple indicators or their collective ranges of variability (Funtowicz and Ravetz, 1990). The aim is also to focus deliberate, self conscious, critical attention on the specific *conditions* under which different metrics (and their associated findings) are best justified (Stirling, 2008).

It has been widely documented across a diverse range of areas of appraisal, that there are often strong institutional pressures artificially to reduce appreciations of uncertainty and complexity in evaluation, in order to justify particular favoured interpretations (Collingridge, 1982). It is in light of this problem that we deliberately use a 'plural and conditional' framework (rather than an unqualified multiplicity of indicators), in order to increase the accuracy and robustness of policy appraisal (Stirling, 2010). By explicitly discriminating between multiple, contrasting quantitative characterisations of disciplinary diversity, coherence and intermediation, each with its associated rationale and applicability, we can better document the specific phenomena under scrutiny, and also contribute methodologically towards the general ends of addressing bias and ensuring legitimacy when using scientometric indicators.

Diversity

A given body of research, as represented for example in the publications of a university department, can be considered as more interdisciplinary if it publishes in diverse disciplines and the publications are coherent in the sense of linking the various disciplines. Diversity is a multidimensional property which has three attributes (Stirling, 1998; 2007): *Variety,* the number of categories of elements, in this case, the disciplines into which publications can be partitioned. *Balance,* the distribution across these categories, in this case, of output publications, or references in, or citations of, these (see details in methods, below). *Disparity*, the degree of distinctiveness between categories, in this case, the cognitive distance between disciplines as measured by using bibliometric techniques (Leydesdorff and Rafols, 2009).

An overlay representation of publications in the map of science captures these three attributes (Rafols et al., 2010; see Figure 1). It shows whether the publications (or references or citations) of a department scatter over many or a few disciplines (*variety*), whether the proportions of categories are evenly distributed (*balance*) and whether they are associated with proximate or distant areas of science (*disparity*). Since this is a multidimensional description, scalar indicators will either have to consider one of the attributes or make a compositional choice spanning the various possible scaling factors. Most previous studies on interdisciplinarity used indicators that rely on variety or balance (e.g. Larivière and Gingras, 2010), or combinations of both such as Shannon entropy (e.g. Carayol and Thi, 2005; Adams et al., 2007) – but crucially missed taking into account the disparity among disciplines. In doing so, they implicitly consider as equally interdisciplinary a combination of cell biology and biochemistry (related fields) and a combination of geology and psychology (disparate fields). Only recently, new indicators incorporating disparity were devised, using the metrics of similarity behind the science maps (Porter et al., 2007; Rafols and Meyer, 2010). This operationalisation of diversity also allows us to visualize processes of knowledge diffusion (rather than integration), by looking at the disciplinary distribution of cites to a topic or organisation's papers (Liu et al., in press).[8]

Following Yegros-Yegros et al. (2010), we employ indicators that explore each of the dimensions separately and in combination. As a metric of distance we use $d_{ij} = 1 - s_{ij}$ with s_{ij} being the cosine similarity between categories *i* and *j* (the metrics underlying the global science maps), with p_i being the proportion of elements (e.g. references) in category *i.* We explore the following indicators of diversity:

Variety (number of categories)	n
Balance (Shannon evenness)	$-\frac{1}{n}\sum_i p_i \ln p_i$
Disparity (average dissimilarity between categories)	$\frac{1}{n(n-1)}\sum_{i,j} d_{ij}$
Shannon entropy	$-\sum_i p_i \ln p_i$
Rao-Stirling diversity	$\sum_{i,j} p_i p_j d_{ij}$

Coherence

The term coherence refers to the extent to which the disciplines are connected to one another *within* the subset under study. Whereas measures of diversity are well established, measures of coherence (and intermediation) are still at an exploratory stage.[9] Here, to capture coherence we compare the *observed* average distance of

[8] In the case of research topics, also by exploring changes in the distribution over time (see Kiss et al., 2010; Leydesdorff and Rafols, 2011b).

[9] Rafols and Meyer (2010, pp. 273-274) operationalised coherence as the similarity (according to bibliographic coupling) among publications in a set, with the aim of revealing coherence of topics, rather than disciplinary coherence. Here, we use SCs as units of analysis since the question is whether units are linking or not the *disparate* disciplines in which they publish. This is why we look at expected/observed distances in citations.

cross-citations as they actually occur in the publications in question ($\sum_{i,j} p_{ij}d_{ij}$), with the *expected* average distance ($\sum_{i,j} p_i p_j d_{ij}$). This formulation assumes that within the set, the citations given by discipline *i* and received by discipline *j* are expected to be proportional to the product of their number of references in the set ($p_{i,j} = p_j p_j$). The observed/expected ratio shows whether the unit under investigation is linking or not the distant disciplines within its publication portfolio. By using a measure of diversity (Rao-Stirling) in the denominator, we ensure that this measure of coherence is orthogonal to diversity.

Coherence $$\frac{\sum_{i,j} p_{ij}d_{ij}}{\sum_{i,j} p_i p_j d_{ij}}$$

Intermediation

Intermediation aims to capture the degree to which a given sets of publication is distant from the most intensive areas of publication —those dense areas of the map representing the central disciplinary spaces. Since this measure is highly sensitive to artefacts created by the process of classification, we carry out the analysis at a finer level of description, namely the journal level (i.e. we use each journal as a separate category). We propose to use two conventional network analysis measures to characterise the degree to which an organisation's publications lie in these 'interstitial' spaces. The first is the clustering coefficient cc_i, which identifies the proportion of observed links between journals over the possible maximum number of links (de Nooy et al., 2005, p. 149). This is then weighted for each journal according to its proportion p_i of publications (or references, or cites), i.e. $\sum_i p_i cc_i$. The second indicator of intermediation is the average similarity of a given journal to all other *N* journals ($\frac{1}{N}\sum_j s_{ij}$) weighted by the distribution of elements (p_i) across the categories.[10]

Average similarity $$\sum_i p_i \left(\frac{1}{N} \sum_j s_{ij} \right)$$

[10] The robustness of clustering coefficient and average similarity as indicators of intermediation needs to be confirmed in further studies. They describe low-density landscapes –which are not always associated (as they can be shown to be in this case study) with intermediate or brokering positions. We thank Paul Wouters for this insight.

3.3 The assessment of performance

Because the contributions of scientific organisations are so diverse, their evaluation is necessarily complex. It becomes even more so if the evaluator attempts to capture societal contributions (Donovan, 2007; Nightingale and Scott, 2007). Since our research interest lies only in exploring the possible disadvantage IDR experiences in research assessment (rather than wider societal impact), we focus on conventional and widely used indicators specifically of scientific quality. These measures aim to capture quality according to what Alvin Weinberg (1963) called 'internal criteria', i.e. by means of criteria generated within the scientific field.

The first conventional indicator we use is the mean rank in the ABS journal rankings of the publications of a given research unit. The ABS journal rankings are 'a hybrid, based partly upon peer review, partly upon statistical information relating to citation [i.e. on the Thompson-Reuters Impact Factor (IF)], and partly upon editorial judgements' (ABS, 2010, p.1). It is created by leading academics belonging to BMS of the ABS – thus it follows *internal* criteria. The function of these journal rankings is to indicate 'where best to publish', to inform library purchases and staffing decisions such as 'appointment, promotion and reward committees' and to help to aid 'internal and external reviews of research activity and the evaluation of research outputs' (ABS, 2010, p. 1). Beyond being simply a matter of high correlation (between RAE results and ABS ranking), these rankings are an explicit part of the BMS 'culture' and are routinely used for recruitment and promotion.

A second conventional indicator is the mean number of citations per publication. Narin and Hamilton (1996, p. 296) argued that bibliometric measures of citations to publications provide one *internal* measure of the impact of the contribution, and hence a proxy of its *scientific* performance. The number of citations per publication (or citation impact) is neither an indicator of quality nor importance. Instead it is a reflection of one form of influence that a publication may exert, which can be meaningfully used in evaluation provided some caveats are met (see detailed discussion in Martin and Irvine, 1983, pp. 67-72; Leydesdorff, 2008).

One of the key caveats of citation per paper as a performance indicator is that different research specialties adopt contrasting publication and referencing norms, leading to highly diverse citation propensities. Hence, some form of normalisation to adjust for such differences between fields is '[p]erhaps the most fundamental challenge facing any evaluation of the impact of an institution's program or publication' (Narin and Hamilton, 1996, p. 296). The most extensively-adopted practice is to normalise by the discipline to which is assigned the journal where the article is published. Here, the field-normalised cites/paper was calculated by dividing the cites of a given paper by the average cites per paper of the publications of a disciplinary category (data obtained from the 2009 Journal Citation Report, JCR).

Given data availability, we rely on Web of Science (WoS) Subject Categories (SC)[11] as disciplinary categories. Although grossly inaccurate for individual papers, they produce meaningful results for sufficiently large numbers of publications seen from at the scale of global science as a whole (Rafols and Leydesdorff, 2009). One advantage of the SCs is that they are mostly defined at the subdiscipline level (e.g. *Organic Chemistry*), allowing varying degrees of larger 'disciplinarisation' according to their clustering in the global map of science, instead of having to rely on 'essential' disciplinary definitions.

Though widely used, the field normalisation is known to be problematic (Leydesdorff and Opthof, 2010). This is, first, because allocations of journals to disciplines can be made in a variety of contrasting but equally useful ways. There are major discrepancies between various conventional disciplinary classifications, such as the WoS or Scopus' categories, which are designed for retrieval purposes but are not analytically robust (Rafols and Leydesdorff, 2009). A second reason is because some (perhaps especially interdisciplinary) papers may not conform to the conventional citation patterns of a journal — they may have a 'guest' role in a category. For example, publications on science policy in medical journals. As a result of these difficulties, normalisations using different field delineations (or levels of aggregation) may lead to different pictures of citation impact (Zitt et al., 2005; Adams et al., 2008).

[11] This study uses 'Subject Categories' of Web of Science's version 4. Notice that in version 5 (as of September 2011), these categories have been relabelled as 'Web of Science Subject Categories', with WC as new acronym.

A way to circumvent the problem of delineating the field of a publication is to try to normalise from the perspective of the audience, i.e. via those publications citing the publications to be assessed. One way to normalise from the citing-side is by making a fractional citation count, whereupon the weight of each cite is divided by the number of references in the citing publication. Fractional counting was first used for generating co-citation maps by Small and Sweeney (1985). Only recently Zitt and Small (2008) recovered it for normalizing journal 'audience' (following a discussion in Zitt et al., 2005), and then Leydesdorff and collaborators developed it for evaluation purposes at the paper level (Leydesdorff and Opthof, 2010; Zhou and Leydesdorff, 2011).

Citing-side normalisation can be particularly appropriate for interdisciplinary cases (which receive citations from publications with different citation norms) because it normalises in a way that is not dependent on classifications (Zhou and Leydesdorff, 2011; Zitt, in press). Although this corrects for the differences in the number of references in the citing paper, it may not correct for differences in their publication rates. For the purposes of this study, the citing side normalisation is made using only the downloaded citing records (i.e. excluding any cite from the unit being investigated), then giving each a cite weight inverse to their number of references, i.e. $\frac{1}{\# References}$. Only cites with more than 10 references (including self-citations) are used, since papers with less references would have a disproportionately high weight (and tend not be a 'normal' publication outlet).

Following conventional practice, in all cases we use the mean to describe the citation distributions. This has long been widely acknowledged to be a flawed method given the highly skewed nature of citation distributions (Narin and Hamilton, 1996, pp. 295-296; Katz, 2000; Leydesdorff & Bornmann, in press; Leydesdorff and Opthof, 2011). As we will see, it also leads to very high standard errors that can often render the differences between performance indicators statistically non-significant.

Finally, we also include measures based on the journal IF, despite its widely known lack of scientific validity (e.g. Seglen, 1997). We do this for two reasons. First, to examine performance under a widely used indicator. Second, and more importantly, to check if the results of the ABS ranking performance are driven by its partial reliance on journals' IF.[12] We compute the mean IF of the journal of publications, the mean normalised by the SC of the publication journal, and the mean IF of the citing journals.

3.4 Data

We investigate three of the centres of IS in the UK: the Manchester Institute of Innovation Research (MIoIR) at the University of Manchester (formerly known as Policy Research in Engineering, Science and Technology, PREST), SPRU (Science and Technology Policy Research) at the University of Sussex and the Institute for the Study of Science Technology and Innovation (ISSTI) at the University of Edinburgh.

The choice was in part by the perceived importance of these centres in the establishment of IS in the UK (Walsh, 2010) and in part determined by lack of coverage in the WoS of more discursive social sciences, which are more prevalent in other centres such as the Institute of Science Innovation and Society (InSIS) at the University of Oxford, whose research tends to concentrate more in the field of 'Science and Technology Studies' (STS) than IS. These IS units are compared with three of the leading British BMS: London Business School (LBS), Warwick Business School (WBS) and Imperial College Business School (formerly Tanaka).

The publications of all researchers identified on institutional websites as members of the six units (excluding adjunct, visiting and honorary positions) were downloaded from Thomson-Reuters Web of Science for the period 2006-2010. The downloads were limited to document types: 'article', 'letter', 'proceedings paper' and 'review'. Publications by a researcher prior to their recruitment to the unit were also included. The download was carried out between 20 and 30 October 2010 (except for SPRU publications downloaded on 22 May 2010 with an update on 26 October 2010). Additionally, publications citing these researchers' publications were also downloaded in the same period (including SPRU's).

[12] We thank an anonymous referee for this suggestion.

The analysis of all other data relating to journals and SCs is based on the CD-ROM version of the Journal Citation Report, following routines explained in previous work (Leydesdorff and Rafols, 2009). A caveat to this approach is that we use the full WoS (~11,000 journals) for the units' data, while relying on JCR data (~9,000 journals) to carry out part of the analysis (such as global maps of science or the field normalisation). In doing this we are assuming that the structure of the WoS and JCR are equivalent.[13]

In order fully to disentangle the analytical results of a unit's publications from the unit's cites, all citing documents from the same unit were removed (i.e. self-citation and cites from institutional colleagues were not included in the citing subset). Due to the retrieval protocol used for the citing papers (researcher-based), those papers repeatedly citing the same author were counted only once, whereas those papers citing collaborations between multiple researchers in the same unit were counted once for each researcher. This inaccuracy only affects the part of the analysis regarding cites (not the publications or references) and is not expected to result is a serious distortion since intra-organisational collaborations represent about 10% of publications.

3.5 Data processing and visualisation

The software Vantage Point (http://www.thevantagepoint.com) was used to process data. A thesaurus of journals to WoS SCs was used to compute the number of aggregated cited SCs in the cited references (Porter et al., 2007, p. 125). The proportion of references which it was possible to assign in this way ranged between 27% for ISSTI to 62% for LBS. These proportions are low partly due to variations of journals names within the references that could not be identified, and partly due to the many references to books, journals and other type of documents not included in the WoS. However, the analysis is statistically robust since between ~1,500 and ~10,300 references were assigned for each unit. A minimum threshold was applied at 0.01% of total publications, in order to remove low counts of SC (statistical noise) in

[13] We thank Thed van Leeuwen at CWTS for making us aware of this potential source of error.

the variety and disparity measures. No threshold was applied in calculating balance, Shannon Entropy, and Rao-Stirling measures; since these are not affected by small number of counts.

The ABS rank for each journal was obtained from the Academic Journal Quality Guide Version 4 (ABS, 2010). The journals are classified into five categories: 1, 2, 3, 4 and 4*. This was used to calculate the average ABS rank for each unit. Each level was weighted according to its ascending ordinal position (1,2,3,4), and 4* rank was given a weight 5. Additionally, SCs were assigned to all journals in the ABS Ranking guide which were in the JCR, which amounted to 60% of the ABS list. This data was used to map the disciplines of each ABS rank, with the node size area corresponding to the proportion of journals in that particular rank belonging to each SC. Cites/paper were computed using the WoS field *Times Cited* (TC) in the WoS record. As a result of the earlier download of SPRU data, the *Times Cited* field of SPRU papers had to be extrapolated.[14] Intermediation measures were computed with Pajek using the journal similarities matrix. The average clustering coefficient (at two neighbours) was computed with a 0.2 threshold.

The freeware Pajek (http://pajek.imfm.si, de Nooy et al., 2005) was used to make all networks except Figure 5. First, disciplinary overlay maps were made by setting the sizes of nodes proportional to the number of references in a given SC, as explained in Rafols et al. (2010)[15], using 2009 data for the basemap (grey background). Second, cross-citations maps (green links) between SC were generated and overlaid on the disciplinary maps in order to generate Figure 3. Lines are only shown if they represent a minimum of 0.2% of cites and more than 5 fold the expected proportion of cross-citation among SCs, in comparison to average WoS cross-citation flows. This shows the extent to which the relations between disciplines are novel or, on the contrary, whether they occur along well trodden paths.

[14] The extrapolation was carried out as follows. In October 2010, 730 unique papers citing SPRU papers were found in the WoS. For the other five units there were an average 8.5% discrepancy between the unique papers found in WoS citing them, and the counts in TC. By using this average discrepancy, 792 cites (730 cites plus the 8.5% discrepancy) were estimated for SPRU. The possible inaccuracy introduced by this extrapolation is well within the standard error (~10%).

[15] The method is made publicly available at http://www.leydesdorff.net/overlaytoolkit/ and http://www.idr.gatech.edu.

The freeware VOSviewer (http://www.vosviewer.com, Van Eck and Waltman, 2010) was used to make a journal map in the journal density format. A sub-set of 391 journals was made from the journals where each unit published (excluding journals which contributed less than 0.5% publications for each single unit) and the top 100 journals which all units (collectively) referenced. The cross-citations between these journals were obtained from the 2009 JCR. This was used to compute the cosine similarities matrix in the cited dimension, which was input into VOSViewer. The size of nodes was determined by the number of publications (or references or cites) per journal, normalised to the sum of all publications (or references or cites) –and overlaid on the basemap. We are currently developing an off-the-shelf application to generate this type of overlay journal maps (Leydesdorff and Rafols, in preparation).

4. Results

4.1 Interdisciplinarity of organisational units

The following sections present the results of this investigation. First we show that IS units are more interdisciplinary than BMS according to three different perspectives and associated metrics.

Diversity and coherence

Figure 3 shows the overlay of the publications of ISSTI (top) and LBS (bottom) over the global map of science – as a representative illustration of the findings in this analysis regarding the general contrast between the three IS units (including ISSTI) and the three comparator BMS (including LBS). The full set of diversity maps for all the organisations can be found at www.interdisciplinaryscience.net/maps and in the supplementary materials.[16] The set of overlay maps were generated for each of the six units and then for the SCs of publications, references and cites (excluding self-

[16] We skip the details of the overlay technique, since it is discussed at length in Rafols et al. (2010). To see the relative positions on the global maps of science see the interactive map http://www.idr.gatech.edu/maps.
The maps complementing Figure 1 can be retrieved in:
http://www.sussex.ac.uk/Users/ir28/IDR/Disciplinary_Diversity.pptx
http://www.sussex.ac.uk/Users/ir28/IDR/Disciplinary_Coherence.pptx

citation). These results show that IS units are cognitively more diverse in the sense that they spread their publications (references, cites) over a wider set of disciplines (variety), do so more evenly (balance), and across larger cognitive distances (disparity). No time trends were found. The differences are more pronounced in the case of publications and cites than for references[17], which tend to be relatively widely spread for both IS and BMS. These insights are shown in the form of indicators in Table 1 and Figure 4.

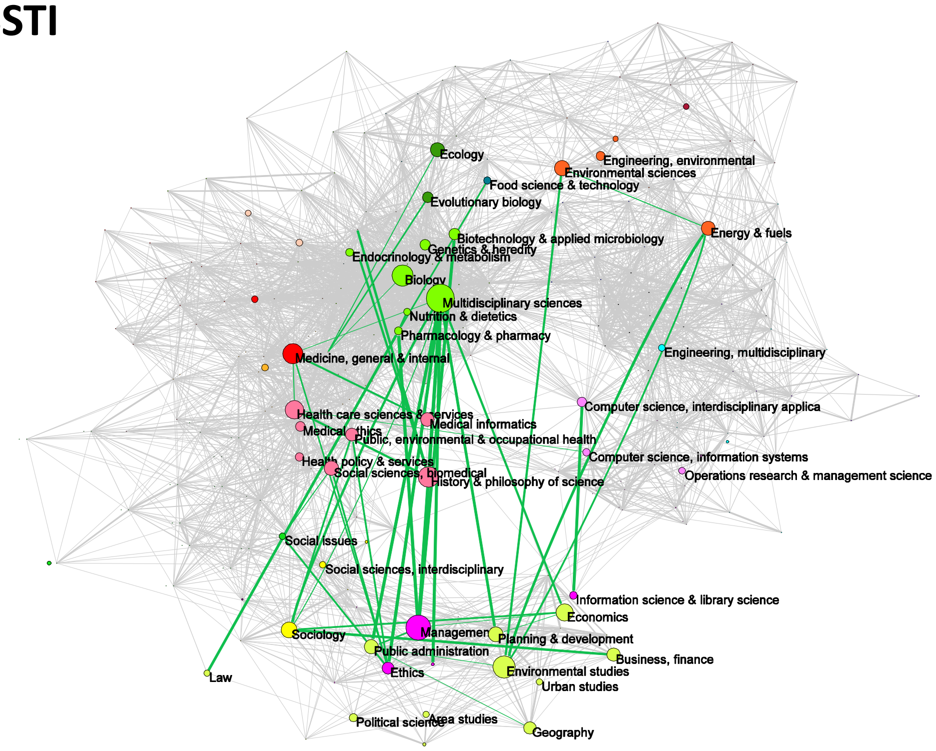
Second, not only are IS units more diverse, but their publications cite more widely across distant SCs than might be expected from the distribution of cross-citations between SC in the WoS. This is shown by the green links overlaid in Figure 3: cross citations between SCs are more than 5-fold the average proportion in the global map of science. For example, ISSTI has major citation flows between management and biomedical sciences, which are rare in the global citation patterns, and SPRU between economics and planning with ecology, environment and energy. This is evidence that these IS units are not only diverse in the sense of 'hosting' disparate disciplines, but are actually linking them. In particular, they play a bridging role between the natural sciences and social sciences.

By contrast, the leading BMS examined here are not only less diverse, but also more fragmented (less coherent) in disciplinary terms, in the sense that they tend to cite more within specialties or disciplines. For example, Imperial is the most diverse of the BMS, thanks in part to its research groups on IS and healthcare management. However, this latter line of research is not strongly linked to other social sciences at Imperial, as shown by the scarcity of cross-citations. In this case, then, co-location of health research and management in the same BMS does not appear to lead to interdisciplinary exchange. The bridging function between the natural sciences and social sciences carried out by IS units is captured by the coherence indicator shown in Table 1 and illustrated in Figure 4.

[17] The results from diversity measures from citing articles (excluding self-cites within units) are important because they refute the possibility that the larger diversity of IS publications is caused by references to the topics of the article (heath or energy), rather than genuine scholarly engagement with distant discipline. This was confirmed by an exploration of the abstracts of articles citing IS units from the natural sciences or engineering. The sample revealed that these articles included both conventional articles embedded in the discipline and policy or opinion papers reflecting on the topics (which might be by other IS scholars). Further research is needed to understand the role of publications and references by IS scholar in the context of practitioners' journals.

Measures such as diversity might have size effects, i.e. they tend to increase or decrease depending of the population size. Since the IS units are between 2 to 4 times smaller than BMS, one might wonder whether size-effects would explain the differences in the diversity measures. However, the most obvious size effect would be that larger units tend to display greater diversity, since they have a higher probability of having a very small proportion of publications/references/cites in some SCs. Since the observed relation is the inverse, i.e. the smaller units have the highest diversity, one can be confident that the results are not an indirect effect of size. Indeed they are evident despite such an effect, and are correspondingly likely to be stronger if size were controlled for. There is no size effect expected in the case of coherence, given that it is computed from a ratio.

ISSTI



LBS

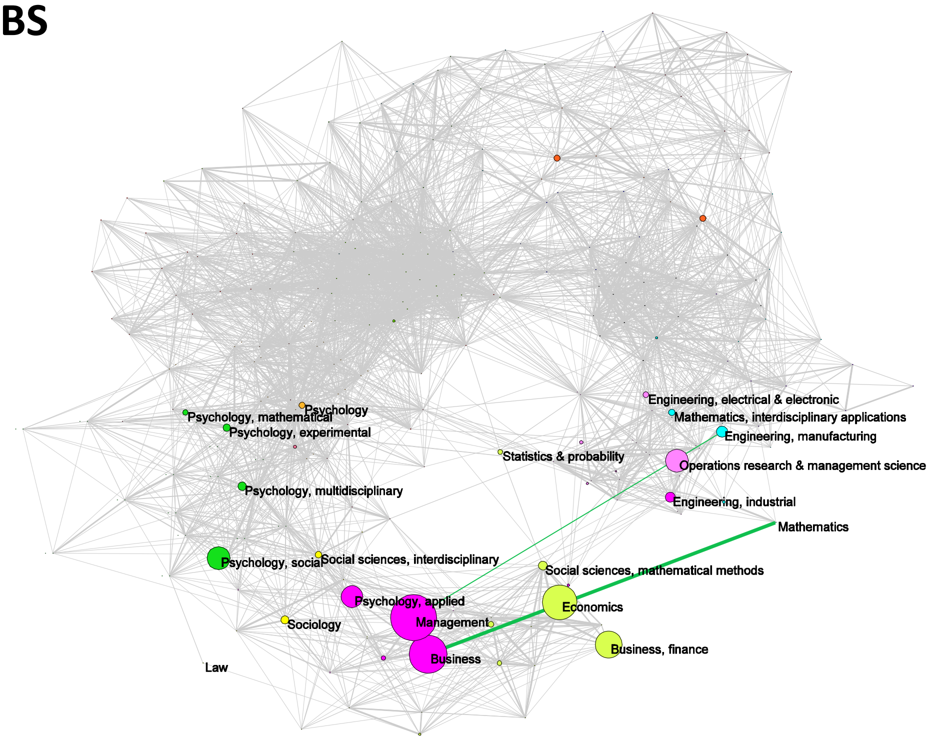


Figure 3. Overlay of number of references on SCs (source) by unit on the global map of science. The extent of referencing (or citing) between SCs (green links) by a given unit is shown only for observed values 5 fold larger than expected**.** Each node represents a subdiscipline (SC). Grey lines portray similarity between SCs. The degree of superposition in the grey background illustrates the degree of similarity between different areas of science for all 2009 WoS data. Diversity of references (spread of nodes over map) and referencing across disparate SCs (amount of cross-linking) are interpreted as signs of interdisciplinarity.

Table 1. Indicators of diversity and coherence for each organisational unit

	Innovation Studies Units			Business and Management Schools		
	Edinburgh ISSTI	**Sussex SPRU**	**Manchest MIoIR**	**Imperial College Business School**	**Warwick Business School**	**London Business School**
# of Publications	129	155	115	244	450	348
SC of Publications						
Variety	28	20	19	15	20	9
Balance	0.653	0.566	0.543	0.485	0.460	0.370
Disparity	0.832	0.839	0.817	0.788	0.770	0.768
Shannon Entropy	3.558	3.243	2.966	2.970	3.078	2.343
Rao-Stirling Diversity	**0.810**	**0.783**	**0.726**	**0.720**	**0.680**	**0.603**
# of References	1737	2409	1558	6017	8044	10381
SC of References						
Variety	28	18	17	17	20	15
Balance	0.510	0.420	0.415	0.347	0.325	0.287
Disparity	0.829	0.842	0.846	0.832	0.780	0.825
Shannon Entropy	4.115	3.575	3.378	3.251	3.153	2.802
Rao-Stirling Diversity	**0.833**	**0.791**	**0.729**	**0.731**	**0.689**	**0.682**
# of Cites	316	767	419	1229	1246	1593
SC of Cites						
Variety	32	21	22	20	24	15
Balance	0.669	0.513	0.505	0.452	0.454	0.379
Disparity	0.852	0.844	0.836	0.819	0.801	0.767
Shannon Entropy	4.222	3.723	3.415	3.482	3.503	2.985
Rao-Stirling Diversity	**0.851**	**0.810**	**0.771**	**0.755**	**0.736**	**0.679**
Cites between SCs						
Coherence	**0.730**	**0.747**	**0.803**	**0.597**	**0.663**	**0.539**

Note: higher values for each metric indicate higher levels of the indicated property.

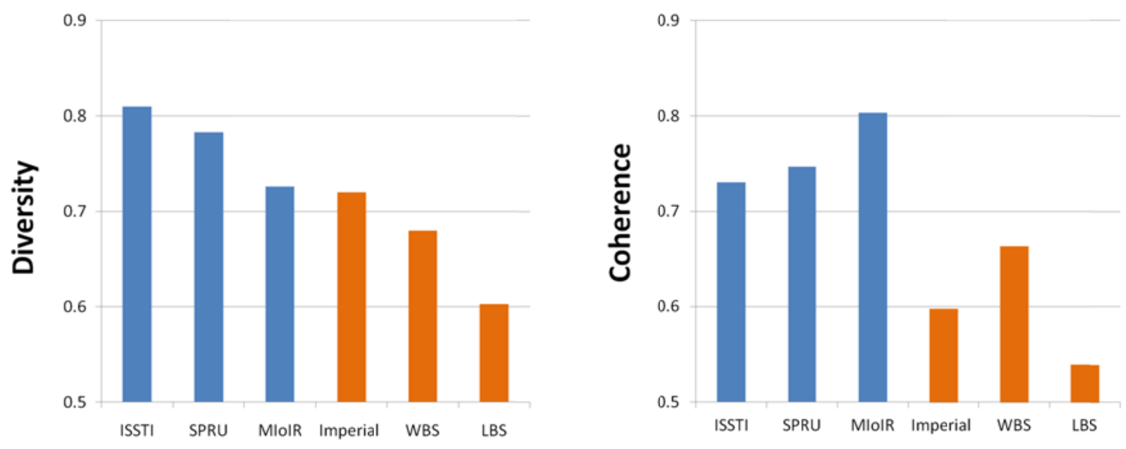


Figure 4. Indicators of Diversity (Rao-Stirling) and Coherence for the publications by organisational unit.

Intermediation

The third property of IDR we want to investigate is whether a given body of research lies within or between existing disciplinary boundaries. For this purpose the WoS SCs are too coarse. Instead of using the SC disciplinary maps, we created maps of the main 391 journals in which the six units examined here publish and reference (see methods). We used the density visualisation option of the software VOSviewer which is helpful to distinguish between the dense areas associated with disciplinary cores, and sparser interstitial areas associated with IDR. To make the basemap, SC-to-SC cross-citation data from the 2009 JCR was used to generate a similarity matrix, which then served as input for the visualisation programme. The publications, references and cites associated with each unit were then overlaid on this map. Notice that this map is on a different basis to conventional journal maps (where positions reflect direct, local similarities, since they have no overlay).

The local journal maps of IS-BMS (Figure 5[18]) show three poles: management, economics, and natural sciences. The sharp polarisation between economics and management is fully consistent with the findings by Van Eck and Waltman (2010, pp.

[18] More maps complementing Figure 3 can be retrieved from: http://www.sussex.ac.uk/Users/ir28/IDR/Intermediation.pptx

529-530). [19] Interestingly, *Research Policy*, which was identified as the most important and central journal of IS by Fagerberg et al. (this volume), occupies an equidistant position between the management and the economics poles – with some pulling towards the natural sciences.

The third pole encompasses the various specific natural sciences studied by these units. The map reveals that within the combined IS-BMS context, journals of different natural sciences are cited similarly, in comparison to the differences among the citations to social science journals. Thus, unlike the economics and management areas, this third pole can be interpreted as an artefact rather than a genuine disciplinary core in its own right. It is nevertheless useful since it provides an axis to show the degree of interaction between the social sciences and the natural sciences. Journals that are more science- than innovation-oriented, such as *Social Studies of Science* and *Scientometrics*, are closer to this pole. The relative position of the different areas is quite consistent with that of the global map of science but here some areas such as business and economics have been 'blown up', whilst the natural sciences have been compressed. The effects of these shifting spatial projections are neutral with respect to the conclusions drawn here.

The overlay maps in Figure 5 show that BMS units mainly publish, reference and are cited by journals in the dense areas of management and economics. IS units, on the contrary, have most of their activity in the interstitial areas lying between management, economics and the natural sciences, that is, in journals such as *Research Policy*, or in journals of application areas such as *Social Science and Medicine* or *Energy Policy*. These differences among units in the position of the journals where they publish can be rendered quantitatively by means of the indicators clustering coefficient and average similarity of the journals as described in the (Table 2 and Figure 6). In summary, what the journal maps show is that IS units carry out their boundary-spanning role, at least in part, by means of interdisciplinary journals.

[19] See Van Eck and Waltman's (2010) interactive maps at http://www.vosviewer.com/maps/economics_journals/

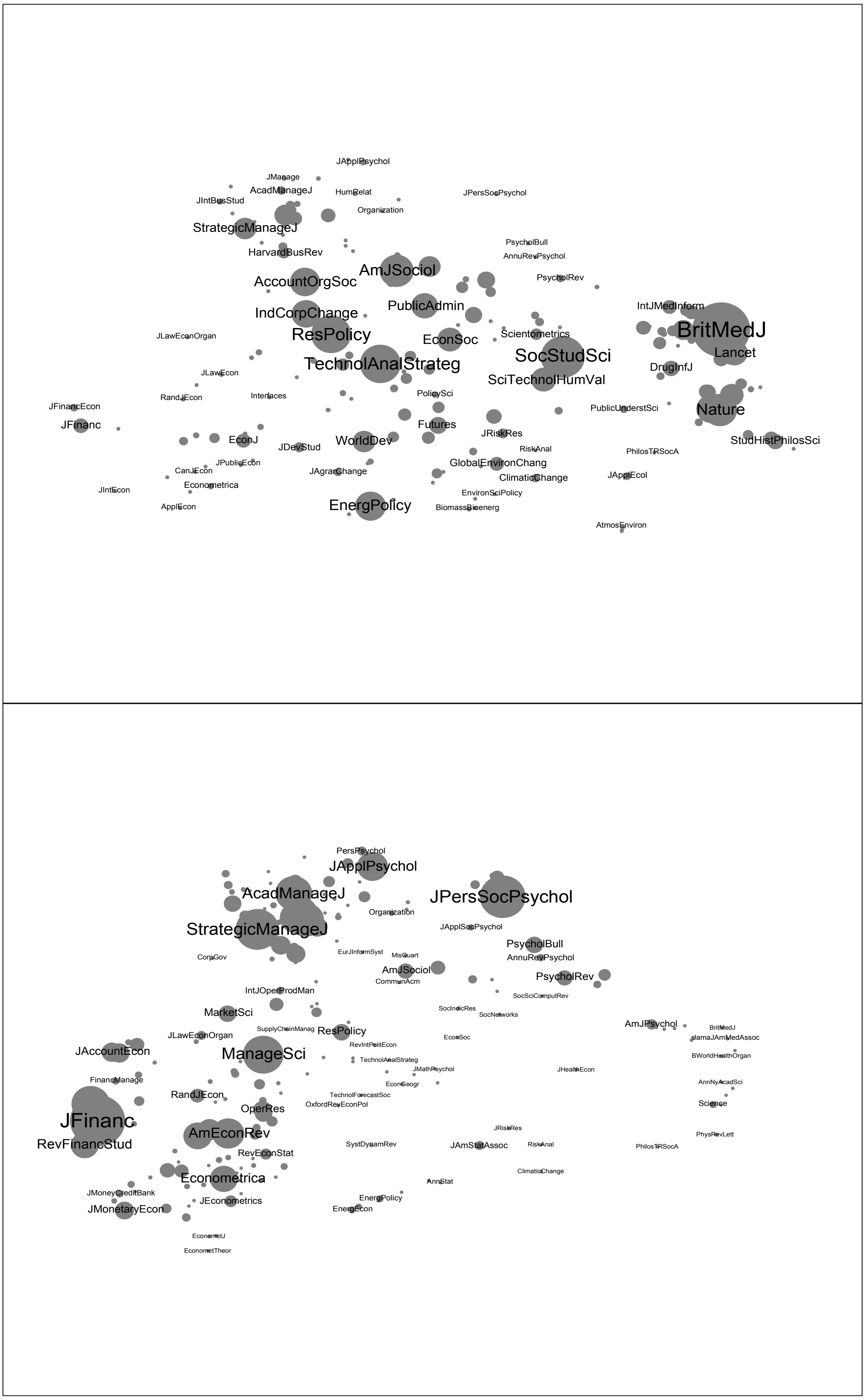


Figure 5. Overlay of the references of ISSTI and LBS publications in a journal map. The map illustrates the similarity structure of the 391 most important journals for all the IS and BMS units analysed. Red areas correspond to high density of journals –indicating areas of mono-disciplinary activity. Green areas show low density. Node size shows the proportion of

a unit's references in a given journal. An intermediate position in the map is interpreted as a sign of interdisciplinarity.

Table 2. Indicators of intermediation by organisational unit.

	Innovation Studies Units			Business and Management Schools		
	ISSTI	**SPRU**	**MIoIR**	**Imperial**	**WBS**	**LBS**
Journals of pubs.						
Clustering Coeff.	0.128	0.098	0.075	0.189	0.165	0.202
Average similarity	0.028	0.034	0.036	0.050	0.045	0.060
Journals of references						
Clustering Coeff.	0.178	0.182	0.166	0.236	0.221	0.235
Average similarity	0.044	0.050	0.058	0.066	0.065	0.068
Journals of cites						
Clustering Coeff.	0.120	0.096	0.074	0.157	0.167	0.183
Average similarity	0.029	0.034	0.037	0.046	0.044	0.055

Note: low values for each metric indicate higher levels of intermediation. Standard errors are not provided because they are negligible (all smaller than 0.07%).

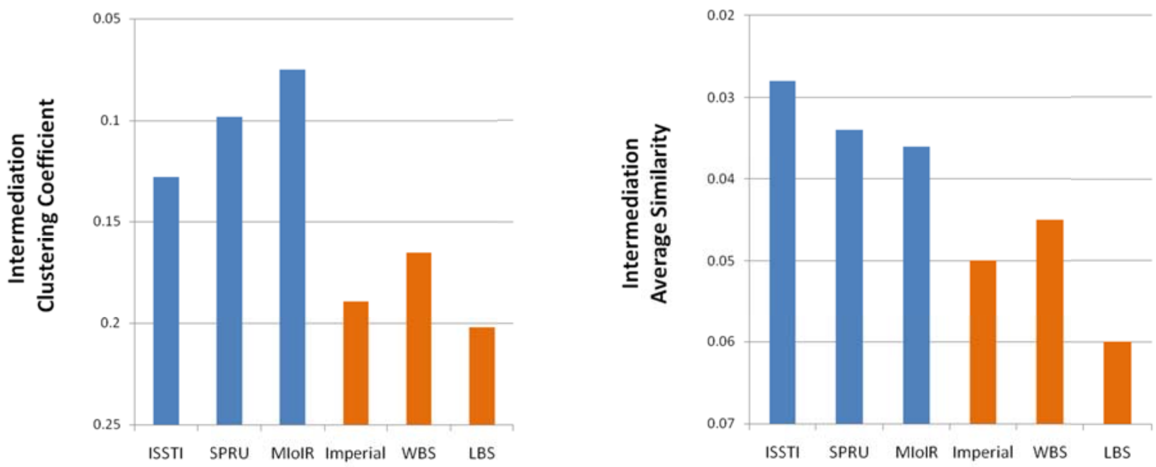


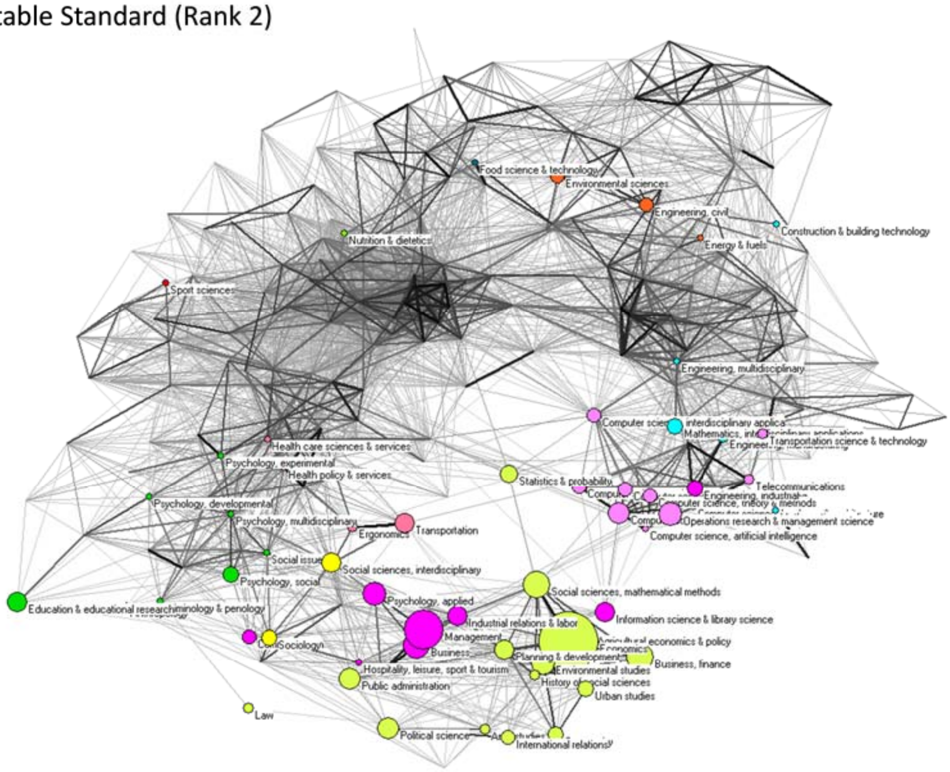
Figure 6. Indicators of intermediation of publications by organisational unit.

4.2 Disciplinary bias in ABS journal rankings

Now we turn our attention to the disciplinary profiles of the journals under different ranks in the ABS classification. For each Rank, from 1 (the lowest quality), to 4* (the highest), we use the JCR to assign journals to SCs. The JCR coverage of the ABS journals was low for rank 1 (14%), but reached an acceptable level for rank 2 (56%), and was almost complete at the highest ranks. We analyze the disciplinary diversity of each rank in terms of its distribution of journals in SCs, following the same protocol as in the previous sections (only now the basic elements are journals, rather than articles). The results are shown in Table 3 and Figures 5 and 6.[20]

[20] Maps complementing Figures 5 and 6 can be retrieved from: http://www.sussex.ac.uk/Users/ir28/IDR/ABS_Ranking_Diversity.pptx

Acceptable Standard (Rank 2)



World Elite (Rank 4*)

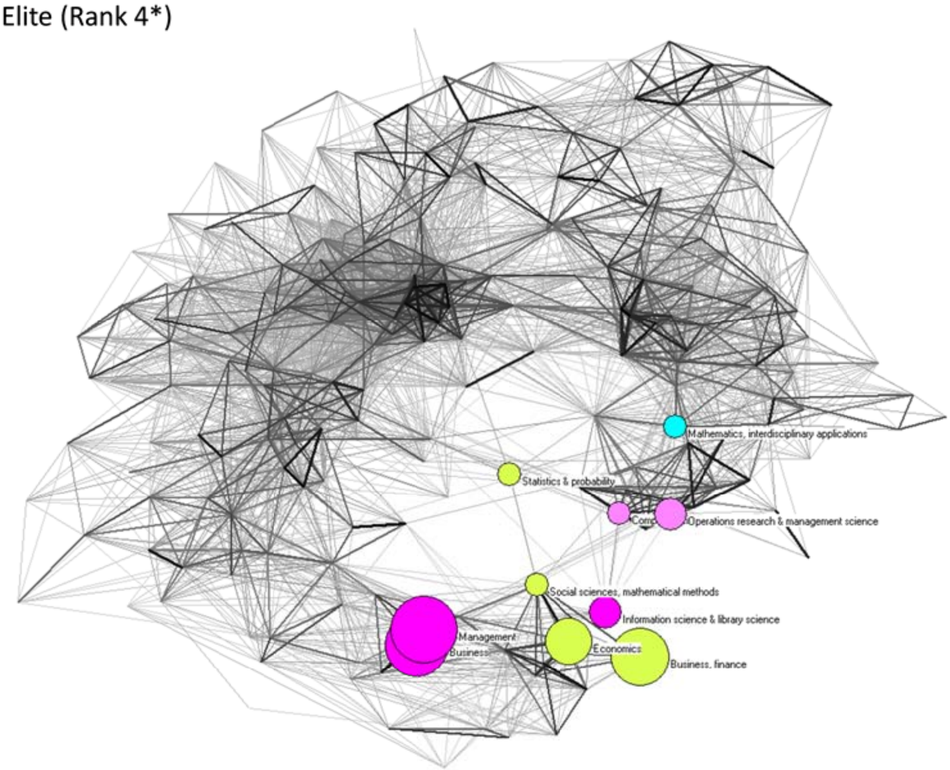


Figure 7. Distribution of journals across different categories for Association of Business Schools' Rank 2 (Acceptable Standard) and Rank 4* (World Elite).

Table 3. Disciplinary diversity indicators of the Association of Business Schools' rankings

	Rank 1 Modest standard	Rank 2 Acceptable standard	Rank 3 Highly regarded	Rank 4 Top in Field	Rank 4* World Elite
# of Journals in JCR	29	166	199	73	21
SC of Journals					
Variety	27	58	56	31	10
Balance	0.797	0.611	0.558	0.606	0.573
Disparity	0.866	0.819	0.792	0.789	0.767
Shannon Entropy	2.979	3.454	3.280	2.940	2.002
Rao-Stirling Diversity	**0.779**	**0.733**	**0.703**	**0.685**	**0.571**

Note: higher values for each metric, indicate higher levels of the indicated property.

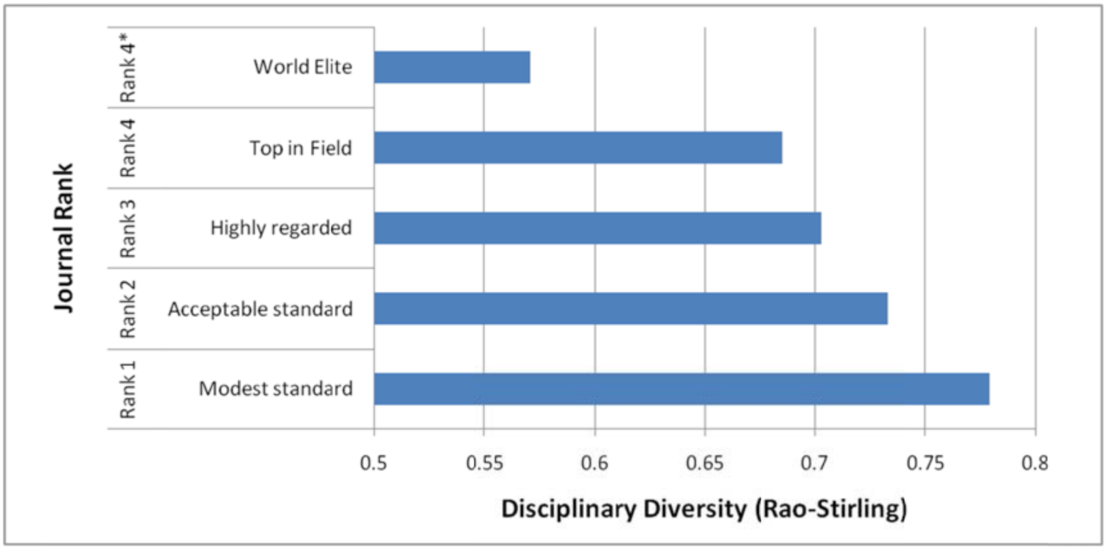


Figure 8. Diversity of the disciplinary distribution of journals for each rank.

These data show that the highest ranking journals are much less diverse than the lowest ranking journals. In particular, the top rank (4*), narrowly focuses on three SCs: management, business and finance. Lower ranks are spread across various social sciences, including economics, geography, sociology, psychology, and some engineering-related fields such as operations research and information science, as well as some application such as environment or food. Thus, while ABS rankings include journals from many disciplines, only some of those in their core subject matters are perceived by ABS as 'World Elite' journals.[21]

[21] The fact that ranks 4* and 4 only contain 21 and 73 journals, in comparison to the more than 100 journals in the ranks 3 and 2, might partly explain why the higher ranks are less diverse. This certainly has an effect in the

4.3 Performance of organisational units

Finally, we can now explore how the disciplinary bias in the ABS journal rankings affects the assessment of organisational units. To do this, we took the mean of the ranks of the journals in which the units publish. In doing so, we first notice a problem of assignation: whereas only 43% of ISSTI or 51% of SPRU journals that are listed in the WoS are also included in the ABS rankings, coverage reaches 79% and 93% of their WoS journals in the case of WBS and LBS, respectively. The results are shown in Table 4 and Figure 9.[22] They conclusively show that the three BMS perform significantly better than the IS units. Within the BMS, the narrow disciplinary profile of LBS achieves a much higher figure than the other two BMS. This is associated with the strong negative Pearson correlation between degree of interdisciplinarity across any metrics and ABS-based performance: -0.78 (Rao-Stirling diversity), -0.88 (coherence), 0.92 (Intermediation, clustering coefficient).

Next we compare the ABS-based performance with citation-based performance. We should emphasize that this performance analysis is only exploratory. Since we are counting cites received by groups of papers in the whole 2006-2010 period and analysing cites received until 2010 instead of using fixed citation windows, the results should be interpreted as only indicative.[23] Although imperfect, the estimates obtained are expected to be sufficiently robust to provide tentative insights and illustrate the inherent difficulties and ambiguities of using citation-based performance indicators.

First, it is important to notice that the standard error is extremely high (in the ~8-18% range) – so high that ranking units becomes problematic. This is the result of the

number of SCs and some effect on Rao-Stirling and Shannon diversity. However, the key insight here comes from understanding whether all the disciplines of the top-rank journals are associated with the same discipline or not. The overlay in Figure 7 supports the insight that highly ranked journals in the ABS list cover a smaller area in the science maps. This should have been reflected in the measure of disparity—the fact that it does not appear is due to that the distance metrics we use is only sensitive to short range differences between SCs (i.e. it gives similar distances when measuring Business vs. Economics, and Business vs. Astronomy). This suggests that there is room for improving the distance metrics used (Leydesdorff and Rafols, 2011a).

[22] Also available in powerpoint format at:
http://www.sussex.ac.uk/Users/ir28/IDR/Performance_Comparison.pptx

[23] Using fixed citation window only allows studies of past research. In this case, we should have studied the period 2001-2005 in order to have 5-year citation windows for each year document. But doing so would have created 'past' portrays of the units, and major hurdles in data gathering due to researchers' mobility.

conventional statistical (mal)practice of using the mean to describe skewed distributions (Leydesdorff and Bornmann, in press). Given these high values of statistical deviation, even for large schools, it is surprising, that citation-based research rankings are allowed to be used for publicity purposes by BMS, or by magazines such as the *Times Higher Education* when ranking universities or departments, without reporting their degree (or lack) of statistical significance.

The analysis shows first that BMS units do not perform better than IS units in terms of raw numbers of cites. Second, normalisation by journal IF shows that IS and BMS units have similar citation frequencies within the journals where they publish (not shown in Figure 9). Third, using a field-based normalisation one obtains a relative decrease of IS performances. One can advance the cause for this result: if IS papers are normalised by the field they publish in, they are doubly disadvantaged in respect both of their publishing in natural sciences (because even if they receive many cites – all else being equal – they tend to be less cited than natural science papers in those journals), or in the social sciences (because they have disproportionate difficulties in publishing in the most prestigious journals – which tend to accrue more citations).[24] Fourth, we use the fractional-counting of cites normalisation, which inversely weighs each citation by the number of references in the citing paper. Under this normalisation, the correlation between citation-based and ABS-based performance completely vanishes. We highlight this as an important result.

Finally, following the multiple indicators tenets, we also provide the performance associated with the journal IF. We stress, however, that IF-based measures have been convincingly proved to be a worse indicator of quality than citations (Seglen, 1997). In this case, nevertheless, the findings are quite similar to those obtained with citations. Overall, the mean IFs of IS publications and citing papers are as high as that of BMS. Interestingly, the standard error of IFs is much higher in IS units than in BMS, which is reflective of more diverse publication behaviour. Again, upon normalisation by field (SC) of publication, IS performance is relatively deflated, while the citing field perspective retains the strong IS performance.

[24] This is suggested by the fact that IS publications have yet less citations for high IF (mainly natural science) journals (1.39 times the IF of journals if IF>5, vs. 1.76 for BMS) yet more citations from low IF journals (4.20 times the IF of journals if IF<0.5, vs. 2.96 for BMS).

In summary, according to the ABS ranks all three BMS units show significantly higher performance than IS units. However, a re-examination of this using various other conventional measures does not provide a clear performance ranking. Raw citation and IF measures put SPRU at the same level as Imperial and LBS, and MIoIR and ISSTI on a par with WBS. The citing-side normalisation inverts the results, putting ISSTI as the best performer. Field normalisation places Imperial and LBS ahead, but without a statistically significant lead over SPRU given the high standard errors. In short, these results show how contrasting, but prima facie equally legitimate, metrics can yield fundamentally different conclusions.[25]

One may speculate on what causes the deflation of IS units' performance when judged by ABS ranking criteria. On the one hand, there is the bias in the ABS list that is generated by the focus on the dominant Business and Management journals – as described in the previous subsection. On the other hand, there may be a more general mechanism, which might reflect the greater difficulty that interdisciplinary papers face in being accepted in mainstream disciplinary fields compared to disciplinary papers of the same quality. This is suggested, yet needs further confirmation, by the high correlation observed (at the aggregated unit level) between the ABS ranks of the 6 organisations, and their field normalised performances, either in terms of field normalised IF (0.920, p=0.009), or, much less significant, for field normalised cites/paper (0.765, □□0.076; compared with a correlation of 0.922 and □□0.009 with raw cites/paper).

These findings also have implications for bibliometric performance measures. Although still an open issue, sophisticated studies strongly suggest that citing-side normalisation provides a more robust measure of citation impact, since it achieves a much more accurate description of the citation context of each individual paper (Zitt and Small, 2008; Zhou and Leydesdorff, 2010). The differences in results we find in this study suggest that further research is needed to investigate whether the

---

[25] It is worth noting that the results are much more stable if we look only into the relative performance of BMS among themselves. In this case LBS and Imperial obtain similar results, with WBS coming third. This observation supports the interpretation that the contrasting results for IS units are due to differences in disciplinary make-up.

conventional field normalisation has been systematically under-estimating interdisciplinary contributions in comparison with the new normalisation method.

The picture that emerges from all the indicators supports the general call for more rigorous ‘plural and conditional’ forms of appraisal mentioned earlier (Stirling, 2008; 2010) as it directly addresses the need explicitly to use ‘partial converging indicators’ in research evaluation (Martin, 1996; Martin and Irvine, 1983). If the ranking is determined more by the choice of indicator than by the content of the research, then the objectiveness of the measures seems questionable as being open to intentional and unintentional bias. Similar warnings about inconsistencies between results and how they depend on the field size and/or classification methods that are used in citation performance normalisation have been repeatedly voiced in the past (Adams et al., 2008; Leydesdorff, 2008; Zitt, 2005; Zhou and Leydesdorff, 2011). This paper confirms quite how problematic the use of single metrics can be.

Yet despite such contrasts, certain specific conclusions in this study transcend the wider ambiguities. These conclusions can be judged to be all the more robust and worthy of confidence because they have been transparently qualified – by declaration of the range of possible interpretations. In this case, no matter how the detailed data is viewed, it is difficult to avoid two conclusions. First that the IS units are more interdisciplinary than the BMS considered here. Second, that the performance of IS units is undervalued in the ABS metrics when compared to a range of citation and IF metrics. In summary, a key robust observation is that while the ABS measure of high performance is likely be to associated with a narrower disciplinary focus (on business, management, finance and economics) it is not necessarily related to high citation performance. This should be of concern, because the latter are generally considered to be a more reliable indicator than journal-based measures (Seglen, 1997).

Table 4. Performance indicators.

	Innovation Studies Units			Business and Management Schools		
	ISSTI	**SPRU**	**MIoIR**	**Imperial**	**WBS**	**LBS**
ABS ranking-based Mean (Std Error)						
Mean ABS rank	2.82 (0.13)	2.65 (0.10)	2.54 (0.10)	3.36 (0.07)	3.01 (0.05)	3.92 (0.05)
% Papers ranked	43%	51%	74%	69%	79%	93%
Citation-based Mean (Std Error)						
Cites/paper	2.69 (0.45)	5.11 (0.59)	3.50 (0.63)	5.30 (0.73)	2.91 (0.23)	5.04 (0.39)
Cites/paper (Journal normalized)	1.99 (0.31)	2.74 (0.36)	2.35 (0.34)	2.69 (0.33)	2.16 (0.16)	2.28 (0.17)
Cites/paper (Field normalized)	1.67 (0.28)	2.79 (0.35)	2.10 (0.43)	3.34 (0.47)	2.11 (0.16)	3.60 (0.28)
Cites/paper (Citing-side normalized)	0.18 (n.a.)	0.12 (n.a.)	0.09 (n.a.)	0.13 (n.a.)	0.07 (n.a.)	0.11 (n.a.)
Impact Factor-based Mean (Std Error)						
Journal IF	2.29 (0.38)	3.14 (0.51)	1.96 (0.34)	2.76 (0.27)	1.65 (0.09)	2.50 (0.09)
Journal IF (Field normalized)	1.17 (0.12)	1.26 (0.11)	0.98 (0.06)	1.46 (0.07)	1.11 (0.03)	1.74 (0.06)
Citing journal IF	3.12 (0.28)	2.45 (0.15)	1.98 (0.11)	2.79 (0.14)	1.79 (0.06)	2.18 (0.05)

Note: The standard deviation of cites/paper with citing-side normalisation is not available because the data on the citing articles was collected in an aggregated form.

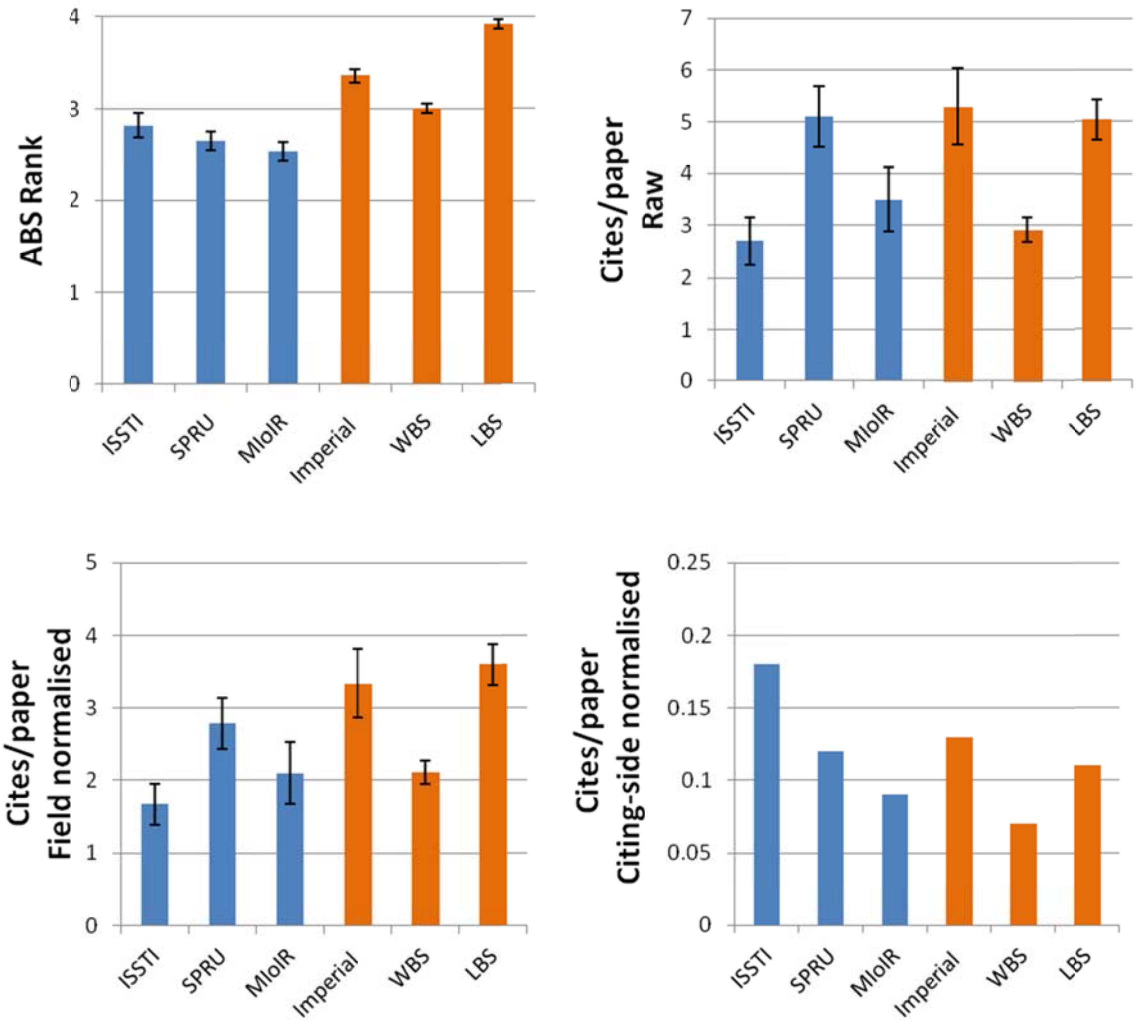


Figure 9. Performance indicators.

5. Discussion: From bias in rankings to suppression of interdisciplinarity

5. 1 Mechanisms of bias amplification

Although the upcoming UK research assessment exercise does not formally rely on rankings, the widespread perception in the field of Business and Management is that the number of publications in top ABS journals will be an accurate predictor of the assessment outcome. As highlighted previously, a number of studies have shown this was the case for the 2008 assessment (Taylor, 2011, pp. 212-14; Kelly et al., 2009; David Storey, personal communication, March 2011).[26]

A number of complementary distorting mechanisms may further amplify the bias highlighted in these results. The first is that the percentage of publications included in the ABS classification is much lower for IS units than for BMS (shown in Figure 10). If each researcher is required to submit up to four articles, then the *average* researcher in an IS unit, evaluated under a Business and Management panel, may need to publish eight articles to ensure four fall within the ABS remit. The alternative, and arguably much more likely scenario, is that she will change her publication patterns, which will require a shift away from IDR towards a more disciplinary focus.

The second mechanism of bias amplification is the exponential scale that the assessment exercise uses to reward perceived quality. In terms of (financial) resource allocation this means that Rank 1 articles have a multiplier of 0 (i.e. they are ignored), Rank 2 articles have a multiplier of 1, Rank 3 article a multiplier of 3, and Rank 4 articles a multiplier of 9. Using such a quasi-exponential scale, the ~50% performance difference in ABS ranks between MIoIR and LBS, becomes a ~120% difference in the resources received by the units (see Figure 10). Given that this process has been cumulative over subsequent assessment exercises, there would be a strong incentive to shift publication, and therefore research patterns.

[26] Business and Management is one of the Units of Analysis where there is also a high correlation (r=0.719) between the number of citations per paper and the 2001 RAE score (Mahdi et al., 2008, Table 3).

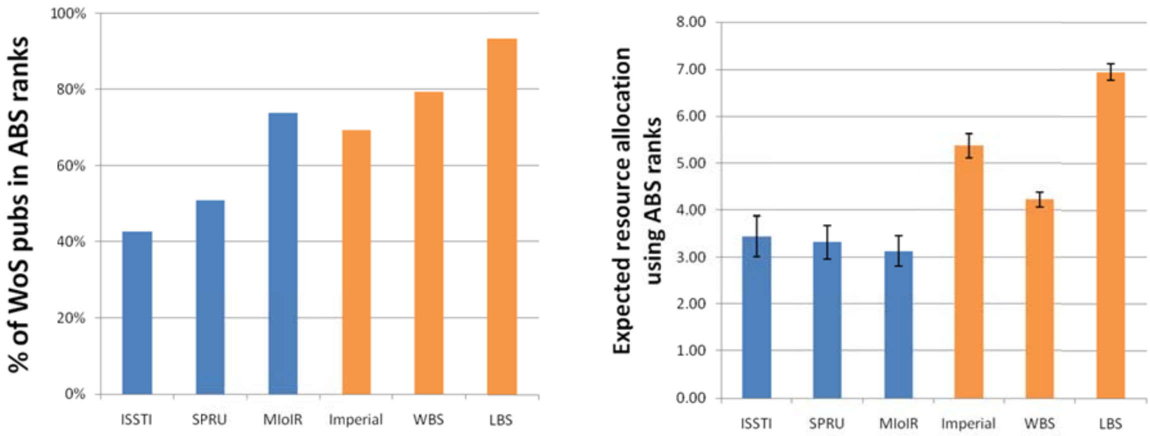


Figure 10. Left: percentage of publications by unit published in journals listed in the ABS rankings. Right: Expected outcome of resource allocation derived from an ABS ranking based assessment exercise. Compare with Figure 9.

5.2 Explanations for differences between IS units and BMS

In principle both IS and BMS might be expected to be equally interdisciplinary, as both deal with complex social issues. Business and management are not traditional scientific units and are to some degree multidisciplinary, with their research outputs published in economics, finance, and to a lesser degree psychology and operations research as well as business and management (though with little cross-linking). However, this is still much less interdisciplinary than IS units. Why are the knowledge base and audiences of IS units so much more diverse than BMS?

Clausen et al.'s (this issue) survey of the drivers and barriers experienced by IS, STS and Entrepreneurship units may help answer this question, as it suggests that a stark contrast exists among them. In IS and STS the main reasons given for units being established were 'need for cross-disciplinarity' and 'new academic knowledge'. By contrast, entrepreneurship units identify the main driver behind their creation as 'new academic teaching program' (Clausen et al., this issue, p. xx). While IS and STS units reported that they had 'met with strong scepticism' from academic departments and 'would have not developed without support from policy-making organizations', entrepreneurship units experienced support from other departments and their university's leadership. Given that entrepreneurship is a field largely embedded in BMS, this attitude is not irrelevant to the wider BMS area.

If IS and STS centres developed to carry out research in response to policy and social questions, then it is no surprise they are often driven simultaneously to engage with diverse disciplines and stakeholders. By contrast, BMS (like entrepreneurship units) developed as centres of professional training and therefore only require a stock of scholars teaching in required fields, without particular need for research integration. Senior BMS scholars have already raised this point and expressed concern about the resulting inability to address important social and managerial issues to a depth appropriate to their full societal complexity (Minzberg, 2000).

The problem-driven nature of nominal IS units may also be the reason why they are seldom 'purely' IS, as defined by the core literature studies of IS. This putative disciplinary core would fall in the middle left area of Figure 5, between the upper management and lower economics poles (see Table 6 in Fagerberg et al., this issue; Fagerberg and Verspagen, 2009). Instead IS units publish over a variety of IS-related fields, including the two management and economics poles, with an important presence in STS as shown in Figure 5, right middle area (see Table 4 in Martin et al.'s core STS literature study, this issue), as well as specifically oriented areas such as health, energy or environment. This is perhaps an important difference between the core literature studies of IS and STS based on handbook chapters (Fagerberg et al., this issue; Martin et al., this issue) and the results obtained here based on the units' journal publications. The former approach emphasises the theoretical foundations of IS and STS and their division, as shown for example, in the relatively small degree of cross-citations (Figure 2, Bhupatiraju et al., this issue; Leydesdorff, 2007b). A focus in the organisational unit, on the one hand, reveals their disparate intellectual debts and allegiances within the social sciences, and on the other hand, their close engagement with practitioners in the areas of energy, biomedical research, health services and environment.

5.3 How general are the findings?

Business and Management is perceived by some analysts as a rather paradoxical case in relation to other disciplines. Given that it is an applied field one might expect to see a highly diverse knowledge base and a plurality of approaches. Instead one finds BMS scholars competing to get published in a small number of very similar journals. This raises a question about the extent to which these findings on bias against IDR in BMS and IS are generalisable or only apply to these fields.

Research on journal ranking in economics (Oswald, 2007) suggests that the findings may at least applicable to related social sciences. In many natural sciences the norm is to use indicators of journal quality such as Thomson-Reuter's Impact Factor (IF) rather than rankings. Could the use of IFs discriminate against IDR? If IDR articles are less likely to be accepted in high IF journal then this might be the case. According to the National Academies (2004, p. 139):

> 'With the exception of a few leading general journals—such as *Science*, *Nature*, and the *Proceedings of the National Academy of Sciences*—the prestigious outlets for research scholars tend to be the high-impact, single discipline journals published by professional societies. Although the number of interdisciplinary journals is increasing, few have prestige and impact equivalent to those of single-discipline journals (…). Interdisciplinary researchers may find some recognition by publishing in single-discipline journals (…), but the truly integrated portion of their research may not be clear too much of the audience or be noticed by peers who do not read those journals.'

The correlation observed in Table 4 between the results of ABS ranks and the mean journal IF (with field normalisation) in comparison to the much lower correlation in citation/paper may be interpreted as supporting the hypothesis of an IF-based bias.

Similarly, although not all the Units of Assessment of the UK's assessment exercise were perceived as disadvantaging IDR departments, the possibility of such a bias has been recurrently raised (Boddington and Coe, 1999) and it remains an open issue across the whole exercise. As Martin and Whitley, (2010, p. 64) noted:

> '…the UK has an essentially discipline- based assessment system for a world in which government policies are trying to encourage more user-focused and often interdisciplinary research. Those who have gone down the user-influenced route frequently conclude that they have ended up being penalized in the RAE process. (…) in practice the heavy reliance on peer review and the composition of RAE panels mean that discipline-focused research invariably tends to be regarded as higher quality.'

In summary, although in other disciplines the bias against IDR resulting from explicit or implicit perceptions of journal quality may not be as manifest or pronounced as in business and management, there are *a priori* grounds to believe that they may nevertheless exist, and generate bias in evaluation. However, further research is needed is order to test these speculations, given the marked differences in the social institutionalisation between (and sometimes within) the various sciences and branches of engineering (Whitley, 2000).

5.4 The consequences of bias

We have so far argued that ABS rankings disadvantage IDR and their use in evaluation would result in a bias against IDR that would in turn result in significant financial implications under current REF procedures.[27] But what would the consequences be for society of such suppression of IDR? A major intent behind both assessment and rankings is to foster competition (which is assumed to be desirable) by providing fair, transparent, accountable rules through which this competition can be managed (Gläser and Laudel, 2007; pp. 108-109).[28] However, many analysts have warned against the 'inadvertent' but 'potentially destructive' consequences of bibliometric rankings for the science system (Weingart, 2005, p. 130; Roessner, 2000) and the capture of evaluation by disciplinary elites (Martin and Whitley, 2010, pp. 64-67).

A first type of consequence may be the creation or reinforcement of dis-incentives for researchers to engage in IDR. Among US sociologists and linguists, for example, Leahey (2007) found that more interdisciplinary (less specialised) researchers tend to earn less.[29] Van Rijnsoever and Hessels (2011) reported that researchers more engaged in disciplinary collaboration benefit more from promotion than those

[27] Notice, that the argument so far has been based on a relatively naive understanding of indicators as simple measurement tools that can have unintended consequences if they are biased. A more politically-nuanced view on the role of indicators would also consider the performative power of journal rankings. Namely, that rather than simply setting standards to help measure quality, they reflect deliberate aims to establish what that quality should be. The ABS rankings guide, for example, state that the function of the rankings is to '[p]rovide an indication of where best to publish' (ABS, 2010, p. 2).

[28] This is based on the assumption that academics are in competition with each other, either as individuals or departments, rather than part of a shared, international, cumulative intellectual endeavour.

[29] Interestingly, this partly explains women's earning differences, given that women tend to be more interdisciplinary (Leahey, 2007; Rhoten and Pfirman, 2007; Van Rijnsoever and Hessels, in press).

engaged in IDR. As noted previously, Lee and Harley have repeatedly argued that bias in the RAEs has focused recruitment in UK economics department towards mainstream economists and against heterodox economists (Harley and Lee, 1997; Lee and Harley, 1998; Lee, 2007). This push towards mainstream is also suggested by an analysis of the economics-related submissions in the new Italian research assessment exercise. Corsi et al. (in press) showed that the percentage of papers in heterodox economics and economic history was much lower in the assessment exercise than in a general economics database such as EconLit – suggesting a selection bias towards the more narrowly disciplinary specialties within economics, such as econometrics and finance.

Second, the bias may generate a process of intellectual inbreeding, where the efforts to increase the quality of research end up creating a self-reinforcing narrowing down of conceptions of quality, that effect the content of research (Mirowski, 2011). The definition of quality may move from the definitions imposed by disciplinary elites, even further, to definitions imposed by the structure of the audit process itself. A number of prominent management scholars have already expressed concerns that this is happening and that some parts of management research are becoming an irrelevant game structured by the academic job market and school rankings rather than by research ingenuity or concern about real-world issues (Minzberg, 2000; Willmott, 2011a, 2011b; Tourish, in press; Alvesson and Sandberg, 2011).[30]

Thirdly, since socially relevant research almost inevitably involves navigating across or between multiple disciplines, a reduction in IDR may shift research away from complex social questions. For example, a study by Goodall (2008), reported that 'over the past few decades, only 9 articles on global warming or climate change have appeared in the top 30 business and management titles', out of ~31,000 papers in total (p. 417). By contrast, more than 2,000 publications had appeared on the topic in 'journals that are peripheral to the main social science disciplines' (p. 415). Goodall attributes this dearth of publication on climate change in top journals to their valuation of theory over practical issues, political bias and associated career incentives.

[30] The ample interest and support generated by these concerns was shown in the special Symposium within the large conference EGOS 2011 organised by Alvesson and Sandberg of, to debate the issue with the editors of the *Journal of Management Studies* and *Organization Studies*. Or by the recent publication of a special issue of *Organization* (see Willmott, 2011a and ensuing papers).

Patenaude (2011) recently confirmed Goodall's findings and showed further that this lack of research interest may spill over into teaching, as MBA curricula also display limited interest in climate change (p. 260). Patenaude partly attributes such biases to corporate values and beliefs as well as to existing academic incentives and communication channels.

Lastly, this sort of bias may generate a reduction of the cognitive diversity of the entire science system. Diversity in the science system is important from an ecological or evolutionary perspective because: (i) notions of quality are dynamic (what is now marginal may quickly become highly significant); (ii) diversity helps prevent paradigmatic lock-in; (iii) it fosters the appearance of new types of knowledge (Salter and Molas-Gallart, 2002; Stirling, 1998, pp. 6-36; Stirling, 2007). All in all, there are good reasons to be concerned about the findings in this paper.

6. Conclusions

This empirical investigation has responded to wider concerns that have been raised in science policy debates about evaluation of IDR. It has illustrated a more rigorously 'plural and conditional' approach to research evaluation, making use of 'partial convergent indicators'. By using a suite of innovative maps and metrics the paper has demonstrated that IS units are more interdisciplinary than leading BMS under various perspectives. In particular, it has shown that ABS rankings commonly used in BMS have a disciplinary bias which translates very directly into a low assessment of interdisciplinary units' performance. Furthermore, we have shown that this low assessment is not warranted by citation-based performance indicators, which are generally considered to offer more robust measures of quality. In this way, the study suggests that the use of ABS rankings serves systematically to disadvantage this form of IDR in the BMS context. This finding needs to be tested in a wider setting, in particular in the natural sciences, to establish its robustness and the extent to which the problems identified here are generalisable. The main caveats are that citation data was harvested for only for a short period after publication, without fixed citation windows; and that we used conventional, mean-based performance measures, instead of more advanced, distribution-based measures (Leydesdorff and Bormann, in press).

These quantitative findings support the well-established picture, evident from qualitative investigations in science studies and in policy (National Academies, 2004), that criteria of excellence in academia are based on disciplinary standards, and that this hinders interdisciplinary endeavours in general, and policy and socially relevant research in particular (Travis and Collins, 1991; Langfeldt, 2006). Few previous quantitative investigations have investigated bias against IDR (Porter and Rossini, 1985; Rinia et al., 2001a). Consequently, this study constitutes one of the first and most thorough explorations of whether one of the criteria most widely used in research assessment, namely journal rankings, may result in a bias. We find strong evidence that it does.

In recent decades criteria of quality have become institutionalised in the form of rankings that can have major (often negative) reputational and funding implications. The use of this kind of simple ranking procedure is predicated on the implicit or explicit assumption that the resulting ranks constitute *objective assessments* that can be treated as robust proxies for academic excellence. The empirical results obtained here challenge such general claims to objectivity. They suggest instead that the resulting picture presents a rather narrow and idiosyncratic view of excellence. To the extent that ABS style rankings are increasingly used to evaluate individual and organisational research performance, it does seem possible to identify a prima facie hypothesis that this practice exercises a suppressive effect on IDR.

In summary, this paper shows that when journal rankings are used to help determine assignments of esteem and resources, they can suppress forms of interdisciplinarity that are otherwise widely acknowledged to be academically and socially useful. Important implications arise, both for research evaluation in the specific fields in question, as well as for wider investigations to inform the more general governance of science and technology using metrics to capture multidimensional qualities that cannot be reduced to a single indicator.

Supplementary materials

The full suite of maps (diversity, coherence and intermediation) for each unit and perspective (publications, references and cites) is available at: http://www.interdisciplinaryscience.net/maps

Also, in powerpoint format:

Disciplinary Diversity:

http://www.sussex.ac.uk/Users/ir28/IDR/Disciplinary_Diversity.pptx

Disciplinary Coherence:

http://www.sussex.ac.uk/Users/ir28/IDR/Disciplinary_Coherence.pptx

Intermediation:

http://www.sussex.ac.uk/Users/ir28/IDR/Intermediation.pptx

ABS Ranking Diversity:

http://www.sussex.ac.uk/Users/ir28/IDR/ABS_Ranking_Diversity.pptx

Comparison of Units' Performances:

http://www.sussex.ac.uk/Users/ir28/IDR/Performance_Comparison.pptx

Acknowledgments

We thank colleagues from various British BMS schools for feedback. We are grateful to Jan Fagerberg for detailed comments and Diego Chavarro for designing the website www.interdisciplinaryscience.net. IR and AO were funded by the US NSF (Award no. 0830207, http://idr.gatech.edu/) and the EU FP7 project FRIDA (grant 225546, http://www.fridaproject.eu). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the funders.

References

ABRC (1990). Peer Review: a Report to the Advisory Board for the Research Councils (ABRC) from the Working Group on Peer Review (Boden Report). ABRC, London.

ABS (2010) *Academic Journal Quality Guide. Version 4*. Association of Business Schools (downloaded on 1st December 2010 from http://www.the-abs.org.uk/?id=257)

Adams, J, Jackson, L. and Marshall, S. (2007) *Bibliometric analysis of interdisciplinary research. Report to the Higher Education Funding Council for England.* Evidence, Leeds.

Adams, J., Gurney, K. and Jackson, L. (2008) Calibrating the zoom – a test of Zitt's hypothesis. *Scientometrics* 75(1), 81-95.

Alvesson, M. and Sandberg J. (2011) Has management studies lost its way? Ideas for more imaginative and innovative research. Paper presented in the 2011 EGOS (European Group for Organizational Studies) Conference. Stockholm, July 2011.

Balakrishnan, A., Pena, V. and Lal, B. (2011) Measuring the Interdisciplinarity of a Federal R&D Program Using Bibliometrics and Science Overlay Maps. Paper presented in the 2011 *Atlanta Science and Innovation Policy Conference.* Atlanta, September 2011 (downloaded on 28 September 2011from http://conferences.library.gatech.edu/acsip/index.php/acsip/ATLC11/paper/view/596/481).

Barry, A. , Born, G. and Weszkalnys, G. (2008) Logics of interdisciplinarity, *Economy and Society* 37, 1, 20-49.

Bhupatiraju, S., Nomaler, O. Triulzi, G. and Verspagen, B. (this issue) Knowledge flows – Analyzing the core literature of Innovation, Entrepreneurship and Science and Technology Studies. *Research Policy.* This issue.

Boddington, A. and Coe, T. (1999) *Interdisciplinary Research and the Research Assessment Exercise.* Evaluation Associates Ltd., Report for the UK Higher Education Funding Bodies (downloaded on 25 April 2011 from 195.194.167.103/Pubs/1_99/1_99.doc).

Boix Mansilla, V. (2006) Assessing expert interdisciplinary work at the frontier: an empirical exploration. *Research Evaluation* 15(1), 17-29.

Bordons, M., Morillo, F. and Gomez, I. (2004) Analysis of cross-disciplinary research through bibliometric tools. In: Moed, H.F., Glänzel, W. and Schmoch, U. (Eds), *Handbook of quantitative science and technology research 10031003,* Kluwer, Dordrecht, 437-456.

Börner, K., Contractor, N., Falk-Krzesinski, H.J., Fiore, S.M., Hall, K.L., Keytonn, J., Spring, B., Stokols, D., Trochim, W. and Uzzi, B. (2010) A multi-level systems perspective for the science of team science. *Science Translational Medicine* 2, cm24.

Braun, T. and Schubert, A. (2003) A quantitative view on the coming of age of Interdisciplinarity in the sciences, 1980-1999. *Scientometrics* 58, 183-189.

Bruce, A., Lyall, C., Tait, J. and Williams, R. (2004). Interdisciplinary integration in Europe: the case of the Fifth Framework programme. *Futures* 36, 457-470.

Campbell, D. T. (1969) Ethnocentrism of disciplines and the fish-scale model of omniscience. *Interdisciplinary relations in the social sciences*. Sherif and Sherif, Chicago, Aldine, 328-348.

Carayol, N. and Thi, T.U.N. (2005) Why do academic scientists engage in interdisciplinary research?, *Research Evaluation* 14(1), 70-79.

Clausen, T.H., Fagerberg, J., Gulbrandsen, M. (this issue) When invisible colleges become visible: the emergence of new research units within innovation studies, entrepreneurship studies and STS. *Research Policy.* This issue.

Cech, T.R. and Rubin, G.M. (2004) Nurturing interdisciplinary research. *Nature Structural and Molecular Cell Biology* 1166-1169.

Collingridge, D. (1982) *Critical Decision Making: a new theory of social choice*, Pinter, London.

Collini, S. (2008) *Absent Minds: intellectuals in Britain*. Oxford University Press, Oxford.

Corsi, M., D'Ippoliti, C., and Lucidi, F. (In press) Pluralism at Risk? Heterodox Economic Approaches and the Evaluation of Economic Research in Italy. *The American Journal of Economics and Sociology.*
DOI: 10.1111/j.1536-7150.2010.00754.x

Cummings, J.N. and Kiesler, S. (2005) Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science* 35 (5), 733-722.

Cummings, J.N. and Kiesler, S. (2007) Coordination costs and project outcomes in multi-university collaborations. *Research Policy* 36, 1620-1634.

de Nooy, W., Mrvar, A. and Batagelj, (2005) *Exploratory social network analysis with Pajek.* Cambridge University Press, New York.

Donovan, C. (2007) Introduction: Future pathways for science policy and research assessment: metrics vs peer review, quality vs impact. *Science and Public Policy* 34(8), 538–542.

ERC (2010) European Research Council (ERC) Grant Schemes. Guide for Applicants for the Starting Grant 2011 Call (downloaded on 29 September 2011 from ftp://ftp.cordis.europa.eu/pub/fp7/docs/calls/ideas/l-gfastg-201101_en.pdf (downloaded on 31-12-2010).

EURAB (2004) *Interdisciplinarity in Research*, European Union Research Advisory Board (EURAB) (downloaded on 25 April 2010 from http://ec.europa.eu/research/eurab/pdf/eurab_04_009_interdisciplinarity_research_final.pdf).

Fagerberg, J. and Verspagen, B. (2009) Innovation studies -The emerging structure of a new scientific field. *Research Policy* 38, 218–233.

Fagerberg, J. Fosaas, M and Sapprasert, K. (this issue) Innovation: Exploring the knowledge base. *Research Policy*. This issue.

Freeman, L.C. (1977). A set of measures of centrality based on betweenness. *Sociometry* 40(1), 35–41.

Fuchsman, K. (2007) Disciplinary realities and interdisciplinary prospects, *The global spiral* (downloaded on 24 April 2011from http://www.metanexus.net/Magazine/tabid/68/id/10192/Default.aspx).

Funtowicz, S. and Ravetz, J. (1990) *Uncertainty and Quality in Science for Policy*. Kluwer, Amsterdam.

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

Gläser, J. and Laudel, L. (2007) The social construction of bibliometric evaluations. In: Whitley, R. and Gläser, J. (Eds). *The Changing Governance of the Sciences*. *The Advent of Research Evaluation Systems.* Springer, Dordrecht, 101–123.

Goodall, A.H. (2008) Why Have the Leading Journals in Management (and Other Social Sciences) Failed to Respond to Climate Change? *Journal of Management Inquiry* 17(4), 408-420.

Harley, S. and Lee, F. S. (1997) Research Selectivity, Managerialism, and the Academic Labour Process: The Future of Nonmainstream Economics in U.K. Universities, *Human Relations* 50, 1425–60.

Hamilton, K.S., Narin, F. and Olivastro, D. (2005) *Using Bibiliometrics to Measure Multidisciplinarity.* ipIQ, Westmont, NJ.

Hollingsworth, R. and Hollingsworth, E.J. (2000) Major discoveries and biomedical research organizations: perspectives on interdisciplinarity, nurturing leadership, and integrated structure and cultures. In: Weingart, P. and N. Stehr, (Eds), *Practising Interdisciplinarity*. University of Toronto Press, Toronto, pp. 215-244.

Heinze, T., Shapira, P., Rogers, J. D. and Senker, J. M. (2009). Organizational and institutional influences on creativity in scientific research. *Research Policy* 38, 610-623.

Hemlin, S., Allwood, C. M. and Martin, B. R., (Eds). (2004). *Creative knowledge environments: the influences on creativity in research and innovation*. Edward Elgar, Cheltenham, UK.

Hessels, L.K. (2010) *Science and the struggle for relevance. PhD Dissertation.* University of Utrecht, Utrecht (downloaded on 30 April 2011 from http://rathenau.academia.edu/LaurensHessels/Papers/338879/Science_and_the_Struggle_for_Relevance).

Huutoniemi, K. (2010) Evaluating interdisciplinary research. In: Frodeman, R., Klein, J.T., and Mitcham, C. (Eds). *Oxford Handbook of Interdisciplinarity.* Oxford University Press, Oxford, pp. 309-320.

Huutoniemi, K., Klein, J. T., Bruun, H. and Hukkinen, J. (2010). Analyzing interdisciplinarity: Typology and indicators. *Research Policy* 39(1), 79-88.

Jacobs, J. A. and Frickel, S. (2009). Interdisciplinarity: a critical assessment. *Annual Review of Sociology* 35, 43-65.

Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy* 27(1), 23-36.

Katz, J.S. and Martin, B.R. (1997) What is research collaboration? *Research Policy* 26, 1-18.

Kelly, A., Morris, H. and Harvey, C. (2009) *Modelling the Outcome of the UK Business and Management Studies RAE 2008 with reference to the ABS Journal Quality Guide* (downloaded on 29 April 2011 from http://www.the-abs.org.uk/files/RAE2008_ABS2009_final.pdf).

Kiss, I. Z., Broom, M., Craze, P. G. and Rafols, I. (2010). Can epidemic models describe the diffusion of topics across disciplines? *Journal of Informetrics* 4(1), 74-82.

Langfeldt, L. (2006) The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessments, *Research Evaluation* 15(1), 31-41.

Laudel, G. and Origgi, G. (2006) Introduction to a special issue on the assessment of interdisciplinary research, *Research Evaluation* 15(1), 2-4.

Larivière, V. and Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact, *Journal of the American Society for Information Science and Technology* 61(1), 126-31.

Leahey, E. (2007) Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings, *American Sociological Review* 72, 533-561.

Lee, C. (2006). Perspective: peer review of interdisciplinary scientific papers. *Nature Online* (downloaded on 28 April 2011 from http://blogs.nature.com/nature/peerreview/debate/comments/2006/06/perspective_peer_review_of_int.html, .

Lee, F.S. (2007) The Research Assessment Exercise, the State and the Dominance of Mainstream Economics in British Universities, *Cambridge Journal of Economics* 31, 309–25.

Lee, F.S. and Harley, S. (1998) Peer Review, the Research Assessment Exercise and the Demise of Non-Mainstream Economics, *Capital and Class* 66, 23–51.

Levitt, J. and Thelwall, M. (2008) Is multidisciplinary research more highly cited? A macrolevel study, *Journal of the American Society for Information Science and Technology* 59(12), 1973-84.

Leydesdorff, L. (2007a). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the American Soc. for Information Science and Technology* 58(9), 1303-1319.

Leydesdorff, L. (2007b). Mapping Interdisciplinarity at the Interfaces between the *Science Citation Index* and the *Social Science Citation Index*. *Scientometrics 71*(3), 391-405.

Leydesdorff, L. (2008). *Caveats* for the Use of Citation Indicators in Research and Journal Evaluation. *Journal of the American Society for Information Science and Technology 59*(2), 278-287.

Leydesdorff, L., and Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology* 60(2), 348–362.

Leydesdorff, L., and Opthof, T. (2010). Normalization at the field level: fractional counting of citations. *Journal of Informetrics* 4(4), 644-646.

Leydesdorff, L., and Opthof, T. (2011). Remaining problems with the "New Crown Indicator" (MNCS) of the CWTS. *Journal of Informetrics 5*(1), 224-225.

Leydesdorff, L. and Rafols, I. (2011a). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics* 5(1), 87-100.

Leydesdorff, L. and Rafols, I. (2011b). The local emergence and global diffusion of research technologies: an exploration of patterns of network formation. *Journal of the American Society for Information Science and Technology* 62(5), 846–860.

Leydesdorff, L., and Bornmann, L. (in press). Integrated Impact Indicators (I3) compared with Impact Factors (IFs): An alternative design with policy implications. *Journal of the American Society for Information Science and Technology*.

Leydesdorff, L. and Rafols, I. (in preparation) Interactive and Online Journal Maps of Science Overlayed to the Global Map of Science using Web-of-Science Data.

Liu, Y., Rafols, I. and Rousseau, R. (in press). A framework for knowledge integration and diffusion. *Journal of Documentation*.

Llerena, P. & Meyer-Krahmer, F. (2004) Interdisciplinary research and the organization of the university: general challenges and a case study. In: Geuna, A., Salter, A.J. & Steinmueller, W.E. (Eds). *Science and Innovation. Rethinking the Rationales for Funding and Governance.* Edward Elgar, Cheltenham, UK, pp. 69-88.

Lowe, P. and Phillipson, J. (2006). Reflexive interdisciplinary research: the making of a research programme on the rural economy and land use. *Journal of Agricultural Economics* 57(2), 165-184.

Lowe, P. and Phillipson, J. (2009). Barriers to research collaboration across disciplines: scientific paradigms and institutional practices. *Environment and Planning A* 41, 1171-1184.

Mahdi, S., D'Este, P., Neely, A. (2008) Citation Counts: Are They Good Predictors of RAE Scores? A bibliometric analysis of RAE 2001. Cranfield University (downloaded on 26 April 2011 from https://dspace.lib.cranfield.ac.uk/handle/1826/2248).

Mallard, G., Lamont, M. & Guetzkow, J. (2009) Fairness as Appropriateness : Negotiating Epistemological Differences in Peer Review. *Science, Technology and Human Values* 34(5), 573-606.

Martin, B. R. (1996) The use of multiple indicators in the assessment of basic research. *Scientometrics* 36(3), 343-362.

Martin, B. R. and Irvine, J. (1983) Assessing basic research. Some partial indicators of scientific progress in radio astronomy, *Research Policy* 12, 61-90.

Martin, B. R. and Whitley, R. (2010). The UK Research Assessment Exercise. A case of regulatory capture?. In: *Reconfiguring knowledge production: changing authority relations and their consequences for the sciences.* Whitley, R., Gläser, J. and Engwall, L. (Eds). Oxford University Press, Oxford, 51-80.

Martin, B. R., Nightingale, P. and Yegros-Yegros, A. (this issue) Science and Technology Studies: Exploring the knowledge base. *Research Policy*. This issue.

Metzger, N. and Zare, N. (1999). Interdisciplinary research: From belief to reality. *Science* 283(5402), 642-643.

Minzberg, H., (2000) *Managers not MBAs*, Oxford: Oxford University Press.

Mirowski, P, (2011), *Science-mart*, Harvard University Press, Cambridge, MA.

Molas-Gallart, J. and A Salter (2002) Diversity and excellence: considerations on research policy, *IPTS Report* 66(5), 5-13.

Muller, B. (2009), Does Interdisciplinarity Lead to Higher Employment Growth of Academic Spinoffs?, *ZEW Discussion Paper* No. 09-087, Mannheim. Available at ftp://ftp.zew.de/pub/zew-docs/dp/dp09087.pdf. Accesses on 26 April 2011.

Narin, F., and Hamilton, K.S. (1996) Bibliometric performance measures. *Scientometrics,* 36 (3), 293-310

National Academies (2004) *Facilitating Interdisciplinary Research.* Committee On Science, Engineering, And Public Policy (Cosepup) Committee on Facilitating Interdisciplinary Research. National Academies Press, Washington, DC.

Nightingale, P. and Scott, A. (2007) Peer review and the relevance gap: ten suggestions for policy-makers. *Science and Public Policy* 34(8), 543–553.

Oswald, (2007) An examination of the reliability of prestigious scholarly journals: evidence and implications for decision-makers, *Economica* 74, 21–31.

Page, S.E. (2007) *The difference. How the power of diversity creates better groups, firms, schools, and societies*. Princeton University Press, Princenton, NJ.

Patenaude, G. (2011) Climate change diffusion: While the world tips, business schools lag. *Global Environmental Change* 21, 259-271.

Porter, A.L., and Rossini, F., (1985) Peer-review of interdisciplinary research proposals. *Science, Technology, and Human Values* 10, 33–38.

Porter, A. L., Roessner, J. D., Cohen, A. S. and Perreault, M. (2006) Interdiscipinary research: meaning, metrics and nurture. *Research Evaluation* 15, 187-196.

Porter, A. L., Cohen, A. S., Roessner, J. D. and Perreault, M. (2007) Measuring researcher interdisciplinarity, *Scientometrics* 72(1), 117-47.

Porter, A.L. and Rafols, I. (2009) Is Science Becoming more Interdisciplinary? Measuring and Mapping Six Research Fields over Time, *Scientometrics* 81(3), 719-45.

Price, D. de S. (1984) The science/technology relationship, the craft of experimental science, and policy for the improvement of high technology innovation. *Research Policy* 13, 3-20.

Rafols, I 2007. Strategies for knowledge acquisition in bionanotechnology: Why are interdisciplinary practices less widespread than expected?. *Innovation: the European Journal of Social Science Research* 20(4), 395-412.

Rafols, I. and Leydesdorff, L. (2009) Content-based and Algorithmic Classifications of Journals: Perspectives on the Dynamics of Scientific Communication and Indexer Effects. *Journal of the American Society for Information Science and Technology* 60(9), 1823-1835.

Rafols, I. and Meyer, M. (2010) Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience, *Scientometrics*, 82(2), 263-87.

Rafols, I, Porter, A. L. and Leydesdorff, L. (2010) Science Overlay Maps: A New Tool for Research Policy and Library Management, *Journal of the American Society for Information Science and Technology* 61(9), 871–1887.

Rinia, E.J., Van Leeuwen, T.N., Van Buren, H.G., and Van Raan, A.F.J. (2001a) Influence of interdisciplinarity on peer-review and bibliometric evaluations in physics research, *Research Policy* 30(3), 357-61.

Rinia, E.J., Van Leeuwen, T.N., and Van Raan, A.F.J. (2001b) Impact measures of interdisciplinary research in physics. *Scientometrics* 53(2), 241-248.

Rhoten, D and A Parker (2006) Risks and rewards of interdisciplinary research path, *Science*, 306, 2046.

Rhoten, D. and Pfirman, S. (2007) Women in interdisciplinary science: Exploring preferences and consequences. *Research Policy* 36 56–75

Rhoten, D., O'Connor, E., Hackett, E.J. (2009) Originality, Disciplinarity and Interdisciplinarity. The Act of Collaborative Creation and the Art of Integrative Creativity. *Thesis Eleven* 96; 83-108.

Roessner, D. (2000) Quantitative and qualitative methods and measures in the evaluation of research. *Research Evaluation* 8(2), 125–132,

Sanz-Menéndez, L., Bordons and M. Zulueta, M.A. (2001) Interdisciplinarity as a multidimensional concept: its measure in three different research areas. *Research Evaluation* 10(1), 47-58.

Seglen, P.O. (1997) Why the impact factor of journals should not be used for evaluating research. *BMJ* 314, 498–502.

Shinn, T., Joerges, B. (2002), The transverse science and technology culture: dynamics and roles of research-technologies, *Social Science Information* 41 (2), 207-251.

Small, H. and Sweeney (1985) Clustering the science citation index using co-citations. I. A comparison of methods. *Scientometrics* 7 (3-6), 391-409.

Stirling, A (1998). On the economics and analysis of diversity, SPRU Electronic Working Papers, 28 (downloaded on 29 September 2011 from http://www.sussex.ac.uk/Units/spru/publications/imprint/sewps/sewp28/sewp28.pdf).

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of The Royal Society Interface* 4(15), 707-719.

Stirling, A. (2008) Opening Up and Closing Down: power, participation and pluralism in the social appraisal of technology, *Science Technology and Human Values* 33(2), 262-294.

Stirling, A. (2010) Keep it Complex, *Nature* 468, 1029–1031.

Stokols, D., Hall, K.L, Taylor, B.K., Moser, R.P. (2008) The Science of Team Science. Overview of the Field and Introduction to the Supplement. *American Journal of Preventive Medicine* 35, S77-S89.

Taylor, J. (2011) The Assessment of Research Quality in UK. Universities: Peer Review or Metrics? *British Journal of Management* 22, 202-217.

Travis, G.D.L. and Collins, H.M. (1991) New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System. *Science Technology Human Values* 16, 322-341.

Tourish, D. (In press) Leading questions: Journal rankings, academic freedom and performativity. What is, or should be, the future of Leadership? *Leadership.*

Van Eck, N.J. and Waltman, L. (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 523–538.

Van Rijnsoever, F.J. and Hessels, L.K, (2011) Factors associated with disciplinary and interdisciplinary research collaboration, *Research Policy* 40(3), 463-472.

Wagner, C.S., Roessner, J.D., Bobb, K., Klein, J.T. Boyack, K.W., Keyton, J., Rafols, I. and Börner, K. (2011) Approaches to Understanding and Measuring Interdisciplinary Scientific Research (IDR): A Review of the Literature, *Journal of Informetrics* 5(1), 14-26.

Walsh, V. (2010) Innovation studies: The emergence and evolution of a new field of knowledge. Paper for EXPLORE Workshop, Lund 7-8 December 2010.

Willmott, H. (2011a) Journal list fetishism and the perversion of scholarship: reactivity and the ABS list. *Organization* 18, 429-442.

Willmott, H. (2011b) Listing perilously. *Organization* 18, 447-448.

Weinberg, A.M. (1963) Criteria for scientific choice, *Minerva* 1, 159-171.

Weingart, P. (2005) Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62(1), 117-131.

Youtie, J, Iacopetta, M. and Graham, S. (2008) Assessing the nature of nanotechnology: can we uncover an emerging general purpose technology? *Journal of Technology Transfer* 33, 315–329.

Yegros-Yegros, A., Amat, C. B., D'Este, P., Porter, A. L. and Rafols, I. (2010). Does interdisciplinary research lead to higher scientific impact? STI Indicators Conference, Leiden.

Zhou, P., & Leydesdorff, L. (2011) Fractional counting of citations in research evaluation: A cross- and interdisciplinary assessment of the Tsinghua University in Beijing. *Journal of Informetrics 5*, (in press); doi:10.1016/j.joi.2011.1001.1010.

Zitt, M. (in press) Behind citing-side normalization of citations: some properties of the journal impact factor. *Scientometrics.* DOI 10.1007/s11192-011-0441-7.

Zitt, M., Ramanana-Rahary, S. and Bassecoulard E. (2005) Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics* 63 (2), 373-401.

Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the American Society for Information Science and Technology* 59, 1856–1860.