How rankings can suppress interdisciplinarity.  
The case of innovation studies and business and management

Ismael Rafols¹,², Loet Leydesdorff³, Alice O’Hare¹, Paul Nightingale¹ and Andy Stirling¹

¹i.rafols@sussex.ac.uk (Corresponding author)
SPRU — Science and Technology Policy Research, University of Sussex, Brighton, BN1 9QE (England)

²Technology Policy and Assessment Center, Georgia Institute of Technology, Atlanta, GA 30332, USA

³loet@leydesdorff.net
University of Amsterdam, Amsterdam School of Communications Research, 1012 CX Amsterdam (The Netherlands)

DRAFT (24 February 2011)

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

Abstract

In the context of increasingly competitive assessments frameworks, academic institutions are crafting strategies to improve their performance. Incentives to faculty to publish in high rank journals figures prominently among the policies developed by university managers. This scientometric investigation provides some elements for reflection on the potential impact on interdisciplinary practices, by comparing innovation studies units with business and management schools. First, we use various mappings and metrics to show that the innovation studies are consistently more interdisciplinary than business and management schools. Second, we provide evidence that the journals in the top ranks of the Association of Business Schools’ rankings span a less diverse set of disciplines than lower ranked journals. Third, we show that this bias results in a more favourable performance assessment of the more disciplinary-focused business and management schools. Fourth, we demonstrate that a citation-based analysis of the units’ performance challenges the ranking-based assessment. In summary, the investigation illustrates how allegedly ‘excellence-based’ journal rankings have a bias in favour of mono-disciplinary research and how this negatively affects the assessment of interdisciplinary organisations. We conclude that this case study illuminates a general mechanism through which unduly narrowly-conceived rankings can suppress interdisciplinary research.

Introduction

In a moment in which science is under pressure to be relevant to society, interdisciplinary research (IDR) is often praised for its contributions towards generating scientific breakthroughs (Hollingsworth and Hollingsworth, 2000), addressing societal problems (Lowe, 2006) and fostering innovation (Gibbons et al, 1994). Reasons given for these kinds of benefit include 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), thus rejuvenating science and contributing to its ‘health’ (Jacobs and Frickel, 2009, p. 48). In practice, however, IDR efforts are often found wanting, accused of being too risk averse, lacking under disciplinary notions of quality or not meeting policy expectations (Bruce et al., 2004, pp. 468-469).
Irrespective of perspective, IDR presents important downsides (Rhoten and Parker, 2006; Llerena and Mayer-Krahmer, 2004; Katz and Martin, 1997). First, there are coordination costs, namely the difficulties of managing knowledge integration which are common in various kinds of team work and collaboration (Cumming and Kiesler, 2005; Rafols, 2007). Second, there are institutional costs, which arise due to the institutionalisation of science in terms of disciplines. These include, for example, poor career structures for academic interdisciplinary researchers, low esteem by colleagues, discrimination by reviewers in proposals or difficulty in publishing in prestigious journals (Bruce et al. 2004, p. 464). These barriers to interdisciplinarity are not only viewed as problematic by fringe researchers struggling with mainstream disciplines, but recognised as such by top policy-makers and scientific elite (Metzger and Zare, 1999). For example, the report Facilitating Interdisciplinary Research by the US National Academies provides a thorough review of the barriers and initiatives to lower them, including a mention to the hurdle posed by high-ranking journals:

‘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.’

National Academies (2004, p. 139)

While these institutional barriers are often acknowledged, driving mechanisms are neither well documented nor deeply understood. In this UK-based case study, we provide novel quantitative evidence of an institutional barrier to IDR by exploring the conflict between the push for excellence in academia generally – focusing particularly here on business and management schools (BMS) – and the pursuit of a specific form of IDR in departments or institutes of innovation studies (IS). Under current funding conditions in the UK, many IS units have been (at least partly) incorporated into BMS (e.g. in Oxford, Imperial, Manchester, Cardiff and recently Sussex). BMS face particularly acute pressure to achieve high performance in publication rankings, both for reputational purposes and due to financial incentives associated with assessment procedures of the national funding council HEFCE, now referred to as the ‘Research Excellence Framework’ (REF). Given the disciplinary organisation of the assessment panels of the previous Research Assessment Exercises (RAE), IDR departments are perceived as being at a disadvantage (Martin and Whitley, 2010, p. 64):

‘…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.’

---

1 Here innovation studies is very broadly defined. The institutes we investigate are active in the overlapping fields of science technology studies, science policy, and innovation studies.

That evaluation of IDR is problematic is not a surprise—rather it is a natural consequence of IDR. Any evaluation needs to take place over established standards. These standards can be defined within a discipline, but what standards should be used for research in between or beyond disciplinary practices? A variety of studies have found that what happens, even in the case of multidisciplinary panels, is that IDR ends up being assessed under one of the disciplinary perspectives to which it relates (Mallard et al., 2009).

Here, we investigate quantitatively the relationship between interdisciplinarity in IS and (perceived) performance as shown by the journal rankings provided by the Association of Business Schools (ABS). The results show that ABS journal rankings strongly favour business and management disciplinary approaches—and thus disadvantage IS units by comparison with more traditional BMS. We suggest that this case is an example of a much wider phenomenon: the ‘ethnocentrism of disciplines’ associated with reinforced mainstream styles of research (Campbell, 1969).

This paper also makes a contribution to the use of scientometrics in assessment, following the proposal by Martin (1996) of a reliability test based on the convergence of multiple indicators. This investigation illustrates a robust use indicators when various perspectives converge to a same conclusion (as on the findings on interdisciplinarity presented here)—and a questionable use, when different approaches lead to contradictory insights (as on findings on performance reported).

The assessment of performance and interdisciplinarity

The assessment of scientific performance and interdisciplinarity are highly controversial. This should come as no surprise, given that their definition is disputed and agreed that they are multidimensional concepts. For the assessment of scientific performance, we follow convention and compute the mean ABS rank and the number of cites per paper. The open question is how to normalise cites by discipline. The most extensively adopted practice is to normalise by the discipline where the article is published. Though widely used, this is known to be problematic for two reasons. First, due to the heterogeneity of research even within disciplines. Second, because some papers do not conform to the disciplines of the journal—they have a guest role. This is the case, for example, with publication on science policy in biological journals. To achieve a more accurate normalisation, Leydesdorff & Opthof (2010) have proposed a fractional counting of cites, whereupon each cite is divided by the number of references of the citing publication.

The conceptualisation of interdisciplinarity is equally ambiguous, plural and controversial — inevitably leading to a lack of consensus on indicators. Even within bibliometrics, the operationalisation of IDR remains contentious (see Wagner et al. 2011 for a review that emphasises the plurality of perspectives; also Bordons et al., 2004) and defies uni-dimensional descriptions (Leydesdorff and Rafols, 2011; Huutoniemi et al. 2009). We propose to investigate interdisciplinarity from two perspectives, which we claim to be of general applicability. First, by means of the widely used conceptualisation of interdisciplinarity as knowledge integration (NAS, 2004; Porter et al., 2006), which underpins the logics of accountability and innovation (Barry et al. 2008). Second, by means of the conceptualisation of interdisciplinarity as a form of research that lies outside or in between established practices, i.e. as intermediation (Leydesdorff, 2007).
The understanding of interdisciplinarity as integration suggests looking at the distribution of components (disciplines) that have been integrated under a body of research (as shown by given output, such as a reference list). We do so by using the concepts of diversity and coherence, as illustrated in Figure 1. A full discussion on how diversity and coherence may capture knowledge integration was introduced in Rafols and Meyer (2010); here a summary is presented\(^3\). It was argued that the larger majority of bibliometric and econometric studies of interdisciplinarity have relied on indicators of diversity such as Shannon entropy and Simpson diversity (equivalent to economics' Herfindahl index). The concept of diversity, ‘an attribute of any system whose elements may be apportioned into categories’ (Stirling, 2007), allows study the distribution of disciplines of to which parts of a given body of research can be assigned. However, knowledge integration is not just about how diverse the knowledge is, but about making connections between various bodies of knowledge. This means assessing the extent to which the relations between disciplines are made along already trodden paths of whether they are novel.

---

\(^3\) See an abstract, documentation based, framework in Liu et al. (under review).
As a comparison of Figures 1 and 2 aims to illustrate, knowledge integration (as capture via diversity and coherence) and intermediation are two distinct processes associated with interdisciplinary practices—although there is often overlap between knowledge integration and intermediation. Knowledge integration 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 applications of lab on a chip, which draws both on micro-fabrication and biomedical expertise (Rafols, 2007). Intermediation 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 are related (mainly providing a service contribution) to different major disciplines. Intermediation is also typical of what Barry et al. (2008, p. 29) called ‘agonistic/antagonistic mode of research, that springs from a self-conscious dialogue with, criticism of or opposition to the intellectual, ethical or political limits of established disciplines’. These ‘antagonistic’ modes of research are seldom captured in conventional classification categories—this is why we will investigate intermediation a lower level of aggregation than diversity and coherence.

Although IDR is often associated with collaboration, several authors noted that interdisciplinary research can be conducted in different ways or research modes, including different very different forms of collaboration (Palmer, 1999; Laudel, 2001; 2002; Rafols, 2007; Barry et al. 2008).

Next, we proceed to describe in more detail how the concepts of diversity, coherence and intermediation are operationalised. As we see, the advantage of mobilising the general concepts of diversity and coherence rather than ad-hoc indicators, is that it allows rigorous choice and comparison of different mathematical forms that are equally consistent with the processes we seek to capture.

**Diversity**

A given body of research, as represented for example in the publications of a university department, is seen 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 as a proxy 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 missed to take into account the disparity among disciplines. In doing so they implicitly consider as equally interdisciplinary a combination of cell biology and biochemistry and one of geology and psychology. Only recently new indicators incorporating disparity were devised, using the metrics of similarity behind the maps of science (Porter et al., 2007; Rafols and Meyer, 2010).

This operationalization of diversity also allows to visualization potential processes of knowledge diffusion (rather than integration), but looking at the disciplinary distribution of cites to the papers of a topic or organisation (Liu et al., under review)\(^4\).

Following Yegros et al. (2010), we here investigate 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:

1. **Variety** (number of categories)

   \[
   \text{Variety} = \frac{1}{n} \sum_{i=1}^{n} p_i \ln p_i
   \]

2. **Balance** (Shannon evenness)

   \[
   \text{Balance} = -\frac{1}{n} \sum_{i=1}^{n} p_i \ln p_i
   \]

3. **Disparity** (average dissimilarity between categories)

   \[
   \text{Disparity} = \frac{1}{n(n-1)} \sum_{i,j} d_{ij}
   \]

4. **Shannon entropy**

   \[
   \text{Shannon entropy} = -\sum_{i} p_i \ln p_i
   \]

5. **Rao-Stirling diversity**

   \[
   \text{Rao-Stirling diversity} = \sum_{i,j} p_i p_j d_{ij}
   \]

6. **Coherence**

   Coherence aims to capture the extent to which the included disciplines are connected to one another. One way to look at coherence is to compare the observed average distance of cross-citations as they actually occur in the publications in question with the average distance of cross-citations that one would obtain (the ‘expected distance’) if simulated cross-citations are

\(^{4}\) In the case of research topics, also by exploring changes in the distribution over time (see Kiss et al., 2010; Leydesdorff and Rafols, forthcoming).
generated across the categories following the distribution of cross-citations found for all the publications in the WoS for 2009. Such estimate is computed taking into account that the expected proportion of citations from SCs i to j, \( p_{ij} \) (expected), is equal to the proportion of citations made from i, \( p_i \), multiplied by the conditional probability that citations go to j when they originate in i, \( p_{ij} | p_i \), namely \( p_{ij} \) (expected) = \( p_i p_{ij} | p_i \). The conditional probabilities \( p_{ij} | p_i \) are assumed to be those from all the observed cross-citations in the WoS. In summary, the measure of relative coherence is the ratio of observed of expected distance of cross-citations.

\[
Coherence \ (obs/exp) = \frac{\sum_{i,j} p_{ij} d_{ij}}{\sum_{i,j} p_i p_{ij} | p_i d_{ij}}
\]

**Intermediation**

Intermediation aims to capture the degree to which a given category of publication is distant from the areas of close-knit publications and cross-citations — those dense areas of the map representing the central disciplinary spaces. Since this measure is highly sensitive to the creation of artefacts due to classification, we here carry out the analysis at the finer level of description, namely the journal level (i.e. we use each journal as a separate category).

We use conventional network analysis measures to characterise the degree to which the publications of an organisation lie in these ‘open’ (or ‘instersticial’) spaces. The first is the clustering coefficient \( cc_i \), which identifies the proportion of observed links between categories over the possible maximum number of links. This is then weighted for each category (now an individual journal), according to its proportion \( p_i \) of publications (or references/cites), i.e. \( \sum_i p_i cc_i \). The second indicator is the average similarity (degree centrality, \( \sum_i ^{N} \frac{s_{ij}}{N} \)) weighted by the distribution of elements across the categories.

\[
Average \ similarity = \sum_i p_i \sum_j ^{N} \frac{s_{ij}}{N}
\]

**Methods**

**Data**

We investigate three of the leading British Business Schools, namely London Business School (LBS), Warwick Business School (WBS) and Imperial College Business School. From innovation studies, we study the Institute for the Study of Science Technology and Innovation (ISSTI) at the University of Edinburgh, SPRU (Science and Technology Policy Research) at the University of Sussex and the Manchester Institute of Innovation Research (MiIoIR) at the University of Manchester. The publications of all researchers identified on institutional websites as members of the six units (excluding adjunct, visiting and honorary positions) were downloaded from the Web of Science (WoS) for the period 2006-2010, limited to document types: ‘article’, ‘letter’, ‘proceedings paper’ and ‘reviews’. Publications

\(^5\) Other measures of coherence can be devised. For example, another proxy of coherence is to compare the observed average distance of cross-citations with the average distance of cross-citations that one would obtain if simulated cross-citations are generated randomly across the categories where there are publications (such as simply to reflect the relative magnitudes of the respective disciplines). This is: \( \frac{\sum_{i,j} p_{ij} d_{ij}}{\sum_{i,j} p_i p_{ij} | p_i d_{ij}} \). For the cases under study, the two measures gave similar insights.
by a researcher previous to their recruitment to the unit were also included. The download was carried out between 20th and 30th 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). In order to fully disentangle results of publications from citing articles, all cites coming from the same unit were removed. 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 has little effect on the results since intra-organisational collaborations are only about 10% of the total cites and in any case, it just affects one part of the analysis regarding cites (not the publications or references).

**Data processing and indicators of diversity and coherence**

The software Vantage Point was used to process data. A thesaurus of journals to WoS Subject Categories (SCs) was used to compute the cited SCs from the cited references. 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 within the references of journals names 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. In order to avoid counting SCs with very low proportions of references, a minimum threshold for counting an SC in the variety and disparity measures was applied at 0.01% of total publications. No threshold was applied in calculating balance, Shannon Entropy, and Rao-Stirling measures; since these inherently take into account the proportion of elements in categories.

**Disciplinary overlay maps**

The software Pajek was used to make all networks except the heat maps. First, disciplinary overlay maps were made as explained in Rafols et al. (2010), 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.

**Journal maps and indicators of intermediation**

The freeware VOSViewer ([http://www.vosviewer.com/](http://www.vosviewer.com/)) was used to make a journal map in the heat-map format. A sub-set of 391 journals was made from the journals where each unit published (excluding journals <0.5% publications per unit) and the top 100 journals which all units (collectively) referenced. The cross-citations between these journals were obtained from 2009 Journal Citation Report (JCR) also available from the WoS. 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/references per journal/cited journal, normalised to the sum of all publications/references. Intermediation measures were computed with Pajek using the journal similarities matrix. The average clustering coefficient (at 2 neighbours) was computed with a 0.2 threshold.

---

6 The method is made publicly available at [http://www.leydesdorff.net/overlaytoolkit/](http://www.leydesdorff.net/overlaytoolkit/)
Analysis of ABS rankings and performance measures

The ABS rank for each journal was obtained from the Academic Journal Quality Guide Version 4 [http://www.the-abs.org.uk/?id=257](http://www.the-abs.org.uk/?id=257). This was used to calculate the average ABS rank for each unit. For simplicity, 4* rank were converted into 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 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.7

The journal field-normalised cites/paper was made by dividing cites/paper by the average Impact Factor (IF) of a SC (i.e. $\frac{\text{Cites}}{\Sigma_{\text{SC}} \text{Publications in a given SC}}$). The citing field-normalised was made using only the citing records downloaded (i.e. excluding unit-wide self-cites), then giving each a cite weight inverse to their number of references, i.e. $\frac{1}{\# \text{References}}$. Only cites with more than 10 references were used, since papers with less are expected not to be a ‘normal’ publication outlet and have a disproportionately high value.

Results: Interdisciplinarity of organisations

The following sections present the results of this investigation. First we show that IS units are more interdisciplinary than BMS in general 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 each unit is shown in [www.interdisciplinaryscience.net/maps](http://www.interdisciplinaryscience.net/maps) and as supplementary materials.8 We skip the details of the overlay technique, since it is discussed at length in Rafols et al. (2010)9. These overlay maps were generated for the six units using the SCs of publications, references and cites (excluding self-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). The differences are more pronounced in the case of publications and cites than for references.

---

7 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, that there was, in average a 8.5% discrepancy between the unique papers found in WoS citing them, and the counts in TC –due to the fact that one paper can cite several papers from one organisation. By using this average discrepancy, 792 cites (730 cites plus the 8.5% discrepancy) were estimated. The possible inaccuracy introduced by this extrapolation is well within the standard error.

8 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_Diversity.pptx) [http://www.sussex.ac.uk/Users/ir28/IDR/Disciplinary_Coherence.pptx](http://www.sussex.ac.uk/Users/ir28/IDR/Disciplinary_Coherence.pptx)

9 To see the relative positions on the global maps of science see the interactive map [http://www.idr.gatech.edu/detail.php?tab=1&id=1](http://www.idr.gatech.edu/detail.php?tab=1&id=1).
which tend to be relatively widely spread both for 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, which show which 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’ various disciplines, but are actually doing interdisciplinary work. In particular, they play a bridging role between the natural sciences and social sciences.

On the contrary, the leading BMS examined here are not only less diverse, but also more fragmented in disciplinary terms, in the sense that they tend to cite more within disciplines. For example, Imperial is the most diverse of the BMS, thanks in part to its research on health services, but this line of research is not strongly linked to other Imperial social sciences, as shown by scarcity of cross-citations. The bridging function carried out by IS units is captured by the coherence indicator shown in Table 1 and Figure 4. ISSTI and SPRU have 13% and 5.4% more cross-citations than would be expected, whereas LBS just makes about half its expected proportion cross-citations.

Measures such as diversity might have size effects, i.e. tend to increase or decrease depending of the population size. Since the IS units are between 2 to 4 times larger than BMS, one might wonder if size-effect might explain the differences in the diversity measures. However, the expected size effect would be that larger units tend to have larger measures of 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 certain that the results are not an indirect effect of size. There is no size effects expected in the case of coherence, given that it is computed from a ratio.
Figure 3. Overlay of SCs of references by a unit on the global map of science (grey background). Cross-citations are shown (green links) only for observed values 5 fold larger than expected.
Table 1. Indicators of diversity and coherence for each organisational unit

<table>
<thead>
<tr>
<th></th>
<th>Innovation Studies Units</th>
<th>Business and Management Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Edinburgh ISSTI</td>
<td>Sussex SPRU</td>
</tr>
<tr>
<td># of Publications</td>
<td>129</td>
<td>155</td>
</tr>
<tr>
<td>SC of Publications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>Balance</td>
<td>0.653</td>
<td>0.566</td>
</tr>
<tr>
<td>Disparity</td>
<td>0.832</td>
<td>0.839</td>
</tr>
<tr>
<td>Shannon Entropy</td>
<td>3.558</td>
<td>3.243</td>
</tr>
<tr>
<td>Rao-Stirling Diversity</td>
<td>0.810</td>
<td>0.783</td>
</tr>
<tr>
<td># of References</td>
<td>1737</td>
<td>2409</td>
</tr>
<tr>
<td>SC of References</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>28</td>
<td>18</td>
</tr>
<tr>
<td>Balance</td>
<td>0.510</td>
<td>0.420</td>
</tr>
<tr>
<td>Disparity</td>
<td>0.829</td>
<td>0.842</td>
</tr>
<tr>
<td>Rao-Stirling Diversity</td>
<td>0.833</td>
<td>0.791</td>
</tr>
<tr>
<td># of Cites</td>
<td>316</td>
<td>767</td>
</tr>
<tr>
<td>SC of Cites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>32</td>
<td>21</td>
</tr>
<tr>
<td>Balance</td>
<td>0.669</td>
<td>0.513</td>
</tr>
<tr>
<td>Disparity</td>
<td>0.852</td>
<td>0.844</td>
</tr>
<tr>
<td>Rao-Stirling Diversity</td>
<td>0.851</td>
<td>0.810</td>
</tr>
<tr>
<td>Cites between SC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>1.131</td>
<td>1.054</td>
</tr>
</tbody>
</table>

NB: 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.
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 (see methods). In this case we used the visualisation software VOSviewer, since it allows us to make a ‘heat map’ depicting the density of nodes and links of different parts of the map. This visualisation is helpful to distinguish between dense areas (associated with disciplinary cores), and sparser interstitial areas (associated with IDR). To make the map we followed again the overlay technique: cross-citation data from the WoS 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 is on a different basis to conventional journal maps (where positions reflect direct similarities, since they have no overlay).

The IS-BMS journal maps (Figure 5\textsuperscript{10}) show three poles: management, economics, and natural sciences. This latter encompasses the various particular natural sciences in which these focal units work. This 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 with the natural sciences that social sciences have. More science-oriented journals such as Social Studies of Science are closer to this pole. The relative position of the different areas is 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 publish, reference and are cited by journals in the dense areas of management and economics. The partial exception is Imperial, with a research subgroup that is active in health sciences. IS units, on the contrary, have most of their activity in the interstitial areas lying between management, economics and the natural sciences, in journals such as Research Policy, or in journals of application areas such as Social Science and Medicine or Energy Policy. This difference between the degree of activity in intermediation is shown by the indicator of clustering coefficient and the average similarity of the journals (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.

\textsuperscript{10} More maps complementing Figure 3 can be retrieved from: http://www.sussex.ac.uk/Users/ir28/IDR/Intermediation.pptx
Figure 5. Overlay of journals in the references by ISSTI and LBS on the heat map based on the citation-similarities between journals (based on WoS 2009).
Table 2. Indicators of intermediation by organisational unit.

<table>
<thead>
<tr>
<th></th>
<th>Innovation Studies Units</th>
<th>Business and Management Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISSTI</td>
<td>SPRU</td>
</tr>
<tr>
<td>Journals of pubs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coeff.</td>
<td>0.128</td>
<td>0.098</td>
</tr>
<tr>
<td>Average similarity</td>
<td>0.028</td>
<td>0.034</td>
</tr>
<tr>
<td>Journals of references</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coeff.</td>
<td>0.178</td>
<td>0.182</td>
</tr>
<tr>
<td>Average similarity</td>
<td>0.044</td>
<td>0.050</td>
</tr>
<tr>
<td>Journals of cites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coeff.</td>
<td>0.120</td>
<td>0.096</td>
</tr>
<tr>
<td>Average similarity</td>
<td>0.029</td>
<td>0.034</td>
</tr>
</tbody>
</table>

NB: low values for each metric indicate higher levels of intermediation. Standard errors are not provided because they are all smaller than 0.07%, i.e. negligible.

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

Disciplinary bias in 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 used the JCR to assign journals to SCs. The coverage of assignment was low for rank 1 (14%), but reached an acceptable level for rank 2 (56%), and was almost complete at the highest ranks. Then, we looked at the disciplinary diversity of each rank, by looking at 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 7 and 8.11

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

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

<table>
<thead>
<tr>
<th></th>
<th>Rank 1 Modest standard</th>
<th>Rank 2 Acceptable standard</th>
<th>Rank 3 Highly regarded</th>
<th>Rank 4 Top in Field</th>
<th>Rank 4* World Elite</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Journals</td>
<td>205</td>
<td>295</td>
<td>231</td>
<td>73</td>
<td>21</td>
</tr>
<tr>
<td>% of Journals in JCR</td>
<td>14%</td>
<td>56%</td>
<td>86%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>SC of Journals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>27</td>
<td>58</td>
<td>56</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>Balance</td>
<td>0.797</td>
<td>0.611</td>
<td>0.558</td>
<td>0.606</td>
<td>0.573</td>
</tr>
<tr>
<td>Disparity</td>
<td>0.866</td>
<td>0.737</td>
<td>0.657</td>
<td>0.755</td>
<td>0.767</td>
</tr>
<tr>
<td>Shannon Entropy</td>
<td>2.979</td>
<td>3.454</td>
<td>3.280</td>
<td>2.940</td>
<td>2.002</td>
</tr>
<tr>
<td>Rao-Stirling Diversity</td>
<td><strong>0.779</strong></td>
<td><strong>0.733</strong></td>
<td><strong>0.703</strong></td>
<td><strong>0.685</strong></td>
<td><strong>0.571</strong></td>
</tr>
</tbody>
</table>

NB: 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 rankings span a less diverse set of disciplines than lower rankings. 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.
Performance assessment 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 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 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. The results conclusive 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 the cites received in 2010 instead of using fixed citation windows, the results should be interpreted as only indicative. Although imperfect, the estimate obtained is expected be sufficiently robust as to provide tentative insights and illustrate the inherent difficulties and ambiguities of using citation-based performance indicators.

Following conventional practice we use the mean to describe the distribution. This is widely acknowledged to be a flawed method, given the highly skewed nature of citation distributions (see Katz, 2000; Leydesdorff and Opthof, 2011). The result of this conventional statistical (mal)practice is that the standard error of the mean is very high (in the ~8-18% range) – so high that ranking units becomes problematic (a major concern in policy-oriented assessment which might deserve a study of its own).

The analysis shows, first, that in terms of raw number of cites, BMS do not perform better than IS units, although there is a weak correlation with ABS-performance (0.47). Second, using a normalisation based on the field of publication (the average impact factor of the SC of publication), one obtains a relative improvement of BMS performances, with a 0.76 correlation with ABS-performance. One can advance a cause for this result: if IS papers are normalised by field, they are doubly disadvantaged in respect both of their publishing in natural sciences (because even if they receive many cites, they may – all else being equal – tend to be less so than natural science papers), or in the social sciences (because they have disproportionately difficulties in publishing in the most prestigious journals). Third, we use a normalisation recently proposed by Zhou and Leydesdorff (2011) which weighs each citation by the number of references in the citing paper. In doing so, it achieves a much more accurate description of the citing context of each individual paper. Most interestingly, under this normalisation, the correlation between citation based and ABS-based performance vanishes to a negligible -0.03.

---

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

13 The problem of using fixed citations is that it only allows to make 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 the gathering of the publications.
In summary, this exploratory analysis of different performance measures highlights the problems of commensurability in appraising IDR publications, and challenges the performance assessment of ABS-rankings. In short, a high performance in ABS terms is a mark of disciplinary compliance, but is not necessarily related to high citation performance.

Table 4. Performance indicators.

<table>
<thead>
<tr>
<th></th>
<th>Innovation Studies Units</th>
<th>Business and Management Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISSTI</td>
<td>SPRU</td>
</tr>
<tr>
<td><strong>ABS ranking-based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ABS Rank</td>
<td>2.82 (0.13)</td>
<td>2.65 (0.10)</td>
</tr>
<tr>
<td>% Papers Ranked</td>
<td>43%</td>
<td>51%</td>
</tr>
<tr>
<td><strong>Citation-based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cites/paper</td>
<td>2.69 (0.45)</td>
<td>5.11 (0.59)</td>
</tr>
<tr>
<td>Journal field normalized Cites/paper</td>
<td>1.67 (0.28)</td>
<td>2.79 (0.35)</td>
</tr>
<tr>
<td>Citing field normalized Cites/paper</td>
<td>0.18 (n.a.)</td>
<td>0.12 (n.a.)</td>
</tr>
</tbody>
</table>

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

Figure 9. Performance indicators.
The bias in the performance measures introduced by the ABS ranks is amplified in the case of current assessment of units by complementary distorting mechanisms. The first one, already mentioned, is that the percentage of publications included in the ABS classification is much lower for IS units than for BMS (shown in Figure 10). In an evaluation such as REF, where each researcher can submit up to four articles, this means that an IS unit researcher evaluated under a Business and Management panel, may need to publish eight articles so that four of them fall within the ABS remit (unless she is willing to change her publication patterns).

The second mechanism is the exponential scale that assessment exercises tend to use to reward perceived quality. In terms of resource (financial) allocation this means, for example, that Rank 1 articles have a 0 multiplier (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 exponential scale, the ~50% performance difference in ABS ranks between MIoIR, becomes a ~120% difference in the resources received by the units (see Figure 10).

Although the upcoming UK’s departmental assessment exercise (REF) does not formally rely on quantitative indicators such as rankings, the widespread perception in the field of Business and Management is that number of publications in top ABS ranks is an accurate predictor of the assessment. Under these conditions, it is no surprise that there is major pressure on researchers to target their publications to ‘high ranking’ journals. It follows, from our previous evidence on disciplinary bias, that this is likely to result in a suppression of interdisciplinarity.

**Figure 10. The effect of rankings on the outcomes of assessment exercises.**

**Conclusions**

This empirical investigation has demonstrated that IS units are more interdisciplinary than leading BMS under various perspectives. It has shown that ABS rankings have a disciplinary bias which translates very directly into a low assessment of interdisciplinary units’ performance. We have shown that this low assessment is not warranted by citation-count performance. In this way, the present pilot study suggests that the use of ABS rankings serves systematically to disadvantage against IDR – a finding that might be tested in analysis of a wider array of BMS-related IDR. To the extent that ABS ranking are becoming increasingly used to evaluate individual and organisational research performance in this field, it does seem likely that they have a suppressive effect on IDR, including that in the IS field.
From a qualitative perspective these findings are not new. Science studies and policy documents have longed observed 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 (Bruce et al. 2004, National Academies, 2004; Metzger and Zare, 2003). In recent decades these 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 ranking procedure is predicated on the assumption that the resulting ranks constitute *objective assessments* that can be treated as robust proxies for academic excellence. These empirical results challenge such claims to objectivity and suggest that the resulting picture presents a rather narrow and idiosyncratic view of excellence. When used in helping to determine assignments of esteem and resources, rankings that remain uncorrected for these effects can have the effect of suppressing forms of interdisciplinarity that are otherwise widely acknowledged to be academically and socially positive.

**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](http://www.interdisciplinaryscience.net/maps)

**Also, in powerpoint format:**


Intermediation: [http://www.sussex.ac.uk/Users/ir28/IDR/Intermediation.pptx](http://www.sussex.ac.uk/Users/ir28/IDR/Intermediation.pptx)


Comparison of Unit’s Performances: [http://www.sussex.ac.uk/Users/ir28/IDR/Performance_Comparison.pptx](http://www.sussex.ac.uk/Users/ir28/IDR/Performance_Comparison.pptx)

**Acknowledgments**

IR and AO were funded by the US NSF (Award no. 0830207, [http://idr.gatech.edu/](http://idr.gatech.edu/)) and the EU FP7 project FRIDA (grant 225546, [http://www.fridaproject.eu](http://www.fridaproject.eu)). We thank Jan Fagerberg for comments.

**References**


