Universiteit Leiden
The Netherlands

Discarding the 'basic science/applied science' dichotomy: A Knowledge Utilization Triangle classification system of research journals

Tijssen, R.J.W.

Citation

Tijssen, R. J. W. (2010). *Discarding the 'basic science/applied science' dichotomy: A Knowledge Utilization Triangle classification system of research journals*. Leiden: SW Centrum Wetensch. & Techn. Studies (CWTS). Retrieved from https://hdl.handle.net/1887/15046

Version:	Not Applicable (or Unknown)
License:	[Leiden University Non-exclusive license]()
Downloaded from:	https://hdl.handle.net/1887/15046

Note: To cite this publication please use the final published version (if applicable).

Discarding the ‘basic science/applied science’ dichotomy: A Knowledge Utilization Triangle classification system of research journals

Robert J.W. Tijssen

CWTS Working Paper Series

Paper number	CWTS-WP-2010-002
Publication date	March 2, 2010
Number of pages	25
Email address corresponding author	tijssen@cwts.leidenuniv.nl
Address CWTS	Centre for Science and Technology Studies (CWTS) Leiden University P.O. Box 905 2300 AX Leiden The Netherlands www.cwts.leidenuniv.nl

Universiteit Leiden

CWTS

Centrum voor Wetenschaps- en Technologie Studies

Discarding the 'basic science/applied science' dichotomy: A Knowledge Utilization Triangle classification system of research journals

Robert J.W. Tijssen

Center for Science and Technology Studies (CWTS), Leiden University, The Netherlands. E-mail: tijssen@cwts.leidenuniv.nl

This paper introduces a new system for classifying scholarly journals in terms of their degree of 'application orientation'. The method extends earlier models and journals classification systems that were designed to tackle the crude duality between 'basic research' and 'applied research'. The metrics-based classification system rests on a 'Knowledge Utilization Triangle' typology, which distinguishes three types of co-existing knowledge application domains: 'clinical', 'industrial' and 'civic'. The empirical data for each journal metrics relate to the institutional origin of authors who publish their research papers in scientific journal literature.
The case study applies 'clinical relevance' and 'industrial relevance' to 11 000 journals indexed by the *Web of Science* database. The resulting multidimensional classification system of journals comprises of six *Journal Application Domain* categories. Macro-level trend analysis of the WoS-indexed research publication output by JAD category reveals redistributions within global science during the years 1999-2008, with a slight increase of output published in 'industrially relevant' journals.

Introduction

Science is a complex adaptive system. Statistical studies of the system's structural properties are stifled by fundamental conceptual and methodological problems. One of them is the commonly adopted definition of 'basic research' and 'applied research'. The traditional understanding of 'basic science' relates to the relevance for problems internal to science, while 'applied science' corresponds to relevance external to science itself. [1]

Developing quantitative measures that truly capture the differences between basic research and applied research are impossible for a lack of comprehensive and systematic data. As a result, progress on empirical modeling of science systems, and how science evolves, is hampered by this crude and outdated dichotomy between 'basic research' and 'applied research'.

Rather than using the 'basic-applied' perspective of knowledge production, and (possible) economic usage, this paper opts for an analytical perspective focusing on knowledge producers (authors of research publications) and usage within the scientific communities (absorbing and using the contents of publications). And rather than trying to classify research activities intrinsically according to the associated type of research, this approach offers the required empirical information on research characteristics that is amenable to systematic measurement. It offers the possibility to develop a metrics and measurement criteria ranging across the spectrum from 'basic' to 'applied'. The next section provides a brief overview of these concepts and earlier work on classification systems. The following section introduces the methodological approach for the new journal-based classification system. The Findings section applies the classification system to a large set of journals indexed by the *Web of Science* database. In the final section the paper concludes by reviewing the method's current methodological constraints as well as its potential for further applications and improvements.

Previous work

The concept of 'basic research' has been a topic on science policy and research funding agenda's for many decades. Attempts to arrive at a standardized definition, mostly for the purpose of statistical surveys, were often a subject of

1 From a purely epistemological perspective all sciences could be seen as 'applied', in the sense that they have to be relevant to some purpose.

academic studies and heated debate among statisticians, especially within the USA and the *Organisation for Economic Cooperation and Development* (Godin, 2003). Many taxonomies and classification systems dealing with the various types of research have been proposed and rejected; almost 40 years ago, Rothschild (1972) identified as many as 45 in the literature.

The leading 'official' definition nowadays, endorsed by many statistical offices worldwide and institutionalized in their statistical surveys, was developed by the OECD and features in its *Frascati Manual* (OECD, 2002): '*Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts, without any particular application or use in view*.' The manual elaborates by stating: '*The reference to no 'particular application in view' in the definition of basic research is crucial, as the performer may not know about actual applications when doing the research or responding to survey questionnaires. The results of basic research are not generally sold but are usually published in* scientific journals or circulated to interested colleagues.' The OECD, fully aware of the definition's shortcomings, suggests the following split with the different types of research: '*Pure basic research is carried out for the advancement of knowledge, without seeking long-term economic or social benefits or making any effort to apply the results to practical problems or to transfer the results to sectors responsible for their application*', and '*Oriented basic research is carried out with the expectation that it will produce a broad base of knowledge likely to form the basis of the solution to recognized or expected, current or future problems or possibilities*', while '*Applied research is also considered as original investigation to acquire new knowledge. It is, however, directed primarily towards a specific practical aim or objective*.' In this Manual, the OECD also recognizes the need for further elaboration with regards to 'applied research': '*Applied research is undertaken either to determine possible uses for the findings of basic research or to determine new methods or ways of achieving specific and predetermined objectives. It involves considering the available knowledge and its extension in order to solve particular problems. In the business enterprise sector, the distinction between basic and applied research is often marked by the creation of a new project to explore promising results of a basic research programme',* where '*The results of applied research are intended primarily to be valid for a single or limited number of products, operations, methods or systems. Applied research gives operational form to ideas. The knowledge or information derived from it is often patented but may be kept secret*.'

Despite all the OECD's good intentions and efforts, these definitions still lack sufficient clarity in terms of to measurable aspects of research inputs, outputs or impacts. As a consequence, statisticians, scientometricians and other analysts have been forced to contend themselves with these rather fuzzy notions and descriptions, which in turn continues to hamper thorough comparative statistical analyses of science (Godin, 2003). Fortunately, some conceptual and methodological progress has been made outside the realm of the statistical offices to devise categories of research activities that transcend the 'basic/applied' duality. OECD's 'oriented basic research' concept resonates with views of Stokes (1997) who introduced the concept of 'Pasteur's Quadrant' to question the legitimacy of this duality for describing and analyzing the general patterns and trends within contemporary science. The quadrant is framed by two key questions regarding the ultimate goal of research activities: is it driven by quest for fundamental understanding (yes/no), and are there considerations for use (yes/no)? The quadrant comprises of three meaningful categories:

- Pure applied research;
- Use-inspired research;
- Pure basic research.

This extension from two to three categories introduces the 'context of application' perspective and creates a middle ground of 'use-inspired' categories within a classification system. However, similarly to the OCEC definitions, Stokes' conceptual framework lacks an analytical model to translate these categories into systematic large-scale collection of empirical data and comparative quantitative measurement.

The only method available thus far, is one of developed by *CHI Research* (Philadelphia, USA) in the 1980s.[2] This classification system of science is based on expert assessments of individual research journals, which are assigned to one of four categories ('levels' in CHI terminology) according to a journal's degree of 'appliedness' as reflected in its contents (Noma, 1986; Hamilton, 2003). *CHI Research* designed two related classification systems. The field-specific classification system dealt exclusively with the biomedical fields, where a journal is labeled by one of the following levels:

1. Clinical observation;

---

2 *CHI Research* was a research-based consultancy company founded in 1968 to provide information services to both government and private clients in the USA. Its staff members, including the Director Francis Narin, were active contributors to the scientometrics community from the early 1970s, to the late 1990s. Francis Narin has now retired and CHI Research no longer exists.

2. Clinical observation and investigation;
3. Clinical investigation;
4. Basic biomedical research.

CHI's 'generic' classification system covers all other fields of science:

1. Applied technology;
2. Engineering science - technological science;
3. Applied research - targeted basic research;
4. Basic scientific research.

Both classification systems have been used extensively within a range of empirical studies, both by staff members at *CHI Research* (e.g. Narin and Rozek, 1988; Hicks and Hamilton, 1999) as well as others (e.g. Brusoni and Geuna, 2003; Lewison and Paraje, 2004; Lim, 2004), mainly to describe and analyze macro-level features of science as represented in bibliometric studies of the research literature. Systematic measurements and metrics require objective statistical data that are generally seen as valid empirical evidence. Various unresolved issues however remain with regard to the validity of those four categories and their discriminatory power, and therefore, by extension, the empirical base and conceptual relevance of the classification system as such. Unfortunately, the documentation available in the open literature provides no clarification as to the underlying rationale or theory employed by *CHI Research* to design these classification systems, nor to the methodological rigor of the way in which journals were attributed to categories. Not surprisingly, this method has not been embraced by government statistical offices.

Nonetheless, for lack of feasible alternatives, using research journals as an entry point to collect comprehensive empirical data remains an attractive approach for developing more elaborate classification systems of science.

Methodology

Some of methodological problems can be partially circumvented by introducing a general typology of journals based on the concept of the 'institutional research environment', i.e. the organizational environment in which research activities are conducted. The notion of classifying research activities according to the institutional environment of researchers dates back to the 1960s (Reagan, 1967), an idea further explored by a series of studies during the 1970s (e.g. Falk, 1973; Brooks, 1980; Langenberg, 1980).

Two of those environments are particularly relevant: the business sector and the medical sector. Both are characterized by a framework of governance and managerial structures, incentive systems, organizational routines and economic forces that affect the choice of research objectives and the extent to which knowledge creation processes are aimed at meeting the need for practical applications by (end) users outside academia. The common underlying assumption is that an applications-oriented research environment is a sufficiently robust proxy measure for applications-oriented research, especially within aggregate-level large-scale systematic studies that comprise a large quantity of research articles produced within these environments. In the case of the business sector (industry mainly), the research agenda's and activities are devoted to topics and outputs with possible medium-term or longer-term commercial applications in the market place. In the medical sector the applications should related to public health care, such as improvements of medical technologies, novel drugs, or therapeutic treatments.

Journals that contain publications (co-)authored by the corporate sector authors, are therefore more likely to represent 'applications-oriented science'[3] relevant for (science-based) industry and other business sector organizations. Peer-reviewed learned journals with a large share of corporate authors likely to perceive industrial R&D staff as one of their major 'target groups'. Likewise, journals publishing many research articles originating from staff employed by general hospitals and medical centers represent 'applications-oriented science' relevant for clinical practice. A third (minor) domain of science-based knowledge utilization is the 'civic sector': government agencies, ministries, and other public sector organizations. Journals with authors from this sector would represent 'applications-oriented science' for a wide variety of topics ranging from public management and policy, economic forecasting, legal issues, to accounting and taxation. Some journals may have no publications listing corporate sector addresses or medical sector addresses, or at best negligible fractions. These journals, with a predominantly academic authorship (including university medical centers), are hence considered to be 'basic' – i.e. more oriented towards curiosity-driven, exploratory 'discovery-oriented research'. Figure 1 presents a stylized graphical model based on a 'Knowledge Utilization Triangle' typology of science-based application domains, where 'scientific research' is located at its core as the source of information and knowledge while the different types of application domains define its edges. The triangle comprises a spectrum of research activities and application domains without clear-cut boundaries.

3 'Applications-oriented science' is meant to encompass to a range of synonyms used within the academic and statistics literature, such as 'directed research', 'strategic research' and 'mission-targeted research'.

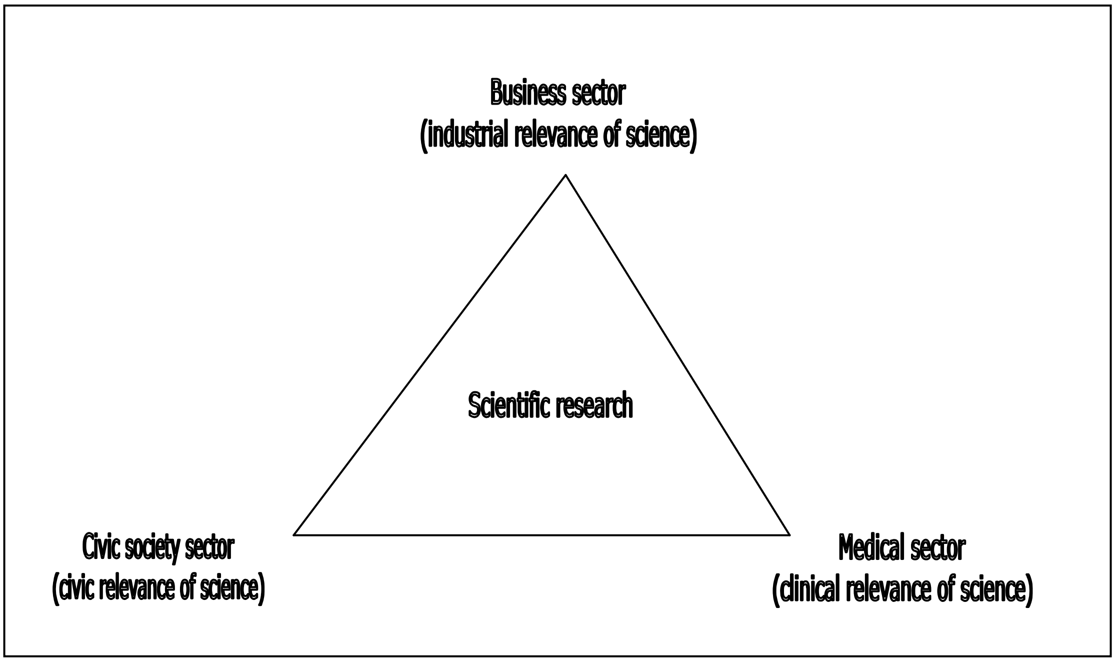

FIG. 1. Knowledge utilization triangle: application domains of scientific research.

Based on the scores along each hypothetical axis, radiating of the core, one can profile each journal according to its degree of relevance within major domains. Journals with a high score in either domain, and hence located closest to an edge of the triangle, are classified as 'application oriented'; those with a low score are seen as 'discovery oriented' and reside in the centre of this triangle. Those core journals are predominantly journals with 'academic' contributors employed by universities or other public sector research-performing organizations. The corresponding journal classification system distinguishes three major domains of research applications: the medical and health sector, the business sector, and the civic society sector. Accordingly, the classification system assigns each journal three percentages, each ranging from 0 to 100:

- *Clinical relevance* - % of author affiliate addresses referring to a general hospital, medical center or clinic;
- *Industrial relevance* - % of author affiliate addresses referring to a business enterprise on another type of private sector organization;
- *Civic relevance* - % of author affiliate addresses referring to local, national or supranational government agencies, ministries, or other civic society organizations (e.g. societies, trade unions).

This classification methodology is platform-independent. It can be applied to any source consisting of bibliographic information on author affiliate addresses in (open access) journals or other sets of documents that describe research findings. It is therefore suitable and applicable to any comprehensive bibliographical database containing published research outputs, notably the large multidisciplinary databases such as Thomson Reuters *Web of Science* or Elsevier's *SCOPUS*, but also disciplinary databases such as *PubMed/Medline*. The next section describes a first application of the journal classification system: the *Web of Science* database.

Application

Web of Science database

The current edition of the classification system was applied to scholarly journals indexed by Thomson Reuters' *Web of Science* database (WoS).[4] All bibliographic records of a WoS-indexed publications, across the 10 years time-period 1999-2008, were scanned for author affiliate addresses. The data collection was restricted to journals of sufficient size for statistical analysis, i.e. those with a total of publications containing more than 50 different author affiliate addresses in 1999-2008. Journals that operated under one or more consecutive names, or name variants, during this period were merged into a single 'standardized' journal title. The smallest journals, those that were discontinued by publishers or very intermittently indexed by the WoS, were discarded. A total of 11 558 journals were entered into the analysis.

Note that the coverage of 'civic sector' research publications is limited within the WoS given the type of journals represented in this database. There are relatively few serials in the database with a strong focus on (local) societal issues and practical problems that would attract a large 'civic sector' authorship. It would require a significant extension of the database (trade journals, professional journals and report series) to provide the same level of coverage as the medical research literature and the industrial research literature. The remainder of this case study is therefore restricted to clinical relevance and industrial relevance.

As for the operationalization of industrial relevance, each author affiliate address was assigned to the corporate sector if the publication was registered within CWTS' *Corporate Research Publications* database (see Tijssen et al., 2009). The top 3 journals with the largest share of corporate addresses are: *Japan*

---

4 Pertains to the Web of Science (WoS) operated by CWTS under a license agreement with the database producer *Thomson Reuters*. The WoS used in this study includes the SCI-Expanded, SSCI and AHCI databases.

Telecommunications Review; *Review of the Electrical Communications Laboratories*; *Fujitsu Scientific & Technical Journal*.

In the case of clinical relevance, each address was assigned to the medical sector if the bibliographic information within the CWTS/WoS information system complies with two selection criteria:

1. Author address information includes one or more of the following character strings: 'Hosp', 'Klinikum', 'Kliniken ', 'Med Ctr', 'Ctr Med', 'Coll Med', 'Med Coll', 'Clin Ctr', 'Hlth Ctr', 'Infirm', 'Infirmary', 'Policlin', 'Ziekenhuis', 'Hop ', 'Spital', 'Osped ', 'Psychiat Klin' and 'Oncol Ctr';
2. CWTS has not assigned the publication to the higher education sector.

The top 3 journals with the largest degree of 'clinical relevance' are: *King Faisal Specialist Hospital Medical Journal*; *Journal of the American Podiatry Association*; *Journal of the Maine Medical Association*.

Introducing the second criterion excludes university hospitals and academic medical centers. These organizations are assumed to be primarily engaged basic 'discovery oriented' research. The distinction between university hospitals and other university organizations is often also blurred in author affiliate addresses because academics often simply list their university as affiliate address, with no mentioning of the hospital, clinic or other medical centre. General hospitals that are assumed to be primarily engaged in 'application oriented' clinical research and applications of medical science as compared to their academic counterparts. However, the exclusion of the university hospitals and medical centers from the medical sector remains problematic given the lack of a clear-cut division between the university-affiliated medical institutes and those outside the university system. In practice, similar application domains occur in both types of organizations, notably the intermediate 'transfer' orientation often referred to as 'translational research'. Also, it is not uncommon for research staff having affiliations in both kinds of institutes, nor is uncommon that general hospitals participate in academic research projects (e.g. in clinical trials).

Figure 2 depicts the scatter of the journals on the 'clinical relevance' axis and the 'industrial relevance' axis. Both distributions are extremely skewed towards the bottom end, reflecting the predominantly academic authorship of the WoS-indexed journals. Both distributions have a mode equal to 0, a median value (50% percentile) of 1.2. The mean score on 'clinical relevance' equals 4.4; whereas 'industrial relevance' mean equals 3.7. The distributions clearly indicate that each metric represents a distinct knowledge utilization dimension (the Pearson correlation coefficient between both metrics equals r=-0.14). Many of the journals with relatively high scores on both axes are at the intersection of industry and medicine, including research fields relevant to clinical trials of new drugs, and likely to include research articles (co-)produced by research staff at biopharmaceutical companies as well as by hospital staff (further elaborated in Table 3).

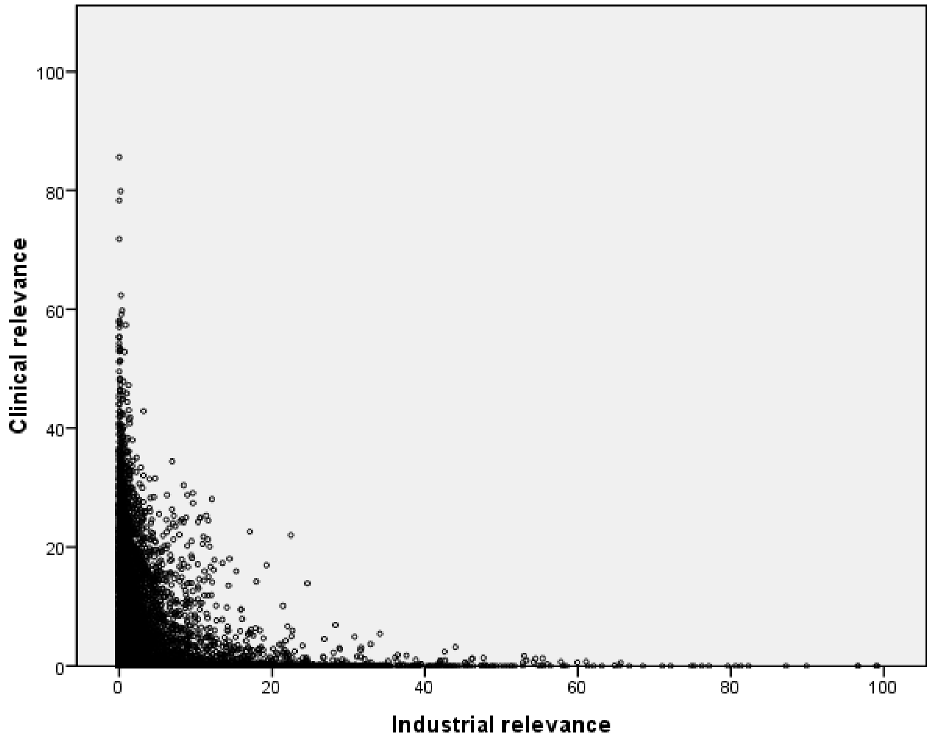

FIG. 2. Journal distribution across knowledge utilization domains (% of addresses referring to the corporate sector or medical sector).

Comparison with the CHI Research classification system

How do these metrics relate to CHI Research's 4-level classification system? Only a subset of 5517 WoS-indexed journals can be labeled by CHI Level, where more recently indexed journals will be under-presented. The CHI-indexed journals were distributed across the four levels as follows: 1. n=978 (18%); 2. n=1 375 (25%); 3. n=1 369 (25%); 4. n=1 795 (32%). Cross-distributions of the relevance scores and these levels are displayed in Figures 3a and 3b.

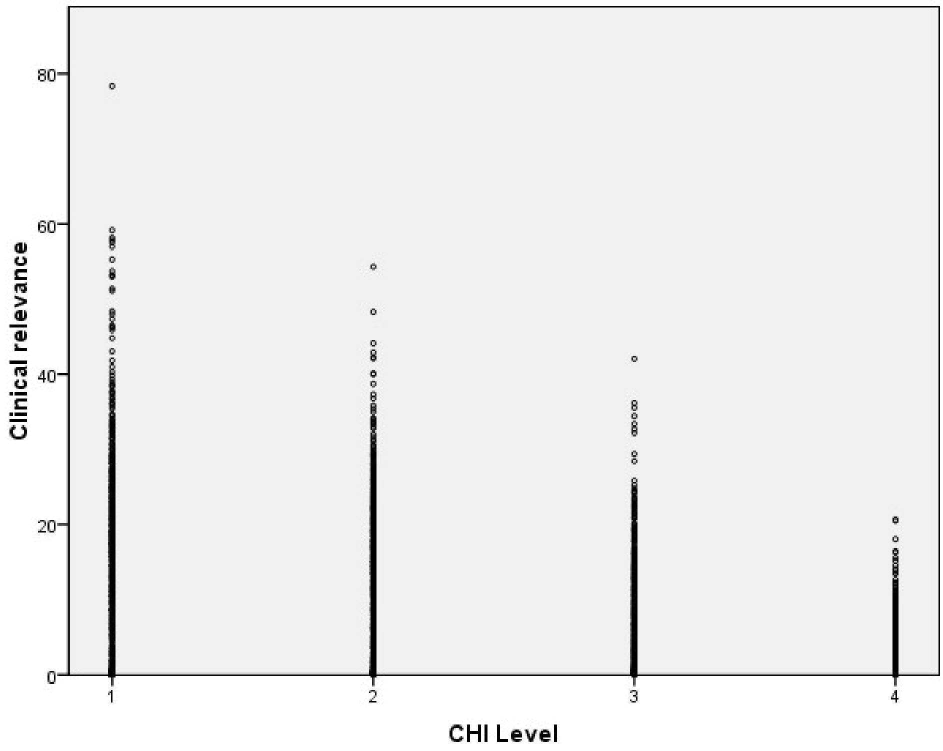

FIG 3a. Distribution of scores on clinical relevance per CHI level (% of addresses referring to the medical sector)

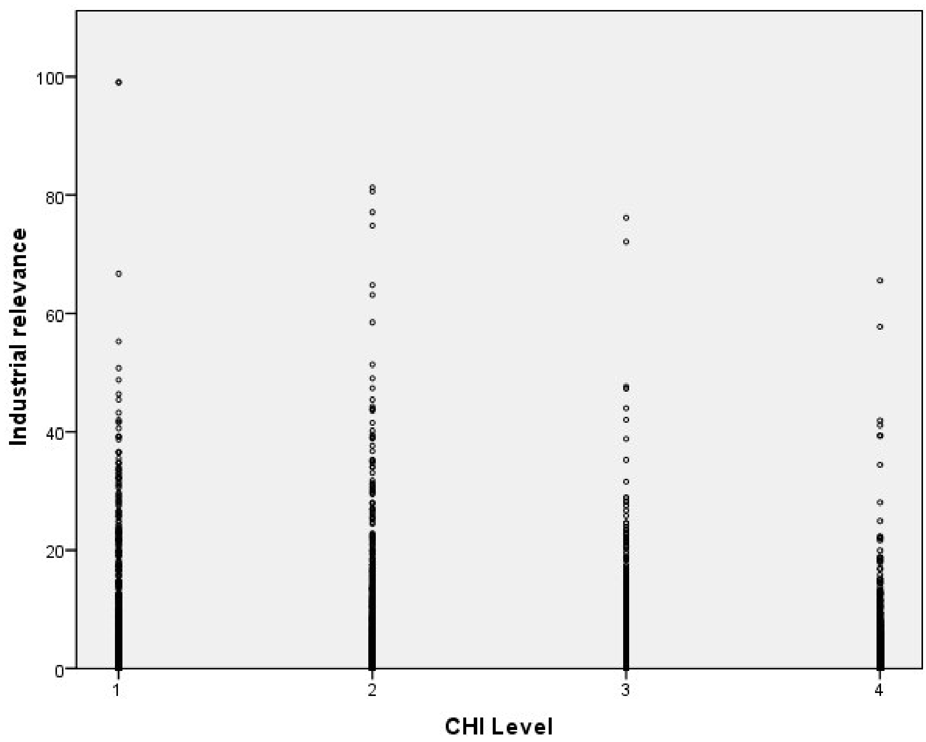

FIG 3b. Distribution of scores on industrial relevance per CHI level (% of addresses referring to the corporate sector)

Figure 3a, relating to clinical relevance, shows the expected outcome: as the CHI level become more 'basic' (Level 4), the share of addresses referring to the medical sector declines. There is a considerable spread within each level, with the weakest correspondence between CHI levels and clinical relevance categories is at its weakest in the 'applied' segment of the journal distribution (Levels 1 and 2). Significant correlation coefficients are found between CHI levels and clinical relevance scores: Pearson correlation coefficient $P=-0.38$ and Spearman's rho coefficient $\rho=-0.19$ (both coefficients are significant at the 0.01 level).

In Figure 3b, dealing with industrial relevance, the same general pattern occurs but the intra-level spread is even larger than in the case of clinical relevance, where the CHI level fails to make any meaningful distinction within the applied Levels 1 and 2. Somewhat lower, but still significant, correlation coefficients are found between CHI levels and the industrial relevance scores: Pearson correlation coefficient $P=-0.19$ and Spearman's rho coefficient $\rho=-0.12$ (both are significant at the 0.01 level).

In view of the fact that the *CHI Research* classification system dates back to the 1980s and 1990s, or at very least back to 2002-2003 (Hamilton, 2003), many journals might have shifted their cognitive emphasis in recent years – some becoming more 'applied', others more 'basic'. Part of the observed discrepancy between the classification systems might therefore allude to changes in journal editorial policies, and associated shifts in content and institutional origins of authors that may have occurred in recent years. However, the major part of the striking differences between the relevance scores and the CHI classification system are left unexplained, which suggests that the discriminatory power of the relevance scores seems superior, especially in identifying and categorizing the applications-oriented journals.

Validation studies

To gauge the validity of the 'industrial relevance' and 'clinical relevance' metrics one needs an independent information source to describe a journal's 'application orientation' profile. The two most obvious options for conducting such cross-validity studies are problematic, if not impossible: one could try to get hold off data from journal publishers to identify the share of subscribers from the medical sector and business sector (which tend to be closely-protected confidential files), or distribute surveys amongst the authorship of journals (large-scale sampling is a huge undertaking, with a large likelihood of suffering from high non-response rates). In both cases the outcome is likely to be subjective and unreliable. Alternatively, one could search other sources of bibliographical information with a bearing on a journal's 'applied/basic' profile.

One option is to compare a journal's 'industrial relevance' score with data extracted from the reference lists of patents. More specifically, patent references that are included in a patent examiner report, and 'cites' a publication in the scientific journal literature, where the journal is indexed by the *Web of Science*. One may safely assume that those journals that are highly cited within patents will relate to, or contributed to, the R&D underlying patented technologies. Hence, these so-called 'non patent references' (NPR) in patents can be used to empirically validate the 'industrial relevance' dimension, thus introducing the hypothesis: *a significant and positive correlation should exist between the share of corporate authors contributing research papers to a journal and the likelihood that research papers within that same journal are cited in the reference lists of patents.*

The NPRs were extracted from the reference lists in all EPO and UPSTO patents granted during the years 1999-2005.[5] 'NRP intensity' is defined the quantity of NPRs to publications in a specific journal divided by the total quantity of publications in that journal within that same time-interval (i.e. 1999-2005). The Pearson correlation across all 11 558 journals between 'industrial relevance' and NPR intensity equaled r=0.22. However the vast majority of the WoS-indexed journals will never be cited within patents, especially the 'basic' journals within fields of science that are disconnected from technological development, notably those within the social sciences and humanities. Hence, the scope of comparison was restricted to those fields of science (i.e. Thomson Reuters *Journal Categories*[6]) that attracted a minimum number of NPRs. The minimum threshold was set at an average of 0.25%, i.e. the NPR count to all journals within a Journal Category relative to the total publication output of all journals in that same category. The distribution of NPRs across journals is highly skewed, both within

---

5 The PATSTAT database (produced by the *European Patent Office*) was used as information source. The matching of NPRs to research papers in WoS-indexed was done by CWTS, with CWTS proprietary software, as a spin-off of an EC/IPTS-ERAWATCH project (EWN-EX24) conducted in 2009 (Tijssen and Van Looy, 2009).

6 These subject categories are imperfect representations of scientific fields, initially designed for, and primarily meant for, information retrieval. They are often very fuzzy defined, and the selected set of journals is neither a complete nor coherent set of all journals that are relevant to the field (e.g. Boyack et al., 2005; Leydesdorff and Rafols, 2009).

and between Journal Categories. The top three journals are *Journal of Biological Chemistry* (assigned to the Journal Category 'Biochemistry & Molecular Biology'), *Applied Physics Letters* ('Applied Physics'), and *Electronics Letters* ('Electrical & Electronic Engineering').

The aggregate result at the level of 31 'technology-related' Journal Categories is displayed in Table 1. These results clearly indicate a statistically significant positive correlation in most of these fields. However, the correlations coefficients vary amongst the fields, where a few fields show no significant correlation at all, while two fields actually exhibit a significant negative correlation.[7] On the whole, these findings confirm the hypothesis that the 'Industrial relevance' indicator reflects the degree of technologically relevant contents of journals.

Unfortunately, no large scale external sources exist, comparable to the patents database, to cross-validate the clinical relevance of journals or the corresponding fields as a whole. The prime candidate for such source would be a worldwide bibliographic database which contains the cited research literature that was extracted from the footnotes or reference lists within professional journals for medical practitioners and within clinical guidelines (e.g. Lewison and Wilcox-Jay, 2003).

The fields with the largest shares of corporate researchers publishing in journals are typically those in which scientific research and new technologies are closely related. These are the fields in which those corporate researchers are likely to be engaged in knowledge transfer from their global research environment to their intramural domains of industrial/technological applications. The number 1 industrial relevant Journal Category is *Petroleum Engineering*, a tiny field within the WoS (nine journals only), with a 33% share of author addresses referring to business enterprises and private sector organizations.

As noted with regard to Figure 1, the industrial application and clinical application domains are intricately linked within the pharmaceutical sector. This is where clinical research conducted within hospitals meets industrial R&D devoted to the development of new therapeutic drugs and innovative treatments.

7 Most notably, the internationally 'hot' field of Nanoscience & Nanotechnology; each of the five journals attributed to this Journal Category have relatively large shares of corporate sector authors and NPRs, but not in any linear relationship. We may assume a positive correlation will emerge when a larger number of nano-journals are assigned to this JC.

TABLE 1. Validating the industrial relevance of science fields: Pearson correlation coefficients (R) between "industrial relevance" and "NPR intensity".

Field of science (Journal Category)	R	Industrial relevance (%)
Biochemical Research Methods	0.97	11
Manufacturing Engineering	0.95	16
Computer Science - Information Systems	0.68	12
Medicine - Legal	0.66	13
Applied Chemistry	0.53	14
Research & Experimental Medicine	0.52	4
Medicinal Chemistry	0.46	20
Materials Science - Coatings & Films	0.45	19
Physics - Atomic, Molecular & Chemical	0.42	3
Computer Science - Artificial Intelligence	0.39	8
Pharmacology & Pharmacy	0.39	10
Computer Science - Theory & Methods	0.38	9
Mathematical & Computational Biology	0.34	5
Acoustics	0.27	9
Biochemistry & Molecular Biology	0.25	4
Electrochemistry	0.22	7
Computer Science - Software Engineering	0.20	13
Computer Science - Hardware & Architecture	0.19	16
Genetics & Heredity	0.15	3
Cell Biology	0.14	2
Biotechnology & Applied Microbiology	0.10	10
Immunology	0.09	4
Electrical & Electronic Engineering	0.06	22
Materials Science - Multidisciplinary	0.04	10
Optics	0.04	10
Imaging Science & Photographic Technology	0.00	17
Virology	-0.01	3
Organic Chemistry	-0.06	6
Telecommunications	-0.06	23
Nanoscience & Nanotechnology	-0.14	18
Biomedical Engineering	-0.17	5

TABLE 2. Application-oriented fields of science.

Fields (Journal Category)	Industrial relevance (%)	Clinical relevance (%)
Top 10 Industrial-oriented		
Petroleum Engineering	33	0
Telecommunications	23	0
Electrical & Electronic Engineering	22	0
Medicinal Chemistry	20	2
Materials Science - Coatings & Films	19	0
Nanoscience & Nanotechnology	18	0
Imaging Science & Photographic Technology	17	1
Computer Science - Hardware & Architecture	17	0
Manufacturing Engineering	16	0
Energy & Fuels	15	0
Top 10 Clinical practice-oriented		
Emergency Medicine	0	35
Orthopedics	1	27
Critical Care Medicine	0	24
Surgery	1	23
Pediatrics	0	23
Rheumatology	1	22
General & Internal Medicine	1	22
Medical Laboratory Technology	3	21
Urology & Nephrology	1	21
Gastroenterology & Hepatology	1	21
Top 5 Industrial/clinically-oriented		
Pharmacology & Pharmacy	10	9
Materials Science - Biomaterials	7	5
Toxicology	9	5
Biochemical Research Methods	11	4
Medicine - Legal	13	4

Table 2 displays the list of top 10 most 'clinically relevant' fields with the highest fraction of addresses to the medical sector; the top 10 'industrially relevant' fields collectively constitute the core of 'transfer sciences' linking science to technological innovation; and the top 5 fields with high scores on both 'clinical relevance' and 'industrial relevance', where we find the fields that play a major role in transferring scientific knowledge, tools and related skills to industrial-medical domain. The 'clinical practice' oriented journals have as much as 20% of their authors from hospitals and clinics. On the whole, the contribution from the medical sector to WoS-indexed journals in the clinically relevant fields is slightly higher than the contribution of corporate researchers in the industrially relevant fields. Note than the 20% share is a conservative estimate since all authors who are faculty members at academic medical centers and university teaching hospitals are not included in the medical sector (these institutions are treated as an integral part of a university system and their publications are classified as 'academic').

Journal Application Domains

We have now moved from the crude two-category 'basic/applied' classification system, via Stoke's three-category system, to CHI Research's elaborate four-category system. Clearly, more differentiation is needed to accommodate the observed diversity within the scholarly journal literature. The statistics and general patterns presented in the previous analyses constitute an evidence-based objective framework for systematically classifying journals according to knowledge application domains. The shape of the distributions depicted in Figure 1 suggests the following six mutually exclusive sets of journals – *Journal Application Domain (JAD)* categories:

A Academic journals: very few or no contributions from industry and private sector organizations, nor from general hospitals and medical centers;

I Industry relevant journals: some contributions from industry;

I^+ Industry practice journals: many contributions from industry;

C Clinical relevant journals: some contributions from general hospitals;

C^+ Clinical practice journals: many contributions from general hospitals;

I-C Industry-clinical relevant journals; some contributions from both industry and general hospitals.

Any data reduction process of numerical scores into such a small number of JAD categories is deemed to introduce a certain measure of arbitrariness and statistical 'noise'. To counter this effect, one should determine the cut-off points conservatively, i.e. reserving the extreme categories for those journals that differ significantly from the mean or median, while the intermediate categories are sufficiently broad to accommodate the wide variety at the center of the distribution. The distributional statistics in Figure 2 and Table 2 suggest the six categories as listed in Table 3, with corresponding integer value cut-off points. These thresholds (3%, 14% and 20%) were validated in an *ad hoc* fashion by inspecting the journal titles, where titles referring to application-oriented research topics and areas were used as a marker. The summary statistics of each category are displayed in Table 4, along with the number of journals they contain.

TABLE 3. Journal Application Domain (JAD) Categories: descriptors and value range specification.

JAD	Descriptor	Industrial relevance (%)	Clinical relevance (%)
A	Academic journals	<3	<3
I	Industrial relevant journals	3-14	<3
I^+	Industry practice journals	>14	<3
C	Clinical relevant journals	<3	3-20
C^+	Clinical practice journals	<3	>20
I-C	Industry-clinical relevant journals	>3	>3

TABLE 4. Summary statistics of JAD Categories.

		Industrial relevance (%)			Clinical relevance (%)		
JAD	WoS journals	mean	median	stan. dev.	mean	median	stan. dev.
A	6 621	1.1	0.6	1.3	0.5	0.0	1.1
I	1 864	11.0	9.4	5.1	0.5	0.0	1.1
I^+	291	39.4	34.7	14.5	0.3	0.0	0.7
C	2 083	1.3	0.9	1.2	13.7	13.5	5.7
C^+	434	0.9	0.4	1.5	32.7	29.8	8.7
I-C	265	8.9	7.6	4.3	11.3	9.9	5.4

Case study

This particular application of the JAD classification system addresses one of the major issues in science policy and in science studies: is contemporary global science becoming more application-oriented? In their study of UK science, Hicks and Katz (1996) argued that the locus of knowledge production might be shifting in the direction of application domains. Based on their analysis of the *Science Citation Index* database (now included in the *Web of Science*), they observed growth rates in publication output above the UK average rate, both by corporate sector authors and by those within medical sector, yet no conclusive indications were found of increased application-orientation within UK science as a whole (Hicks and Katz, 1996).

Appling the JAD system to the WoS enables a check of their hypothesis at the global level. If their claim holds, an increasing share of the global publication output in recent years would be distributed through 'application-oriented' journals and an increasing fraction would finds its way to 'academic journals'. To exclude the effects of database changes, the trend analysis in constrained to a fixed subset of 6 387 journals that published papers in each year during time period 1999-2008.

Table 5 lists the trends in the share of each JAD category within the WoS total annual output. As to be expected, the largest share of knowledge production represented in WoS-journals relates to research findings that were published in 'academic journals'. However, this fraction has slipped noticeably over the last 10 years - the 2% annual growth rate of this category is below the growth rate of the WoS overall. Global science, as represented in by the authorship of WoS-indexed journals, seems to be becoming slightly less 'academic' and more 'application oriented'. A very visible upward trend is found within 'industry relevant science' (I), which has gradually increased its share in worldwide publication output from 20.8% to 23%. Interestingly, a downward trend is found in application-oriented science represented in 'industry practice journals' (I^+), while the fraction of output in 'industry-clinical relevant journals' (I-C) also shows a gradual decline in recent years. Note that these declines are relative; as the WoS continues to expand, the absolute publication output of I^+ journals has how leveled off, after growth until 2006. The absolute numbers of publications in I-C journals decreased for the first time in 2008.

TABLE 5. Trends within the total WoS-indexed publication output per JAD.

	A	I	I+	C	C+	I-C	Total output
Annual trends in share of category (%)							
1999	31.9	20.5	5.4	19.3	9.6	13.3	965 319
2000	32.7	20.8	5.3	19.4	9.8	12.0	978 224
2001	32.4	21.4	5.3	18.8	10.1	12.0	968 213
2002	32.1	20.8	5.0	19.9	10.0	12.3	1 004 564
2003	31.6	21.5	5.4	19.2	10.1	12.3	1 017 459
2004	30.4	21.8	5.1	20.4	9.7	12.7	1 068 419
2005	30.3	21.9	5.2	20.1	10.1	12.4	1 096 571
2006	29.9	22.3	5.3	20.2	10.0	12.3	1 147 780
2007	30.2	22.7	5.1	19.4	10.3	12.1	1 179 017
2008	31.1	23.0	5.1	19.4	9.7	11.7	1 175 900
Mean Annual Growth (%)							
1999-2008	2.0	3.5	1.7	2.4	2.4	0.9	2.2

These subtle shifts suggest a redistribution of publication output within the industrially relevant research domains, where research publication output in journals closest to science-based technological development (Stoke's *Pure applied research*) are very slowly loosing ground to journals that are less oriented to industrial research (Stoke's *Use-inspired basic research*). This 'hidden evolution' within the global research literature mirrors trends in corporate R&D expenditures and research publication output volumes indicating that science-based industries have been gradually scaling down the level of resources devoted to in-house scientific research (Godin, 1996; Varma, 2000; Tijssen, 2004). The observed trends in publication output may also arise (in part) from changes in publication habits and strategies, where corporate research staff is publishing less about their research findings within WoS-indexed journals (all else remaining equal).

The distribution of JAD categories across all fields of science worldwide is shown in Figure 4. This graphical presentation, produced by *VOSviewer*, dispicts the relational network of WoS-indexed journals, where inter-journal distances are calculated according to their co-citation similarities. The graph comprises of the subset of 5 000 WoS-indexed journals that are most highly cited (Van Eck & Waltman, in press). The JAD-dependent color coding of each journal is as follows:

Academic (A) - grey; Industrial relevant (I) – light blue; Industry practice (I+) – dark blue; Clinical relevant (C) – light red; Clinical practice (C+) – red; Industry-clinical relevant (I-C) – pink. World science, as labeled by JAD categories, clearly shows a divide between clinical relevance to the right (within medical and life sciences), and industrial relevance to the left (natural sciences and engineering sciences). The interface between both application domains, industrial-clinical relevance (biopharmaceutical relevant fields), is also found at the right hand side of the graph where it is closely tied to the medical and life sciences. The corresponding interactive *VOSviewer* map, which can be accessed at http://www.vosviewer.com/maps/journal_application_domains/, provides more graphical detail, while allowing users to check the clinical relevance score (% general hospital addresses) and industrial relevant score (% corporate addresses) of each journal.

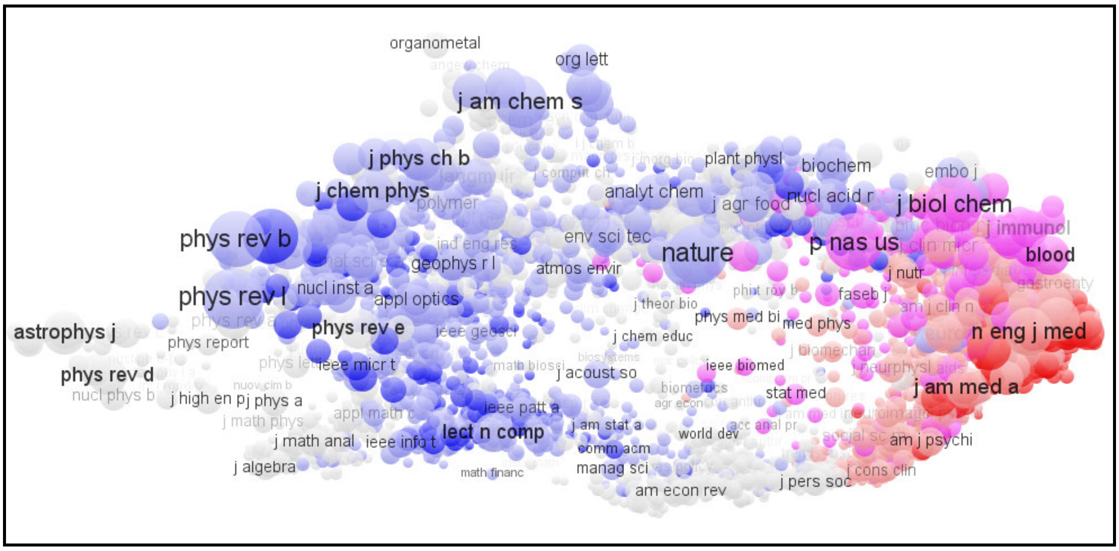

FIG 4. Graphical representation of WoS-indexed journals by Journal Application Category

Clearly, these macro-level findings should be treated with due caution. It remains unclear to what extent journal- or field-specific factors are also at play, such as changes within the editorial policies of journals. And it goes without saying that these findings are platform-dependent and method-dependent. The choice of the *Web of Science* as a frame of reference, and the choice of parameter values for defining the six JAD categories, will determine the outcomes. A different version of the WoS, or another comparable source such as SCOPUS may produce a (slightly) different outcomes. Further applications of this journal classification system, using other or additional information sources, will shed more light on its statistical robustness and its validity for macro-level analyses of science.

Discussion and conclusions

Scientific research constitutes a complex, multifaceted activity - its dimensions are many, several of which are time-dependent and context-dependent. However, complexity need not breed mystery, because relevant dimensions are clearly observable and some are, albeit crudely, measurable in terms of comparative statistics. The model and method presented in this paper taps into one of those measurable dimensions: it introduces a systematic evidence-based classification system of research publication outputs that helps unravel the diversity within science in terms of major application domains. However, the usage of research publications, and institutional sectors of their authors, for capturing and codifying diversity is merely a crude approximation of real life and therefore subject to justified criticism.

The taxonomic principle of the JAD classification system rests critically on two basic assumptions: (i) an entire journal can be assigned to a single JAD category reflecting its major application domain(s); (ii) these domains are reflected in the institutional sector(s) and associated working environments of the authors publishing in the journal. At macro levels of analysis both are likely to hold, i.e. across large sets of journals representing science as a whole, or within fields of science. At meso and micro levels, i.e. within journals and among authors, these assumptions may break down as the degree of variance within the data increases.

Research publications within a journal tend to reflect different stages within knowledge creation and utilization process, and may include research papers from authors covering a variety of working environments and application domains. Obviously the authors may also differ quite a lot in terms of their motives and rationale for doing science and disseminating results in their publications, as do (end)users vary in their the goals and intentions with regards to applications of those findings and outcomes. The institutional profile of a journal according to the affiliate background of its authorship is therefore at best a journal-dependent partial indicator of the associated working environments, application domains and user communities.

The definition of an institutional sector itself depends heavily on the conceptual framework and delineation principle. For example, adopting a broader notion of the 'medical sector', one including academic medical centers, would significantly increase the numbers of 'clinically relevant' journals. And redefining the corporate sector, by for example excluding privately-funded industrial research organizations, will diminish the quantity of 'industrially relevant' journals.

Applying the JAD classification system only to peer-reviewed journals discards relevant segments of the research publication literature. Peer-reviewed scholarly journals, especially those that are indexed for the *Web of Science*, are biased toward English-language 'main stream' research, is comprised of (open access) journals whose business models and editorial policies are focused on wide, international dissemination. In contrast, the published outputs with a view toward practical, short-term applications may also find their way to the open literature through the specialized serials with limited circulation (report series, professional and trade journals), take the form of proprietary documents (patents), or remain restricted access documents (confidential reports).

In spite of these methodological constraints, the JAD classification system represents a significant improvement, at least from an analytical point of view, compared to other measurement models or earlier classification systems that lacked a sufficiently transparent methodology or a sound foundation of empirical data. The omission of a journal classification parameter related to the 'civic relevance' application domains constitutes a noticeable shortcoming in the current classification system, especially so because the societal relevance of science is now high on many research policy agenda's, and an increasingly important component of research evaluation frameworks linked to evidence-driven funding of science. Incorporating 'civic relevance' domains will require considerable investments in data mining of existing bibliographic databases (local or disciplinary), or developing dedicated new databases, that capture the publication output of these researchers and their user communities worldwide.

The 'Knowledge Utilization Triangle' model, its classification system and the precise definition of the JAD categories, is of course open to further research, debate, experimentation and improvement. More methodological work lies ahead of us to develop it into fine-grained and generally-accepted approach. The longer term and more ambitious R&D agenda is to apply such a system at the source of the 'basic science/applied science' dichotomy: the statistical surveys by government authorities and international agencies, and their reports on the state of science.

Acknowledgments

Thanks are due to Arnold Verbeek (IdeaConsult, Brussels), Erik van Wijk, Renald Buter, Nees Jan van Eck, and Ludo Waltman (CWTS, Leiden University) for their valuable data inputs and research assistance. I would also like to thank the reviewers for their insightful comments and suggestions.

References

Bailey, K. (1994). Typologies and taxonomies: an introduction to classification techniques. Thousand Oaks: Sage Publishers.

Boyack, K., Klavans, R., & Börner, K. (2005). Mapping the Backbone of Science. Scientometrics, 64, 351-374

Brooks, H. (1980). Basic and applied research. In: Categories of Scientific Research (pp. 14-18). Washington DC: National Science Foundation.

Brusoni, S., & Geuna, A. (2003). An international comparison of sectoral knowledge bases: persistence and integration in the pharmaceutical industry, Research Policy, 32, 1897-1912.

Falk, C. (1973), An operational, policy-oriented research categorization scheme, Research Policy, 2, 186-202.

Hamilton, K. (2003). Subfield and Level Classification of Journals, CHI Report No. 2012-R. Cherry Hill, NJ: CHI Research Inc.

Godin, B. (1996). Research and the practice of publication in industries, Research Policy, 25, 587–606.

Godin, B. (2003). Measuring science: is there 'basic research' without statistics? Social Science Information, 42, 57-90.

Hicks, D., & Katz, J.S. (1996). Where is science going? Science, Technology & Human Values. 21, 379-406.

Hicks, D., & Hamilton, K. (1999). Real numbers, Issues in Science and Technology. 15, 74-75.

Langenberg, D. (1980). Distinctions between basic and applied research. In: Categories of Scientific Research (pp. 32-36). Washington DC: National Science Foundation.

Lewison, G., & Paraje, G. (2004), The classification of biomedical journals by research level. Scientometrics, 60, 145 –157.

Lewison, G., & Wilcox-Jay, K. (2003), Getting biomedical research into practice: the citations from UK clinical guidelines. In Gohua, J., Rousseau, R., & W. Yishan (Eds.), Proceedings of the 9th International Conference on

Scientometrics and Informetrics (pp. 152–160). Beijing, China: Dalian University of Technology Press.

Leydesdorff, L., & 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, 348-362.

Lim, K. (2004). The relationship between research and innovation in the semiconductor and pharmaceutical industries (1981–1997). Research Policy, 33, 287-321.

Narin, F., & Rozek, R. (1988). Bibliometric analysis of bibliometric analysis of United States pharmaceutical industry research performance. Research Policy, 17, 139-154.

Noma, E. (1986). Subject Classification and influence weights for 3,000 journals. Cherry Hill, NJ: CHI Research/ Computer Horizons Inc.

OECD (2002). Frascati Manual, Paris: Organisation for Economic Cooperation and Development.

Reagan, M. (1967). Basic and applied research: a meaningful distinction? Science, 155, 1383-1386.

Rothschild, L. (1972). Forty-five varieties of research (and development), Nature, 239, 373-378.

Stokes, D. (1997). Pasteur’s Quadrant: Basic science and technological innovation. Washington DC: Brookings Institutional Press.

Tijssen, R.J.W. (2004). Is the commercialisation of scientific research affecting the production of public knowledge? Global trends in the output of corporate research articles, Research Policy, 33, 709-733.

Tijssen, R.J.W., Van Leeuwen, T.N., & Van Wijk, E. (2009). Benchmarking university-industry research cooperation worldwide: performance measurements and indicators based on co-authorship data for the world’s largest universities. Research Evaluation, 18, 13-24.

Tijssen, R.J.W., & Van Looy, B. (2009). Production of data to analyze university references found in patent documents by geographic origin. Third report for IPTS/ERAWATCH NETWORK ASBL (EC Framework Service Contract Nr - 150176-2005-F1SC-BE). Leiden University/Catholic University of Leuven.

Van Eck, N.J., & Waltman, L. (in press). Software survey: VOSviewer, a computer program for bibliometric mapping, Scientometrics.

Varma, R. (2000). Changing research cultures in US industry, *Science,* Technology and Human Values, 25, 395–416.