National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance

Janne Pölönen[1†], Raf Guns[2], Emanuel Kulczycki[3], Gunnar Sivertsen[4], Tim C. E. Engels[5]

[1]Federation of Finnish Learned Societies, Snellmaninkatu 13, 00170 Helsinki, Finland

[2]University of Antwerp, Faculty of Social Sciences, Centre for R&D Monitoring (ECOOM), Middelheimlaan 1, 2020 Antwerp, Belgium

[3]Adam Mickiewicz University, Scholarly Communication Research Group, Szamarzewskiego 89c, 60-568 Poznań, Poland

[4]Nordic Institute for Studies in Innovation, Research and Education (NIFU), P.O. Box 2815,0608 Tøyen, Oslo, Norway

[5]University of Antwerp, Faculty of Social Sciences, Centre for R&D Monitoring (ECOOM), Middelheimlaan 1, 2020 Antwerp, Belgium

Abstract

Purpose: This paper presents an overview of different kinds of lists of scholarly publication channels and of experiences related to the construction and maintenance of national lists supporting performance-based research funding systems. It also contributes with a set of recommendations for the construction and maintenance of national lists of journals and book publishers.

Design/methodology/approach: The study is based on analysis of previously published studies, policy papers, and reported experiences related to the construction and use of lists of scholarly publication channels.

Findings: Several countries have systems for research funding and/or evaluation, that involve the use of national lists of scholarly publication channels (mainly journals and publishers). Typically, such lists are *selective* (do not include all scholarly or non-scholarly channels) and *differentiated* (distinguish between channels of different levels and quality). At the same time, most lists are embedded in a system that encompasses multiple or all disciplines. This raises the question how such lists can be organized and maintained to ensure that all relevant disciplines and all types of research are adequately represented.

Research limitation: The conclusions and recommendations of the study are based on the authors' interpretation of a complex and sometimes controversial process with many different stakeholders involved.

† Corresponding author: Janne Pölönen (E-mail: janne.polonen@tsv.fi).

Research Paper

Practical implications: The recommendations and the related background information provided in this paper enable mutual learning that may feed into improvements in the construction and maintenance of national and other lists of scholarly publication channels in any geographical context. This may foster a development of responsible evaluation practices.

Originality/value: This paper presents the first general overview and typology of different kinds of publication channel lists, provides insights on expert-based versus metrics-based evaluation, and formulates a set of recommendations for the responsible construction and maintenance of publication channel lists.

Keywords   Publication channel lists; Research funding; Scholarly communication; Journal ranking

1   Introduction

This paper provides an overview of scholarly publication channel lists and contributes with a set of recommendations for the construction and maintenance of national lists of scholarly journals and publishers in order to safeguard a balanced representation.

A scholarly publication channel has distinct editorial standards and procedures regarding peer-review and decision-making that all the outputs—articles and books—published in the channel have undergone. The most important and typical kinds of scholarly publication channels are journals and book publishers and their imprints, although other types also exist (e.g. book series, conference proceedings series).

Since the establishment of the first peer-reviewed journals in the 17th century, there has been an immense growth in the number of publication channels specializing in publishing research results (de Solla Price, 1963; Haustein, 2012; Houghton, 1975). Globally, there may be currently over 70,000 academic/scholarly journals (Johnson, Watkinson, & Mabe, 2018). Before the emergence of journals, research results were published in letters and books. Book publishing continues to be important, especially in the social sciences and humanities (SSH) (Engels et al., 2018). Estimates vary, but certainly dozens of thousands of book publishers and imprints are involved internationally and locally in publishing research results in the form of monographs and articles in edited volumes (Giménez-Toledo, Mañana-Rodríguez, & Sivertsen, 2017; Giménez-Toledo et al., 2019).

Efforts to make sense of the number and diversity of scholarly publication channels started relatively early, mostly with a focus on journals. Already in the late 19th century, the Royal Society of London listed scholarly journals, as distinct from professional journals, for the purpose of producing the *Catalogue of Scientific Papers* published globally (Csiszar, 2017). Research libraries have also had an increasing interest, from the point of view of collection management, to list and

National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

prioritize academic/scholarly journals (Nisonger, 1988). The purpose of the first journal ranking produced in 1926 was to determine, based on citations, which chemistry journals were indispensable for a university library with scarce resources (Gross & Gross, 1926). Ulrich's Periodicals Directory, started in 1932, is the most elaborate library directory of over 300,000 serials, including peer-reviewed journals. The International Standard Serial Number (ISSN) has been used since 1975 to identify serial publications—including journals and series of books and proceedings — and issued globally to over 2,000,000 titles.

In 1964, the Institute for Scientific Information (ISI) introduced the Science Citation Index (SCI) of cited references and publications in a selected group of international peer-reviewed journals. The SCI and later sibling citation indexes like the Social Science Citation Index (SSCI), the Arts & Humanities Citation Index (AHCI), and the Emerging Sources Citation Index (ESCI) are nowadays part of the Web of Science (WoS), owned by Clarivate Analytics. Including all four journal lists, WoS currently covers over 21,000 journals. Since 1975, ISI has also published the Journal Citation Reports (JCR), introducing the Journal Impact Factor (JIF) and other metrics that currently rank over 12,000 journals included in SCI and SSCI based on citations. In 2004, Elsevier launched Scopus, a competing index of publications and cited references currently covering almost 23,000 journals from all fields, adding also a suite of citation-based journal metrics. The journal lists of WoS and Scopus are often regarded as the standard lists of qualified international peer-reviewed journals, while journal metrics are frequently used to differentiate, prioritize and rank these journals in specific subject categories.

It has been well-established in bibliometric research, however, that WoS and Scopus cover only a relatively small share of all peer-reviewed publications and their channels, and that there is considerable variation in their representation of research produced in different fields and countries (Archambault et al., 2006; Giménez-Toledo, Mañana-Rodríguez, & Sivertsen, 2017; Hicks, 1999; Hicks & Wang, 2011; Kulczycki et al., 2018, 2020; Larivière & Macaluso, 2011; Nederhof, 1989, 2006; Ossenblok, Engels, & Sivertsen, 2012; Sivertsen & Larsen, 2012; Sivertsen, 2016). There are two main reasons for this. Firstly, to have success on the market, these products not only depend on the coverage, but also the quality and international relevance of their contents, as well as on their production costs. Citation indexing inherits a tradition in which Eugene Garfield (1979) demonstrated that information retrieval theory (Bradford's law of scattering) and citation analysis support the idea of indexing mainly the "core journals" of international interest (Aksnes & Sivertsen, 2019). However, many peer-reviewed journals are entirely, or to some extent, locally oriented in terms of authorship, readership and scope, and thus may be less visible internationally and less frequently cited in international

journals. Consequently, most journals are not included in WoS and Scopus. This is especially common in the SSH and for journals in other languages than English. Secondly, in all fields— but especially in computer science, engineering and SSH— research results are also communicated through other channels, such as peer-reviewed conference proceedings and books. Although both WoS and Scopus also index conference proceedings and books, their coverage of these publication types remains weak in the social sciences and humanities where books are most important (Aksnes & Sivertsen, 2019).

Many institutions rely on the readily available international WoS and Scopus lists of journals, as well as the related journal metrics, in funding, assessment and evaluation procedures. China is an example of a country where WoS-based indicators (Journal Impact Factors, JCR Quartiles, and ESI Highly Cited Papers) have been used at all levels in research evaluation, staff employment, career promotion, awards, university or disciplinary rankings, funding, and resource allocation (Zhang & Sivertsen, 2020). According to a recent survey, around 40% of 129 research intensive institutions in the United States and Canada mentioned impact factors in documents relating to review, promotion, and tenure processes (McKiernan et al., 2019). Also in Europe, 75% of 186 universities responding to the European University Association survey used Journal Impact Factor to evaluate research careers (Saenen et al., 2019).

These practices have prompted strong criticism from the research community. It has been shown that the Journal Impact Factor has serious deficiencies as a tool for assessing the quality of individual outputs (Adler, Ewing, and Taylor, 2008; Amin & Mabe 2000; Seglen, 1997; Zhang, Rousseau, & Sivertsen, 2017). Already in 2012, the San Francisco Declaration on Research Assessment (https://sfdora.org) highlighted the need to assess research on its own merits rather than on the basis of the journal in which the research is published: "Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions." There is also a broader international campaign promoting more responsible use of metrics in research evaluation (Hicks et al., 2015). The same trend is expressed in a recent reform of research evaluation and funding in China that turns away from what has been called "SCI worship" in the country for a long time (Zhang & Sivertsen, 2020).

The demands for a more responsible evaluation culture are highly relevant also regarding the development and use of publication channel lists more generally. These demands cover many other aspects than using journal hierarchies to assess individual articles (Wilsdon et al., 2015):

National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

- Robustness: basing metrics on the best possible data in terms of accuracy and scope;
- Humility: recognizing that quantitative evaluation should support—but not supplant—qualitative, expert assessment;
- Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results;
- Diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system;
- Reflexivity: recognizing and anticipating the systemic and potential effects of indicators, and updating them in response.

Furthermore, the unit of assessment is an important dimension of responsible use of metrics: does assessment concern individual researchers and research groups, departments and faculties, institutions or countries? (Glänzel & Wouters, 2013; Verleysen & Rousseau, 2017). The purpose of the assessment should be considered: is it research evaluation in the sense of learning and improvement and/or funding allocation? (Molas-Gallart, 2012; Sivertsen, 2017). It is also important to consider that citation-based impact factors are not the only means of assessing the quality, impact or prestige of journals and other publication channels. The traditional means of journal assessment also include expert evaluation, both in the form of surveys and expert panel assessment (Ahlgren & Waltman, 2014; Haddawy et al., 2016; Kulczycki & Rozkosz, 2017; Saarela et al., 2016; Saarela, & Kärkkäinen, 2020; Serenko & Dohan, 2011; Walters, 2017). More recently, Wouters et al. (2019) called for a broader and more transparent suite of journal metrics.

This study is structured as follows: first, we present an overview of various publication channel lists on the international, national, and local level. Next, we discuss the ongoing debate on journal evaluation at the national level, using experiences from the Nordic countries as an example. We conclude with a set of recommendations and suggestions for the construction, maintenance, and future development of national lists of scholarly journals and publishers.

2 Typology and overview of publication channel lists

Publication channel lists have been started and are used for different purposes. Consequently, such lists may also have different characteristics. We provide the following typology, which may be useful to describe the most salient dimensions by which publication channel lists can be differentiated.

- *Geographic scope*. A list may be used in an international, a national, or a local context. Note that this refers to the purpose of the list rather than the nature of the channels on it: most national lists also contain international channels and channels from other countries.
- *Selectivity*. This refers to the question if a publication channel list can include all publication channels, at least in theory, or if some inclusion criteria relating to quality or quality assurance processes like peer review are present. In practice, almost all publication channel lists are selective, although the degree of selectivity (the amount and rigour of criteria) may be different.
- *Differentiation*. Many publication channel lists differentiate between publication channels of different levels or classes. Such levels or classes reflect the channels' quality, prestige, internationality or similar aspects. They may be based on expert judgment, one or more bibliometric indicators, or a combination thereof.
- *Composition*. Some publication channel lists are composite, in that they treat groups of publication channels (e.g. local versus international channels, WoS-indexed versus non-indexed journals) in a different way. Other lists are unitary and treat all publication channels uniformly.
- *Ex ante/ex post*. Ex post lists rely on a set of publications, such as all publications in a national database, and only consider those channels for inclusion that are associated with at least one publication in the set. Ex ante lists compile an overview of as many publication channels as are deemed relevant in the context in which they are established.
- *Field coverage*. Some lists aim to cover all fields of research, whereas others deliberately focus on one or a few fields.

Below, we provide an overview of publication channel lists. First, we describe national lists used as tools in research evaluation or performance-based research funding systems. Second, some international lists are characterized. Finally, other international, local and field specific initiatives are presented.

2.1   National lists

During the past two decades, ministries in several European countries have established performance-based research funding systems (PRFS) for the purpose of allocating part of annual core-funding from the government to universities based on bibliometric indicators and other indicators of contributions to research and higher education (de Boer et al., 2015; Hicks, 2012; Jonkers & Zacharewicz, 2016; Sivertsen, 2017; Zacharewicz et al., 2018). Poland established its PRFS in 1991 and started to publish a national list of journals in 1999 (Kulczycki & Rozkosz, 2017). In 2005, Norway introduced a PRFS based on a fixed funding formula, in which

National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

the entire research publication output of the universities from all fields is weighted according to publication type and an expert-based quality rating of journals/series and book publishers as indicated in a comprehensive authority list of publication channels (Schneider, 2009; Sivertsen, 2010; 2018b). Denmark in 2009 and Finland in 2012–2015 have adopted the Norwegian model for all fields.

All three countries use some combination of 2-4 level categories to indicate differentiation between the basic peer-reviewed (level 1) and leading channels (2, 3) according to quality, impact and/or prestige. Some lists also indicate not approved channels (level 0). The assignment of channels to levels is based on expert-evaluation informed, but not constrained, by journal metrics (Aagaard, 2018; Pölönen, 2018; Sivertsen, 2016b; 2017; 2018b). The number of panels and experts differ among the three countries (see Table 1). These three lists can be described as unitary rather than composite, in the sense that they form a single entity with uniform quality rating. They are also produced ex ante, including also publication channels where researchers affiliated with the country's universities have not yet published. All of them were designed to be applied at macro-level, i.e. the unit of assessment is a university, not a faculty/department or an individual researcher.

- *Norway*. The Norwegian Register for Scientific Journals, Series, and Publishers is managed by The National Board of Scholarly Publishing (NPU) and operated by the Norwegian Centre for Research Data (NSD). As of 30 March 2020, the

Table 1.   Organisation of the publication channel lists in Denmark, Finland, and Norway.

		Denmark	Finland	Norway
Organization	Established	2009	2010	2005
	Full-time personnel	1-2.5	2	2
	Expert-evaluators	429	250	331
	Panels	67	23	89
Jourals/series level quotas	Levels	1, 2, 3	0, 1, 2, 3	0, 1, 2
	Basis	World production	World production	World and national, production
	Level 2 share	17.5-22.5%	20%	20%
	Level 3 share	2.5%	5%	
Journals/series number of titles	Level 1-3	20,787	23,596	27,214
	Level 2-3	3,104	3,057	2,111
	Level 2-3 share	15%	13%	8%
Book publishers level quotas	Levels	1, 2	0, 1, 2, 3	0, 1, 2
	Basis	Estimated world production	Number of titles	National production
	Level 2-3 share	20%	10%	20%
Book publishers number of titles	Level 1-3	1,409	1,335	1,693
	Level 2-3	91	106	86
	Level 2-3 share	6%	8%	5%

register includes 35,113 journals/series and 3,215 book publishers that are evaluated and assigned in all fields to three quality level categories (1=normal, 2=high, 0=not peer-reviewed) by experts in 83 field-specific groups. The expert-groups are largely based on pre-existing National academic bodies established by Universities Norway (UHR, the Norwegian association of higher education institutions) for professional and administrative development. The rating of book publishers is decided by the NPU (Aagaard et al., 2014; Sivertsen, 2018b).

- *Denmark*. The BFI list of series and book publishers to support the Bibliometric Research Indicator (BFI) is administered by the Ministry of Higher Education and Science on the basis of recommendations from the 67 Expert Panels composed of researchers appointed by the Universities Denmark. The recommendations are managed and finally decided upon by an Academic Committee representing all eight universities and the major areas of research. The most recent 2018 list includes 20,788 journals/series and 1,410 book publishers assigned to three quality levels (1=normal, 2=leading, 3=top). Level 3 is used only in some fields, and the publication channels not meeting the level 1 criteria are excluded from the list. The book publisher ratings are decided by the Academic Committee (Aagaard, 2018; Sivertsen & Schneider, 2012).

- *Finland*. The Publication Forum list of journals/series and book publishers is produced by the Federation of Finnish Learned Societies (TSV), while CSC —IT Centre for Science is responsible for the technical maintenance of the database containing the register of publication channels. As of 30 March 2020, the list contains 29,604 journals/series/conferences and 3,370 book publishers assigned in all fields to four quality level categories (1=normal, 2=leading, 3=top, 0=other publication channels) by experts in 23 field-specific panels, established by TSV in 2010 for the sole purpose of the channel evaluation. The book publisher ratings are decided collectively by the panel chairs, based on a preliminary proposal of the SSH panel chairs (Auranen & Pölönen, 2012; Pölönen & Ruth, 2015; Pölönen, 2018).

In Poland, Flanders [Belgium] and the Czech Republic, the PRFS is supported with authority lists of publication channels that can be described as composite rather than unitary, in the sense that they are made up of several parts. WoS, Scopus and/or ERIH Plus indexed journals have a different status, sometimes dependent on the JIF or other journal metrics, compared to other journals or book publishers included in the list of peer-reviewed publication channels. These composite lists do not usually contain a differentiation expressed in terms of unitary quality levels or categories; however, the publication channels are differentiated in the PRFS model

National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

by means of the number of points the articles or books published in them generate. The part of the list that is not based on WoS, Scopus or ERIH PLUS, is produced ex post, including only channels where researchers affiliated with the country's universities have published.

- *Poland*. Since 1999, the Polish Journal Ranking (PJR) developed by the Ministry of Science and Higher Education consists of three lists: A = journals with JIF (i.e. covered by WoS), B = other Polish or foreign journals, and C = Journals in ERIH PLUS. The points given to articles in B-list journals are partly based on expert ranking recommendations, however with less points than articles in C and especially A list journals. In 2018, the rules for the journal list were changed and a single list is now based on WoS, Scopus, ERIH PLUS, and lists of Polish journals. Moreover, the Polish Journal Ranking has been complemented with a book publisher list established by experts and the list of conferences based on the DBLP Computer Science Bibliography and the Computing Research and Education Association of Australasia (CORE). Since 2019, the PJR and the book publisher lists are used for funding scientific institutions as well as in promotion procedures (Kulczycki, 2017; 2018; Kulczycki & Rozkosz, 2017; Kulczycki & Korytkowski, 2018).

- *Flanders*. In Flanders [Belgium], the publication database VABB-SHW (and authority lists of peer-reviewed journals and book publishers) was established in 2010 for the SSH fields, to complement a pre-existing PRFS publication and citation indicator for funding of universities based on publications and citations in WoS-indexed journals. The 2019 VABB-SHW list contains 13,640 journals, in which SSH researchers affiliated with Flemish universities have published between 2008 and 2017. Of these, 6,243 (46%) are fully or partially indexed in the WoS (excluding ESCI), and 7,397 (54%) are other journals with ISSN. The non-WoS journals are evaluated by an Authoritative Panel appointed by the Flemish Interuniversity Council (VLIR) in consultation with disciplinary subpanels of experts. The Panel has classified 4,503 journals as peer-reviewed and 2,894 as non-peer-reviewed (Engels & Guns, 2018; Verleysen, Ghesquière, & Engels, 2014).

- *Czech Republic*. The Czech Ministry of Education, Youth, and Sports distributes funding to universities partly based on publication points determined formally by JIF, inclusion in Scopus or ERIH PLUS, or in the authority list of peer-reviewed journals published in Czech. In this case also, the national authority list of Czech journals complements other lists (Good et al., 2015). The list does not include book publishers.

National evaluation agencies have also established authority lists of publication channels in France, Australia, Italy and Spain. They are both composite and unitary

Research Paper

ex ante lists covering all fields or just SSH, and they have been used either to inform expert-based assessment of research units (Australia and France), and/or to assess individual researchers in academic promotion procedures (Italy and Spain).

- *France*. In 2008, the Agence d'Evaluation de la Recherche et de l'Enseignement Superieur (AERES) published an authority list of peer-reviewed journals in the SSH to inform evaluation of research units. The list differentiated journals with three level categories (A, B, C), partly based on the ERIH list, however in 2010 the differentiation was abandoned in most fields. The list was used by AERES to determine actively publishing researchers for the purpose of evaluation of research units (Pontille & Torny, 2010a; 2010b; 2012).

- *Australia*. Since 2010, Australia's national research evaluation framework, the Excellence in Research for Australia (ERA), administered by the Australian Research Council (ARC), has partly relied on an authority list of peer-reviewed journals and conferences established by panels of experts. The original list, covering all fields, used in the ERA 2010 evaluation campaign was differentiated with four level categories (A*, A, B, and C). However, since 2012 an undifferentiated list has been used. The list was not employed in a fixed funding formula but the results were used to inform expert-evaluation of research units (Genoni & Haddow, 2009; Haddow & Hammarfelt, 2018).

- *Italy*. Since 2012, the Agenzia per la valutazione del sistema Universitario e della ricerca (ANVUR) has produced a list of peer-reviewed journals in the SSH for the purpose of assessing applicants' scientific outputs in the framework of Italy's National Scientific Habilitation procedure. The list contains a total of 21,679 Italian and foreign journals with indication of Class A journals based on internationality, which is determined by expert panels (Ferrara & Bonaccorsi, 2016). The criteria for internationality of journals include the use of international experts as reviewers, as well as at least one of the following: indexing in one of the major international databases; significant share of contributions from international authors; and significant share of contributions in languages relevant for the scientific debate in the field (ANVUR, 2019).

- *Spain*. Since 2006, the Fundación Española para la Ciencia y Tecnología (FECYT) has developed within the ARCE project a classification of peer-reviewed SSH journals published in Spain. Around 300 SSH journals published in Spain have obtained the FECYT Quality seal (Sello FECYT) based on 57 formal quality and impact criteria (de Filippo, Aleixandre-Benavent, & Elías Sanz-Casado, 2019). The FECYT list is included in the Clasificación Integrada de Revistas Científicas (CIRC), a composite list integrating SSH journals from various information sources, such as WoS, Scopus, ERIH PLUS and Latindex (Torres-Salinas et al., 2010). In CIRC, journals are differentiated in 5 categories

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

(A+, A, B, C, D) based on their international impact and visibility as measured mainly by impact metrics from JCR and Scopus. The FECYT seal is used as one of the criteria for admitting of highest quality Spanish SSH journals in category B. The purpose of these lists is to support evaluation agencies (CNEAI and ANECA) in assessing merit in the periodic evaluation of researchers.

In China several journal lists are in use. According to Huang et al. (2020) the most influential are: the Chinese Science Citation Database (CSCD, available within WoS and managed by Clarivate in collaboration with the Chinese Academy of Sciences); the Chinese Social Sciences Citation Index (CSSCI); the journal partition table (JPT); the AMI Comprehensive Evaluation Report (AMI); the Chinese STM Citation Report (CJCR); the "A Guide to the Core Journals of China" (GCJC); and the World Academic Journal Clout Index (WAJCI).

2.2   International lists

More comprehensive lists of peer-reviewed publication channels have been constructed and are maintained at international, national and institutional level. Their aim is to list peer-reviewed journals and/or book publishers in certain or all fields to promote SSH publishing (ERIH PLUS), Open access publishing (DOAJ), predatory journals (Cabell's) and regional journals (Latindex). The validation and evaluation of publication channels is typically carried out by experts in the field.

- *ERIH PLUS*. In the early 2000s, the European Science Foundation (ESF) started preparation and expert-panel consultation to produce the European Reference Index for the Humanities (ERIH), the purpose of which was to identify the most important national and international journals publishing in European languages in the humanities and certain social science fields (e.g. psychology, pedagogical and educational research, gender studies, as well as evolutionary and social anthropology, as defined by OECD field of science classifications). The aim was to increase the visibility of non-English publications and to provide a tool for research assessment at the aggregate level. When first published in 2007–2008, ERIH covered 5,172 journals differentiated in three categories (A, B, and C) according to the degree of international reputation (Pontille & Torny, 2010a; Román Román, 2010). The name of the current edition is ERIH PLUS, operated since 2014 by the Norwegian Centre for Research Data (NSD), supported by a network of country-experts, and it covers 7,812 peer-reviewed journals from all SSH fields but without A-B-C-differentiation (Lavik & Sivertsen, 2017; Sivertsen, 2019). Instead, there are six formal criteria for journals to enter the list, and these criteria are checked for each journal.

Research Paper

- *DOAJ*. Since 2003, the Directory of Open Access Journals established at Lund University, Sweden, has provided a community-curated list of peer-reviewed open access journals. DOAJ currently lists over 14,000 open access journals in all fields from over 131 countries publishing in 75 languages. In 2014, DOAJ introduced new tightened inclusion criteria, according to which all journals are reviewed and approved upon application by a group of voluntary associate editors as well as managing editors. There is no differentiation of journals according to impact, quality or prestige, however, a DOAJ Seal is a mark of adhering to editorial and publishing best practices (Olijhoek, Mitchell, & Bjornshauge 2016; Marchitelli et al., 2017).

- *Cabell's Journalytics and Predatory Reports*. Cabell's Journalytics lists over 11,000 peer-reviewed journals across all fields, whereas the Predatory Reports list over 13,000 journals with identified questionable practices according to 60 criteria (Eykens et al., 2019).

- *Latindex*. Established in 1995 by the Universidad Nacional Autónoma de México (UNAM), Latindex is a comprehensive register of scientific, technical-professional and scientific and cultural dissemination journals published in Portuguese or Spanish in Latin America, the Caribbean, Spain and Portugal. Since 2018, a new version called Catalog 2.0 currently includes 7,512 online journals, across all fields, that meet specified requirements including peer-review (Gregorio-Chaviano, 2018).

2.3   Other lists

Many institutions rely on more extensive publication channel lists than WoS and Scopus that are not based on impact factors. In Sweden, for example, several universities have adopted the Norwegian national publication channel list produced for the purpose of funding allocation to universities into their internal evaluation and funding procedures (Hammarfelt et al., 2016). The local use of the national authority lists of publication channels, produced to support funding schemes of universities in Denmark, Finland, and Norway, is attested also in all three countries (Aagaard et al., 2014; Pölönen & Wahlfors, 2016; Sivertsen & Schneider, 2012; Wahlfors & Pölönen, 2018). There are, however, also institutional publication channel lists produced by research organisations or their subunits. Publication channel list produced at University College Dublin is one well-documented example.

- *University College Dublin*. Since 2016, University College Dublin (UCD) has implemented the Output-Based Research Support Scheme (OBRSS) to award individual academic staff members based on number of publications and doctoral students. The scheme ranks publications according to a list of journals and publishers differentiated in three categories (0, 1, and 2). The list contains

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

over 2,500 book publishers and more than 43,000 journals across all fields integrating journals and classifications from the Danish, Finnish and Norwegian lists, as well as Scopus journals and metrics (Cleere & Ma, 2018).

Numerous field-specific journal rankings exist based either on citation analysis or surveys. These are typically published in field-specific journals or journals devoted to bibliometric and scientometric studies. In addition, there are also some internationally renowned field-specific lists based on expert-evaluation, such as the Nature Index in the natural sciences and the Academic Journal Guide by the Chartered Association of Business Schools (AJG).

- *Nature Index*. Since 2014, Nature Research has compiled a database of articles published in high quality journals in the field of natural sciences to assess research excellence and institutional performance. Journals are selected by two panels of independent academics, informed by a global survey of the wider research community. The original list contained 68 journals, and the current edition was expanded to 82 journals in 2018.
- *Academic Journal Guide*. Since 2009, the British Association of Business Schools (ABS) has published the Academic Journal Guide of journals in the field of business and management. The most recent 2018 edition contains 1,561 journals differentiated in five categories (4*, 4, 3, 2, 1) based on expert-evaluation informed with metrics.
- *Journal Quality List*. Since the late 90's, Anne-Wil Harzing, now at Middlesex University in London, compiles and regularly updates the Journal Quality List of journals in Economics, Finance, Accounting, Management and Marketing. It is a collation of rankings of 13 different sources. The 66th edition, published online on 15 February 2020, contains over 900 journals (https://harzing.com/resources/journal-quality-list).

There are also field-specific journal and book publisher ratings developed for institutional assessment, for example, of Dutch research schools:

- *WASS-SENSE*: SENSE Research School in the Netherlands developed a set of performance criteria in 2006 and constructed lists of journals (A, B, and C journals) and publishers (A, B, and C publishers). The ranking of publishers is evaluated yearly. In 2017, *WASS-SENSE* ranking list of publishers has been published for *the WASS and SENSE Dutch Graduate Schools* (http://www.sense.nl/organisation/documentation).
- *CERES*: CERES Research School for International Development of the Utrecht University has designed internal valuation tools for the SSH reseachers and managers. In the framework of this system, two lists have been published:

(1) the list of journals (A, B, C, D, E journals) which comprises of journals indexed in Wos and other academic and non-academic periodicals; (2) the list of book publishers (A, B, C, D, E journals) categorized on the basis of the visibility in Google Scholar (https://ceres.sites.uu.nl/about-the-valuation-system/).

2.4   Criticism of publication channel lists

In this section, we highlight some of the key issues that led to the abandonment of the quality differentiations in some journal lists developed by evaluation agencies, notably ERIH (European Science Foundation), AERES (France) and the Australian Research Council (Pontille & Torny, 2010a). The British Academy considered the ERIH list in its report on peer review in the SSH fields (British Academy, 2007). In 2008, the editors of History & Philosophy of Science journals launched the "save our journals" campaign and demanded the removal of their journals from the ERIH list. As the ERIH list was largely integrated into the AERES list, a similar petition was promoted in France calling for the journal lists to be abandoned (Pontille & Torny, 2010a). Both lists were ripe targets for the increasing criticism among many SSH fields against quantitative metrics and research management. Several problems were identified with the construction of the lists:

- *Confusing criteria and meaning of categories*. While ERIH maintained that ABC categories relate to differences in journal scope and audience, it seemed that the A ("a very strong reputation") and B ("good reputation") categories for international journals implied quality differences, with C ("local and regional significance") apparently a category of inferior quality. As the A category should include journals "regularly cited all over the world", it was also not clear how ratings related to impact factors. Moreover, the criteria of the AERES list seemed inconsistent (scope, level, high impact factor) across fields.

- *Undervaluing journals publishing in local languages*. Although ERIH sought to increase the visibility of European humanities research published in different languages, the assignment of categories to journals appeared to favour English language journals over local language ones. Marginalisation of local language journals would lead, it was feared, to impoverishment of their publication activities. In the AERES list, however, the French language journals in the field of education, for example, were considered to be more favourably assessed (Rey, 2009).

- *Representativeness and transparency of expert panels*. The members of each expert group were displayed on the ERIH website, however, the panels were

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

deemed not representative enough. They were also not selected upon consultation of representative disciplinary organisations. In the case of AERES list, expert group members were originally not made public and the wider research community was not consulted.

- *Unclear or inappropriate use of the list*. Although ERIH was envisaged as research assessment tool at aggregate level, the journal editors anticipated that the list would "provide funding bodies and other agencies in Europe and elsewhere with an allegedly exact measure of research quality". In Poland, for many years, ERIH was used as an assessment tool and ERIH's categories were used as quality categories (Kulczycki & Rozkosz, 2017). In the case of AERES, the distinction between ABC-categories was actually reduced to a binary distinction, in which only outputs in A and B category journals were taken into account in the assessment of researchers and their units.

Similar issues, related for example to the marginalization of locally relevant journals—including those published in English—or transparency of the expert judgment, were discussed in the case of the Australian journal list. The main official reason for removing the 4-tier level ratings (A*, A, B, C) from the ERA journal list was, however, its alleged inappropriate use at the institutional level: "institutional managers targeting journals only from the top 20% of journals and, in many cases, obstructing their academics from seeking to publish in the other 80%" (Dobson, 2011).

It is an interesting question—albeit difficult to answer—why quality differentiated lists were abandoned by evaluation agencies in some countries but are developed and continue to be used in others. Both AERES and ERA produced the lists for the purpose of allocating government funding based on assessment of university units. In France, this involved identification of "publishing" and "non-publishing" researchers. In the Australian ERA, the results based on journal ratings were supposed to inform expert-evaluation of the units. In both cases, there may thus have been concerns that quality of individual outputs would be assessed—misguidedly—based on the journal instead of their contents.

In Italy and Spain, for example, journal ratings are used to support criteria-based assessment of researchers' productivity for promotion or recruitment, and the criteria for the assignment of journals to different categories are very detailed and formalized. In the Nordic countries, the differentiated publication channel lists are used in a fixed funding-formula to distribute funding between universities. These relatively formal procedures are not intended to produce or replace content-based evaluation of research by experts at institutional or individual level.

3 Current debate on journal evaluation at national level: experiences from the Nordic countries

In three Nordic countries, Denmark, Finland and Norway, bibliometric indicators representing research activities are part of the direct funding formula for the annual allocation of block-grant funding to universities (Sivertsen, 2017). Since 2009, Sweden applies an indicator based on WoS publications and citations for the same purpose (Sīle & Vanderstraeten, 2018). Instead, several Swedish institutions apply the "Norwegian list" for local purposes (Hammarfelt et al., 2016).

The three countries applying the "Norwegian model" at the national level use it for institutional funding allocation by linking comprehensive publication data of the institutions, integrated at the national level, to a list of publication channels (journals and book publishers) with level ratings representing all fields. The rating is performed by experts representing the national research community in the field. The ratings together with a definition of scholarly publications determine what outputs count as peer-reviewed publications and how they are weighted in the funding formula. Accordingly, the list of publication channels serves two main purposes: 1) identify reliably peer-reviewed publication channels; 2) indicate in each field the leading publications channels in terms of quality, impact and prestige (Aagaard, 2018; Pölönen, 2018; Sivertsen, 2018b).

Performance-based research funding systems (PRFS) using undifferentiated counts of peer-reviewed publications risk promoting quantity at the expense of quality (Aagaard & Schneider, 2017; Butler, 2003; 2004; Schneider, Aagaard, & Bloch, 2015; Van den Besselaar, Heyman, & Sandström, 2017). In the Norwegian model, the purpose of the quality index with weighted funding-formula is to make it more rewarding for the universities if publication activity takes place in channels with more stringent requirements related to originality, quality, and impact of submitted manuscripts (Norwegian Association of Higher Education Institutions, 2004). In Norway a funding-model including a publication channel rating has been able to foster publication activity without increasing publishing in the low-impact journals, as happened in Australia where model rewarded publication counts undifferentiated by quality index (Butler 2004; Schneider, Aagaard, & Bloch, 2015; for Denmark, see: Ingwersen & Larsen, 2014).

The possible effects of the national level PRFS indicator on the publishing activities, however, are mainly realized locally (Aagaard, 2018; Aagaard et al., 2014; Hammerfelt et al., 2016). Given that universities use different kinds of journal lists and metrics for internal assessment, funding and promotion purposes (e.g. McKiernan et al., 2019), the governmental incentives cannot alone explain local use of indicators or changes in publication practices. In Sweden, for example, several

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

universities use variants of the Norwegian model including publication channel ratings internally, even if this has no budget funding effects (Hammarfelt et al., 2016). In many countries, publication channel lists have also been produced specifically for assessing career promotion (Gimenez-Toledo & Roman-Roman, 2009; Ferrara & Bonaccorsi 2016). Nevertheless, once the PRFS indicator is established with link to government funding, the publication channel list becomes a relevant metric and tool for research evaluation and management also at the local level, even if individual universities in each Nordic country may differ considerably in how they make use of the national publication channel list. More frequent use of national lists is reported in SSH fields than STEM, probably because other comprehensive metrics have been lacking (Aagaard et al., 2014; Aagaard, 2018; Krog Lind, 2019; Pölönen & Wahlfors, 2016; Sivertsen & Schneider 2012; Wahlfors & Pölönen, 2017). Norway and Finland have published guidelines for the responsible use of the publication channel-based indicators (Pölönen, 2018; Publication Forum, 2020; Sivertsen, 2018).

3.1 Expert-evaluation versus metrics

While the involvement of the research community in the production of the indicator is an important hallmark of the model's legitimacy (Ahlgren, Colliander, & Persson, 2012), in academia the use of expert-based evaluation also raises concerns about personal bias and validity (Bornmann, 2011; Haddawy et al., 2016). Expert-based ratings of publication channels are often compared with JIF rankings or other impact indicators based on average citation counts to articles in journals, which are considered objective measures of quality or impact. Correlation between the subjective and objective methods of journal evaluation is a well-established research track (Serenko & Dohan 2011), to which the national ratings provide a new source of data (Ahlgren & Waltman 2014; Haddawy et al., 2016; Kulczycki & Rozkosz, 2017; Pölönen, Leino, & Auranen, 2011; Saarela & Kärkkäinen, 2016; Saarela et al., 2020; Walters, 2017). Low correlations are sometimes critiqued among the research community. When researchers look at the national ratings, it can be regarded a failure of the expert-based ratings if these do not conform to the impact factor ranking order of journals. These debates take place in Norway and Denmark (Sivertsen & Schneider, 2014), and also in the Finnish context it has been suggested that artificial and/or collective intelligence could improve or even replace the expert-based evaluation in the Norwegian model.

Saarela et al. (2016), Saarela and Kärkkäinen (2020) have used novel data-mining and machine-learning techniques to demonstrate that Scopus-based IPP, SNIP and SJR, in combination with the Danish and Norwegian level ratings, allow for good prediction of the Finnish expert-ratings. They show that higher ratings only rarely

diverged from the classification based on impact factors or the other Nordic ratings. In such cases, however, journals frequently used by Finnish researchers, or even by the panellists, appeared to have been favoured. The authors suggest that automatic rules based on impact factors and other Nordic ratings could replace or assist the expert qualitative judgment to improve the transparency and objectivity and to save man-hours and money for Finish researchers.

Another reasoning holds that evaluation by expert panels could be replaced with methods combining popular vote with mechanical application of JIF. According to Erola (2016), the problem with the current expert-ratings in social sciences is that even "entirely unimpactful" journals have a good chance to be assigned to the highest level. Mechanical rating of journals on basis of JIF is not feasible because the indicator is field dependent, and all Finnish language SSH journals would automatically be left outside the higher quality levels. But if ratings were based only on popular vote among the researchers, journals with most Finnish publications might be favoured over high-impact journals. Therefore, Erola suggests that the vote should be used to identify a pool of important channels, from which Finnish language journals would be placed on the higher levels on basis of a popular vote, and other journals would be rated mechanically on basis of their JIF.

In the debate concerning the involvement of panels in the rating of publication channels the JIF is presented as "a technology of distance" in a "struggle against subjectivity" (Beer, 2016; Porter, 1995). The metric characteristics of the JIF do not mean, however, that it necessarily circumscribes the average quality of journals more reliably or appropriately than expert-based ratings. There are large differences between disciplines in coverage and esteem of JIF (or other journal impact indicators). Because the size of a field, the citation culture and the coverage in WoS influence the JIF values, these are not comparable between or even within disciplines. In Denmark, Finland, and Norway, expert-evaluation of publication channels is informed with a range of impact indicators. A major challenge for the panels, however, is to produce a rating that is more balanced between disciplines and specialties than one only based on impact factors. This involves also taking into account the framework of level quotas that increase equality of ratings across panels in the Norwegian model.

It is a demonstration of trust on the part of the governments in Denmark, Finland, and Norway that the national research communities, represented by the expert panellists, are involved in the construction of the funding-model indicator. In each country, researchers are also actively engaged in this process by suggesting additions and improvements to the ratings, as well as by criticising the ratings. Reliance on journal metrics does not increase the legitimacy of the ratings unless there is a wide agreement among researchers in the field or discipline that these metrics accurately

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

reflect the quality or impact of journals. In many fields, especially SSH, legitimacy of rating based on citation-based journal metrics alone would be low. The rating of publication channels in the Norwegian model is a multidisciplinary exercise that necessarily represents a compromise of disciplinary standards of quality that exist in the research community (Lamont, 2010; Sivertsen, 2016).

When researchers confront ratings that seem incoherent from their perspective, they have had little means to engage with the reasons behind those ratings. Apart from the general level criteria that are published, the evaluation process itself remains relatively opaque. As the recent evaluation of the Norwegian publication indicator suggests, increasing transparency can increase legitimacy of the model (Aagaard et al., 2014). To address this issue, the Norwegian Association of Higher Education Institutions implemented a solution making the procedure and groups for expert-panel decisions more transparent in an internet portal open to all researchers: https://npi.nsd.no/ (Sivertsen, 2018). Similar portal has been developed also in Finland, where all the information supporting the panel evaluation is also available for the researchers: http://jfp.csc.fi:8080/en/ (Pölönen, 2018).

The Nordic countries collaborate in order to increase the uniformity and quality of the publication channel data that support the expert-evaluation process. Nordforsk funded a Nordic collaboration project where the publication channel lists from Denmark, Finland and Norway are integrated and level ratings from different countries are compared (Sivertsen, 2016; 2019). Relatively large discrepancies exist between the Danish, Finnish and Norwegian ratings. In the three countries, a total of around 4,000 journals have been identified across all fields as leading journals included in level 2 or 3. Of these journals, 31% have been rated as leading in all three countries, 27% in at least two countries, and 41% in only one of the countries. The same overall pattern is observed, more or less, in all main fields (Pölönen, 2012; Pölönen & Sivertsen, 2017). The causes of these discrepancies have not been fully investigated, but we speculate that, among other things, national publication profiles, the restrictions imposed on evaluation by the level quota framework (see 3.3. below), and evaluation of journals in different disciplinary contexts may play a role. Increasing the uniformity of national ratings is also on the agenda of this Nordic collaboration.

3.2   Coverage of publications in citation indexing services

The Norwegian model is designed to cover all peer-reviewed output types used across fields: articles in journals, proceedings and books, as well as monographs and edited works regardless of publication country or language. Therefore, the Nordic publication channel ratings need to include not only journals but also other publication series and book publishers. The sources of citation data do not provide

full coverage of all publication channels evaluated by the panels. Reliable international citation databases, WoS, and Scopus, have very limited coverage of books and offer no publisher level impact metrics (Gimenez-Toledo et al., 2016). The coverage of WoS and Scopus is limited mainly to international English language journals. In SSH fields the coverage even of these is partial, and is seriously wanting in case of peer-reviewed journals in other languages.

Google Scholar could be a source for citation data for a wider range of publication channels than WoS or Scopus. However, Google Scholar's sources remain beyond control documented, it is burdensome to use for citation analysis at journal or publisher level, and the quality of data is poor and requires manual cleaning (Bakkalbasi et al., 2006; Neuhaus et al., 2016).

Another issue is that JIF does not cover all journals included in WoS: it has been calculated only for journals in the SCI and the SSCI, but not for those in the Arts & Humanities Citation Index. This means that JIF covers only a small share of humanities journals that happen to be included also SSCI. These few journals are more oriented towards the social sciences (Mañana-Rodríguez & Giménez-Toledo, 2013). Using JIF for the humanities therefore creates biases. Scopus based journal metrics—CiteScore, Scimago Journal Rank (SJR) and Source Normalized Impact per Paper (SNIP), are available in all fields but these metrics also suffer from limited database coverage.

3.3   Correlation of Journal Impact Factors and expert-ratings

There are many reasons why expert-ratings do not follow exactly the JIF ranking order. The most important reason is that JIF varies between disciplines and even between specialties within disciplines (Adler, Ewing, & Taylor, 2008; Amin & Mabe, 2000; Seglen, 1997). JIFs are based on citations from articles in journals indexed in the WoS. The larger the share of publications of a field that is covered by indexed journals, the more fully the JIF captures its citation potential. But if a large share of a field's publications in journals, let alone books, is not covered, citations from publications outside the database are not counted toward the JIF of indexed journals. In this case, it is also likely that a sizeable share of references in articles of indexed journals are to publications in journals and books outside the database and do not count toward the JIFs of indexed journals. Journals that publish all or part of articles in languages other than English also suffer from the predominance of English language journals in the international databases (Lange, 1985; Seglen, 1997).

The publication and citation culture plays a role as well. JIF has a relatively small window for citations, as it is based on citations to journal's articles published in the two preceding years. Such a short time window used in calculation of JIFs is

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

favourable to fields, in which citations accumulate relatively fast (Adler, Ewing, & Taylor, 2008; Amin & Mabe, 2000; Seglen, 1997). Citations received after the time window do not count toward the JIF of journals, and in many fields, this includes clear majority of citations. Fields also differ considerably in average number of references per article (Zitt, Ramanana-Rahary, & Bassecoulard, et al., 2005), in average number of authors per article (Amin & Mabe, 2000) and in total number of researchers and publications in the field (Adler, Ewing, & Taylor, 2008; Seglen, 1997). All these differences contribute to variation in the average number of citations per article, which correlates with the average JIF of journals in different fields.

Impact factors in themselves would not produce balanced ratings across different fields, disciplines and specialties. In the Nordic countries, journals are divided for evaluation between field specific panels. In Norway the number of panels is 89, in Denmark 67 and in Finland 23. It is inevitable that variation in JIF values between WoS and Scopus subject categories result in similar variation between panel fields. JIFs of journals rated in a Physics panel are higher than those rated in a Mathematics panel, so it is inevitable that many level 1 Physics journals have higher JIFs than level 2 Mathematics journals. Similar discrepancies are produced across the panel framework. But even within each panel, journals in different subfields may have widely different JIFs.

It also contributes to the difficulty of comparing journals within subfields that journals associated with other fields with relatively high impact factors (typically bio, medical and health sciences) rank higher than the core journals of the subject category. JIFs are also influenced by the research orientation of journals within a field, such as basic-clinical (Seglen, 1997; van Eck et al., 2013), theoretical-empirical, or qualitative-quantitative research. In addition to this, journals publishing review articles gain on average more citations than journals publishing original research papers (Adler, Ewing, & Taylor, 2008; Amin & Mabe, 2000; Seglen, 1997). There can, in short, be multiple reasons why a JIFs ranking order cannot be maintained between or even within panels.

Access to higher level publication channels ought to be equal across fields if lists are used for evaluation or funding among universities with different disciplinary profiles. In the Nordic publication channel lists (Sivertsen, 2018) this balance is achieved by limiting level 2 nominations in such a way that in each panel the level 2 journals publish about the same share of the total world output (Ahlgren, Colliander, & Persson, 2012; Ahlgren & Waltman, 2014). In Norway, panel quotas are based on national output of articles, of which the level 2 journals in all fields may not exceed publishing 20 percent of the articles. In Denmark, panels were at first allowed to rate to level 2 at the most 20 percent of the journals (Sivertsen, 2010). Soon, new quotas were introduced based on the total output of articles, of which

the level 2 journals might not exceed 20 percent (Jensen, 2011). The first rating in Finland was based on percentage of channels but the updated rating published in 2015 was based on total output of articles, of which the level 2 and 3 journals may not exceed 20 percent and the level 3 journals 5 percent (Pölönen & Ruth, 2015).

The rationale behind the article-based quotas is to take into account the size of journals. In some natural and medical science disciplines publication activity concentrates heavily in large leading international journals. Therefore, panel quotas based on the percentage of publication channels result in unbalanced representations of different field's output on level 2 (Ahlgren, Colliander, & Persson, 2012; Ahlgren & Waltman, 2014; Pölönen & Ruth, 2015). For example, 20 percent of the top journals in Physics publish more than half of the world total as well as national journal article output, whereas the same share of journals in SSH fields publish only 30 percent of the output. Article-based level quotas are needed in the Norwegian Model whether or not journal metrics are involved in the rating of publication channels. It follows, however, that in some instances the journal size can become a decisive factor in level 2 nomination if a panel is running out of quota. It is important to notice that the publication counting techniques, including fractionalization, may have to be adjusted to achieve a good balance between all fields (Sivertsen, Rousseau, & Zhang, 2019).

3.4   Information to support expert assessment

Expert ratings and JIFs tend to correlate broadly. In most fields, the average JIFs of higher rated journals are higher than that of lower rated journals (Ahlgren, Colliander, & Persson, 2012; Ahlgren & Waltman, 2014; Pölönen, Leino, & Auranen, 2011), even if the ratings do not follow exactly the JIF ranking order of journals. The reason for this is twofold. In some fields, for instance medicine, experts know JIFs and rely on them also in rating journals. This would probably happen whether or not JIFs were provided for the panels.

In Denmark, Finland, and Norway, JIFs are indeed supplied to all panels (Ahlgren & Waltman, 2014; Saarela et al., 2016; Sivertsen, 2010; 2016). In Norway, originally JIFs were supplied but this has been replaced with Scopus based SNIP, CiteScore, and SJR indicators. In Denmark, panels have been supplied with field-normalized JIFs. In Finland, panels were at first provided JIF, JIF5, SNIP, and SJR. Currently the set of journal indicators provided to panel includes CiteScore, SNIP, and SJR. In Finland, panels were from the start also provided expert ratings of publication channels in Norway and Denmark, as well as Australian and ERIH ratings. The current set of indicators in Finland includes Danish and Norwegian ratings. Now also Denmark and Norway inform panels about the ratings of the same journals in the other Nordic countries. Especially in the SSH fields, other expert ratings are an

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

important addition wherever there is a lack of journal indicators derived from WoS or Scopus.

In all fields, but especially in the SSH, the national publication channel lists and ratings cover the peer-reviewed literature more extensively than the international citation databases and impact factors (Dassa et al., 2011; Hicks & Wang, 2011; Pölönen, Leino, & Auranen, 2011; Sivertsen & Larsen, 2012). It is an important task of the publication channel ratings in the Norwegian model also to distinguish between peer-reviewed and not-peer-reviewed outlets (Pölönen, Engels, & Guns, 2020). This distinction is mainly based on formal criteria that are fairly easy to check, such as use of ISSN/ISBN identifier, and existence of a regular peer-review procedure as well as an expert editorial board. There is also an increasing discussion in the Nordic countries if and how should open access (OA) and open science be integrated into the evaluation criteria. The identification of scholarly journals also involves screening of the national authority lists for so-called predatory journals (Eykens, Guns, & Engels, 2018). The distinction between level 1 and level 2 is more complicated, and involves broad consideration of relative international importance, quality, impact and prestige of journals withing different fields and specialties. The information on ratings from other Nordic countries is helpful in identifying both top- and bottom-tier peer-reviewed journals and book publishers.

Journal metrics and level ratings are supposed to support expert-evaluation, which the expert panelists principally base on their own experience of different publication channels. They may have gained personal knowledge of the editorial and peer-review procedures as editors, editorial board members, reviewers and authors. As active researchers they also read and use large number of articles and books published in different channels. As members of international and national research communities they also learn about reputation of different channels in disciplinary and interdisciplinary contexts.

One major challenge of the Nordic expert panels is to cover a wide range of outlets in their field, not all of which individual panelists have personal experience or knowledge of. Not each and every discipline or specialty has an expert in the panel. Panels need to have input also from the national research communities, of which they are representatives. For example, in Finland, panelists are encouraged to consult other specialists in the field. Some panels and panelist engage local communities more than others, so there is a lot of variation in practice. All Nordic countries producing authority lists also offer individual researchers the option to suggest new additions to the ratings, as well as to suggest upgrades to level ratings.

Expert panels may face pressure from the research community to upgrade channels that are frequently used by their colleagues, to show institutional or disciplinary solidarity. The purpose of JIFs and ratings from other Nordic countries is not to

decide the ratings on behalf of the national experts, but to help them estimate and discuss the relative impact and esteem of journals in the international context. It is the task of the expert-panels in the Norwegian model to know how JIFs work in context of disciplines and specialties under their responsibility. If used with due caution, citation-based metrics can provide valuable information to assist expert evaluation (Hicks et al., 2015). This holds true for the evaluation of journals and book publishers too.

Expert-based ratings and citation-based journal metrics represent in different ways the same dimensions of research quality: solidity, originality, scholarly relevance or practical utility (Gulbrandsen, 2000; Auranen et al., 2013). It has been argued that citations may reflect, with some limitations, scientific impact and relevance but scarcely solidity, originality, and societal value of research (Aksnes, Langfeldt, & Wouters, 2019). While JIF also gives a very narrow representation of the journal quality, it is possible that expert-assessment of publications channels is able to provide a more well-rounded representation of the different dimensions of research quality—it requires further research, however, how the expert-ratings represent research quality.

At macro level, results based on citations and publication channel ratings tend to concur (Ahlgren, Colliander, & Persson, 2012; Auranen & Pölönen, 2012; Auranen et al., 2013; Sandström & Sandström, 2009), even if—of course—the expert-based ratings do not predict the citation counts of individual papers any better than JIF. An analysis of 15,265 Finnish WoS publications from 2011–2013 shows considerably stronger citation impact for articles in higher rated journals compared to lower rated journals (Pölönen & Sivertsen, 2017; for a more complete report of an earlier analysis, see Auranen & Pölönen, 2012; and for similar analysis for Norway, see Aksnes, 2017). This suggests that publication channel ratings can indicate differences in citation impact of publication activity also in natural and medical sciences, where citation-based measurement would usually be preferred to national ratings as quality measures for evaluating or funding research. Also, even if the expert-ratings are often suspected of personal bias in case of specific journals, overall, the expert-evaluation can produce robust macro-level results also from the perspective of the citation analysis.

4   Recommendations

We conclude by presenting a list of recommendations for national publication channel lists based on our experience with scholarly publication channel lists in different countries as well as extensive discussion in the context of the COST-action ENRESSH (European Network of Research Evaluation in the Social Sciences and

National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

Humanities). We only provide general recommendations for the construction and maintenance of publication channel lists that are applicable in variety of geographical contexts. More specific measures will depend on the contexts and purposes of the use of lists. The recommendations pertain to organisation, evaluation, quality control and usage. These recommendations are intended to be useful to all who are engaged in the creation and maintenance of lists of scholarly publication channels.

4.1 Organization

4.1.1 Define and make explicit the purpose of the publication channel list.

A publication channel list is typically constructed to support an evaluative context or funding procedure. Hence, define and clearly state the main purpose at the outset, even when several uses of the publication channel list are envisioned. This should be the purpose guiding the construction and development of the publication channel list, even if there may be other—even unpredicted or unsuitable—uses. Explain how the intended use is responsible in the perspective of recommendations like DORA, the Leiden manifesto, or the Metric Tide report. If certain uses are considered unsuitable, such as the use at individual level, this should be stated explicitly and publicly.

4.1.2 Determine bodies responsible for governance and evaluation

Construction and maintenance of lists requires steering to establish and develop general classification criteria for publication channels, as well as an organisation of field-specific expert-group(s) that are responsible for the evaluation of publications channels. Whether there are pre-existing bodies that can take up new functions or new bodies need to be established for the purpose, state clearly which body is responsible for the steering, and which body for implementing the publication channel list. The steering body requires a broad representation to supervise the disciplinary panels. Also define procedures and criteria for selecting the members for the steering and evaluation groups. Employ secretarial staff to assist the steering body and/or the evaluation process, and clearly define also their role.

4.1.3 Make sure that the publication channel list represents research adequately

The main advantage of a national publication channel list compared to WoS or Scopus is its wider coverage of research outputs and outlets. Make sure that the national channel list includes all serials and book publishers that the researchers affiliated with institutions use for publishing peer-reviewed articles in journals, conferences and books, as well as monographs and edited volumes. In order to

ensure that publication channels from different fields are adequately covered, use both international and national lists to construct the list of journals and book publishers. Use a well-established field-classification system (e.g. OECD FOS) to assign journals/series to different fields, and to specific expert-groups for evaluation. To identify journal field, make use, when possible, of established journal field classifications (e.g. from the ISSN Centre, WoS, Scopus, or ScienceMetrix).

4.1.4 Define principles for cataloguing publication channels

ISSN and ISBN are the standard international persistent identifiers used in publication metadata to connect outputs to publication channels. To ensure the interoperability with publication databases, use ISSN and ISBN to identify serials and book publishers also in the publication channel list. However, take into account their ambiguities. A single journal often has multiple ISSNs (e.g. for print and online versions). As for ISBNs, the ISBN-root is not an unequivocal identifier of a publisher, as books with the same ISBN-root can appear under different publisher and imprint names. Clearly define if the channel list is organised by unique ISSNs and ISBNs, or by unique channels. Also make explicit if the existence of registered ISSN and/or ISBN is a technical defining criterion for a channel, and if there are exceptions (e.g. conferences that use no ISSN or ISBN). Establish regular procedures for keeping the publication channel data up to date and valid. Internal persistent identifiers can be useful.

4.2 Evaluation

4.2.1 Define and clearly state inclusion criteria for publication channels

PRFSs typically use national publication channel lists to identify peer-reviewed articles and books, so the main aim of the national list is to indicate peer-reviewed serials and book publishers. Peer-review practices differ between fields and publication types, so provide a clear definition of peer-review and other possible inclusion or exclusion criteria (such as expert editorial board, local, national or international authorship, "predatory" behaviour, relevance, etc). Also explain clearly how peer-reviewed and not-peer-reviewed channels are indicated in the list (e.g. levels distinction, or complete exclusion of not-approved channels).

4.2.2 Clearly state if the publication channel list indicates/implicates quality differences

Peer-reviewed journals and book publishers differ in terms of quality, impact and prestige as perceived by the research communities. If such logic is relevant for the purpose(s) of the list, clearly define how many quality categories, if any, are used,

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

what are the criteria for differentiating between channels, by what means the differentiated classification is balanced between disciplines (e.g. world production), and how the differences are indicated in the list (e.g. levels distinction). Also explicate how open access and national language channels are treated.

4.2.3   Make explicit the role of expert-judgment and metrics

National lists may contain tens of thousands of publication channels. Therefore, support the expert evaluation by dividing the list in relevant disciplinary categories and with metrics and other relevant information. Provide experts with information on inclusion of journals and book publishers as peer-reviewed channels in international and national lists, as well as bibliometric journal indicators and level ratings from other national lists to support classification of channels into different quality levels. Explain clearly the usefulness and limitations of all information supporting evaluation, and if possible, make the data openly available. The perceived validity of, for example, Journal Impact Factors differs between fields and individuals, so state clearly if some information is used as evaluation criterion or if their role is only to inform expert judgment.

4.3   Quality control

4.3.1   Establish procedures for feedback, updates, and corrections

The landscape of publication channels changes constantly, as journals and book publishers start publishing, end operations, split and merge. Also, peer-review status and perceived quality and prestige of channels may change over time. Establish procedures for regularly adding new channels to the national list, as well as for reviewing and updating the quality levels and inclusion. It is especially important that researchers, whose work constitutes the research output subject to national evaluation or funding procedures, are able to provide feedback on the list. Make sure that feedback from the research community is communicated to the experts responsible for the evaluation of channels.

4.3.2   Make efforts to identify and exclude questionable journals

The publishing model based on author fees (APC, article processing charges) has increased the number of questionable (predatory, grey-zone) journals and book publishers that claim but fail, among other issues, to provide reliable peer review. Characteristic features of such channels include fast processing time of manuscripts, a vague topic, aggressive email marketing, lack of contact information, fake information about the editorial board, database indexing and impact factors. Although questionable channels are often difficult to identify, make effort to keep them away from the category of peer-reviewed channels, e.g. through screening

against both blacklists (e.g. Cabell's Predatory reports) and white lists (e.g. DOAJ; see also Eykens et al., 2019). Support the expert evaluation with information from such sources.

4.3.3   Assess the list and its criteria regularly for possible improvements

A national publication channel list is expected to increase the reliability of identification of peer-reviewed outputs, and possibly also a meaningful and balanced differentiation of peer-reviewed output according to channel quality, impact and prestige across fields. Compare the peer-review status and quality levels in the national list systematically with those in other national lists, as well as with international lists and impact factors. Use national publication data to assess the balance of classification between fields, and to monitor developments in scholarly publishing. Use this information to help experts and steering-bodies to improve the list and its criteria.

4.4   Usage

4.4.1   Make the publication channel list and its basis openly available

Transparency is the key to generating trust and feedback from the research community, as well as to any informed and responsible use of the publication channel list. Establish a website where the information about the organisation, steering and expert groups is available, and the evaluation procedures and criteria are explained. Make also the list of publication channels available on the website as documents (e.g. as an Excel list) or via a searchable interface (e.g. a portal).

4.4.2   Explain the use of the publication channel list in national evaluation or funding procedure

State clearly in what way and why the national publication channel list is used in the evaluation or funding procedure, what is the publication data used, which institutions does it concern, and what is its financial importance. Also make explicit how updated versions of the list apply to outputs from different publication years. Explain to both institutions and researchers how the publication channel list is applied to individual outputs, and how channels are matched with articles and books. If one output can be matched with several channels (e.g. book series, imprint, and publisher), explain how channels are prioritized.

4.4.3   Provide guidelines for the responsible use of the list at institutional and individual level

According to the recommendations of DORA, the Leiden Manifesto for research metrics (Hicks et al., 2015) and the Metric Tide report (Wilsdon et al., 2015), the

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

evaluation of the quality of research at universities or other research organisation units or of individual researchers must primarily be based on expert evaluation, but research metrics can be used to support the evaluation. Explain clearly the limitations of the national publication channel list at different levels and the conditions for its responsible use.

Acknowledgements

Author contributions

References

Aagaard, K. (2018). Performance-based research funding in denmark: The adoption and translation of the norwegian model. Journal of Data and Information Science, 3(4), 20–30.

Aagaard, K, Bloch, C., Schneider J.W., Henriksen, D., Ryan, T.K., & Lauridsen, P.S. (2014). Evaluering af den norske publiceringsindikator. Dansk Center for Forskningsanalyse, Aarhus Universitet.

Aagaard, K., & Schneider, J. (2017). Some considerations about causes and effects in studies of performance-based research funding systems. Journal of Informetrics, 11(3), 923–926. doi:10.1016/j.joi.2017.05.018

Adler, R., Ewing, J., & Taylor, P. (2008). Citation statistics. A report from the International Mathematical Union. www.mathunion.org/publications/report/citationstatistics0

Ahlgren, P., Colliander, C., & Persson, O. (2012). Field normalized rates, field normalized journal impact and Norwegian weights for allocation of university research funds. Scientometrics, 92(3), 767–780. doi:10.1007/s11192-012-0632-x.

Ahlgren, P., & Waltman, L. (2014). The correlation between citation-based and expert-based assessments of publication channels: SNIP and SJR vs. Norwegian quality assessments. Journal of Informetrics, 8(4), 985–996.

Aksnes, D. (2017). Artikler i nivå 2-tidsskrifter blir mest sitter. Forskerforum. https://www. forskerforum.no/artikler-i-niva-2-tidsskrifter-blir-mest-sitert/

Aksnes, D., Langfeldt, L., & Wouters, P. (2019). Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. Sage Open. https://doi.org/10.1177/2158244019829575

Aksnes, D.W., & Sivertsen, G. (2019). A criteria-based assessment of the coverage of Scopus and Web of Science. Journal of Data and Information Science, 4(1), 1–21.

Amin, M., & Mabe, M. (2000). Impact factor: use and abuse. Perspectives in Publishing, 1, 1–6. http://www.elsevier.com/framework_editors/pdfs/Perspectives1.pdf

Research Paper

ANVUR. (2019). Regolamento per la classificazione delle riviste nelle aree non bibliometriche. https://www.anvur.it/wp-content/uploads/2019/02/REGOLAMENTO-PER-LA-CLASSIFI CAZIONE-DELLE-RIVISTE_20022019.pdf

Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing data-bases. Scientometrics, 68(3), 329–342.

Auranen, O., Leino, Y., Poropudas, O., & Pölönen, J. (2013). Julkaisufoorumi-luokitus ja viit-tausindeksittieteellisten julkaisujen laadun mittareina: Web of Science -aineistoon perustuva vertailu. TaSTI Työraportteja 8/2013.

Auranen, O., & Pölönen, J. (2012). Classification of scientific publication channels: Final report of the Publication Forum project (2010–2012). Helsinki: Federation of Finnish Learned Societies. URL: http://www.julkaisufoorumi.fi/sites/julkaisufoorumi.fi/files/publication_ forum_project_final_report_0.pdf

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomedical Digital Libraries, 3, 7. doi:10.1186/1742-5581-3-7

Beer, D. (2016). Metric Power. Palgrave Macmillan UK.

Bornmann, L. (2011). Scientific peer review. Annual Review of Information Science and Technology, 45(1), 197–245. doi:10.1002/aris.2011.1440450112

British Academy. (2007). Peer Review: The challenges for the humanities and social sciences. London: The British Academy. https://www.thebritishacademy.ac.uk/sites/default/files/Peer-review-challenges-for-humanities-social-sciences.pdf

Butler, L. (2003). Explaining Australia's increased share of ISI publications—The effects of a funding formula based on publication counts. Research Policy 32(1), 143–155.

Butler, L. (2004). What Happens When Funding is Linked to Publication Counts? In H.F. Moed, W. Glänzel & U. Schmoch (eds.) Handbook of Quantitative Science and Technology, Dordrecht, The Netherlands: Kluwer Academic Publishers, 340–389.

Cleere, L., & Ma, L. (2018). A Local Adaptation in an Output-Based Research Support Scheme (OBRSS) at University College Dublin. Journal of Data and Information Science, 3(4), 74–84. doi:10.2478/jdis-2018-0022

Csiszar, A. (2017). How lives became lists and scientific papers became data: Cataloguing author-ship during the nineteenth century. British Journal of History of Science, 50(1), 23–60.

de Boer, H.F., Jongbloed, B., Benneworth, P. et al. (2015). Performance-based funding and perfor-mance agreements in fourteen higher education systems—Report for the Ministry of Educa-tion, Culture and Science, CHEPS, University of Twente: Enschede. https://ris.utwente.nl/ws/portalfiles/portal/5139542/jongbloed+ea+performance-based-funding-and-performance-agreements-in-fourteen-higher-education-systems.pdf

de Filippo, D., Aleixandre-Benavent, R., & Sanz-Casado, E. (2019). Categorization model of Span-ish scientific journals in social sciences and humanities. In G. Catalano et al. (eds), Proceedings of the 17th International Conference of the International Society for Scientomet-rics and Informetrics, Vol II. Rome: International Society for Scientometrics and Informetrics, 726–737.

de Solla Price, D. (1963). Little science, big science- and beyond. New York: Columbia University Press.

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

Dobson, I. (2011). Australia: troubled history of an ERA. University World News, 05 June 2011. https://www.universityworldnews.com/post.php?story=201106031829474

Engels, T.C.E., & Guns, R. (2018). The Flemish performance-based research funding system: A unique variant of the Norwegian model. Journal of Data and Information Science, 3(4), 45–60. https://doi.org/10.2478/jdis-2018-0020

Engels, T.C.E., Starčič, A., Kulczycki, E., Pölönen, J., & Sivertsen, G. (2018). Are book publications disappearing from scholarly communication in the social sciences and humanities? Aslib Journal of Information Management, 70(6), 592–607. https://doi.org/10.1108/AJIM-05-2018-0127

Erola, J. (2016). Valitaan lehdille JUFO-taso äänestämällä! https://janierola.net/2016/05/12/valitaan-lehdet-jufo-tasoille-aanestamalla/

Eykens, J., Guns, R., & Engels, T.C.E. (2018). Comparing VABB-SHW (version VIII) with Cabells Journal Blacklist and Directory of Open Access Journals: Report to the Authoritative Panel. ECOOM: Antwerp.

Eykens, J., Guns, R., Rahman, J., & Engels, T.C.E. (2019). Identifying publications in questionable journals in the context of performance-based research funding. PLoSONE. https://doi.org/10.1371/JOURNAL.PONE.0224541

Dassa, M., Kosmopoulos, C., & Pumain, D. (2010) JournalBase—A Comparative International Study of Scientific Journal Databases in the Social Sciences and the Humanities (SSH). Cybergeo: European Journal of Geography, document 484. http://cybergeo.revues.org/22862

Ferrara, A., & Bonaccorsi, A. (2016). How Robust is Journal Ratings in Humanities and Social Sciences? Evidence from a Large-scale, Multi-method Exercise. Research Evaluation, 25(3), 279–291. doi.org/10.1093/reseval/rvv048

Garfield, E. (1979). Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. ISI Press.

Genoni, P., & Haddow, G. (2009). ERA and the ranking of Australian humanities journals. Australian Humanities Review, 46, 7–26.

Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T.C.E. et al. (2019). Taking scholarly books into account, part II: a comparison of 19 European countries in evaluation and funding. Scientometrics, 118(1), 233–251.

Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T.C.E. et al. (2016). Taking scholarly books into account: Current developments in five European countries. Scientometrics, 107(2), 685–699.

Giménez-Toledo, E., Kulczycki, E., Pölönen, J., & Sivertsen, G. (2019). Bibliodiversity—What it is and why it is essential to creating situated knowledge. LSE Impact Blog 5.12.2019. https://blogs.lse.ac.uk/impactofsocialsciences/2019/12/05/bibliodiversity-what-it-is-and-why-it-is-essential-to-creating-situated-knowledge/

Giménez-Toledo, E., Mañana-Rodríguez, J., & Sivertsen, G. (2017). Scholarly book publishing: Its information sources for evaluation in the social sciences and humanities. Research Evaluation, 26(2), 91–101.

Gimenez-Toledo, E., & Roman-Roman, A. (2009). Assessment of humanities and social sciences monographs through their publishers: A review and a study towards a model of evaluation. Research Evaluation, 18, 201–213.

Glänzel, W., & Wouters, P. (2013). The do's and don'ts in individual level bibliometrics. Paper presented at the 14th International Society of Scientometrics and Informetrics Conference, Vienna, Austria. Retrieved from http://www.slideshare.net/paulwouters1/issi2013-wg-pw

Good, B., Vermeulen, N., Tiefenthaler, B., & Arnold, E. (2015). Counting quality? The Czech performance-based research funding system. Research Evaluation, 24(2), 91–105. doi: 10.1093/reseval/rvu035

Gregorio-Chaviano, O. (2018). Evaluación y clasificación de revistas científicas: reflexiones en torno a retos y perspectivas para Latinoamérica. Revista Lasallista de Investigación, 15(1), 166–179. https://dx.doi.org/10.22507/rli.v15n1a12

Gross, P., & Gross, E. (1926). College libraries and chemical education. Science, 66(1713), 385–389. doi:10.1126/science.66.1713.385

Gulbrandsen, M. (2000). Between Scylla and Charybdis—and Enjoying It? Organisational Tensions and Research Work. Science Studies, 13(2), 52–76.

Haddawy, P., Hassan, S.-U., Asghar, A., & Amin, S. (2016). A comprehensive examination of the relation of three citation-based journal metrics to expert judgment of journal quality. Journal of Informetrics, 10(1), 162–173.

Haddow, G., & Hammarfelt, B. (2018). Quality, Impact, and Quantification: Indicators and Metrics Use by Social Scientists. Journal of the Association for Information Science and Technology, 70(1), 16–26. https://doi.org/10.1002/asi.24097

Hammarfelt, B., Nelhans, G., Eklund, P., & Åström, F. (2016). The heterogeneous landscape of bibliometric indicators. Evaluating models for allocating resources at Swedish universities. Research Evaluation, 25, 292–305.

Haustein, S. (2012). Multidimensional Journal Evaluation: Analysing scientific periodicals beyond the impact factor. Berlin/Boston: De Guyter Saur.

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. Scientometrics, 44(2), 193–215.

Hicks, D. (2012). Performance-based university research funding systems. Research Policy, 41(2), 251–261.

Hicks, D., & Wang, J. (2011). Coverage and overlap of the new social science and humanities journal lists. Journal of the American Society for Information Science and Technology, 62(2), 284–294.

Hicks, D., Wouters, P.F., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics: Use these 10 principles to guide research evaluation. Nature, 520(7548), 429–431.

Houghton, B. (1975). Scientific Periodicals: Their Historical Development, Characteristics and Control. Hamden, Conn.: Linnet Books.

Huang, Y., Li, R., Zhang, L., & Sivertsen, G. (2020). A comprehensive analysis of the journal evaluation system in China. Quantitative Science Studies (forthcoming).

Ingwersen, P., & Larsen, B. (2014). Influence of a performance indicator on Danish research production and citation impact 2000–12. Scientometrics, 101, 1325–1344.

Jensen, C.B. (2011). Making Lists, Enlisting Scientists: The Bibliometric Indicator, Uncertainty and Emergent Agency. Science Studies, 24(2), 64–84.

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

Johnson, R., Watkinson, A., & Mabe, M. (2018). The STM Report: an overview of scientific and scholarly publishing, 5th edition. The Hague: International Association of Scientific, Technical and Medical Publishers.

Jonkers, K., & Zacharewitcz, T. (2016). Research Performance Based Funding Systems: A Comparative Assessment. Publications Office of the European Union. doi:10.2760/70120

Krog Lind, J. (2019). The missing link: How university managers mediate the impact of a performance-based research funding system. Research Evaluation, 28(1), 84–93.

Kulczycki, E. (2017). Assessing publications through a bibliometric indicator: The case of comprehensive evaluation of scientific units in Poland. Research Evaluation, 45, 41–52. https://doi.org/10.1093/reseval/rvw023.

Kulczycki, E. (2018). The diversity of monographs: Changing landscape of book evaluation in Poland. Aslib Journal of Information Management, 70(6), 608–622. http://doi.org/10.1108/AJIM-03-2018-0062.

Kulczycki, E., Engels, T.C.E., Pölönen, J. et al. (2018). Publication patterns in the social sciences and humanities: The evidence from eight European countries. Scientometrics, 116(1), 463–486.

Kulczycki, E., Guns, R., Pölönen, J. et al. (2020). Multilingual Publishing in the Social Sciences and Humanities: A Seven-Country European Study. Journal of the Association for Information Science and Technology. https://doi.org/10.1002/asi.24336

Kulczycki, E., & Korytkowski, P. (2018). Redesigning the Model of Book Evaluation in the Polish Performance-based Research Funding System. Journal of Data and Information Science, 3(4), 61–73. https://doi.org/10.2478/jdis-2018-0021

Kulczycki, E., & Rozkosz, E.A. (2017). Does an expert-based evaluation allow us to go beyond the Impact Factor? Experiences from building a ranking of national journals in Poland. Scientometrics, 111(1), 417–442. https://doi.org/10.1007/s11192-017-2261-x

Lamont, M. (2010). How Professors Think: Inside Curious World of Academic Judgment. Harvard University Press: Cambridge Massachusetts.

Lange L. (1985). Effects of disciplines and countries on citation habits. An analysis of empirical papers in behavioural sciences. Scientometrics, 8(3), 205–215.

Larivière, V., & Macaluso, B. (2011). Improving the coverage of social science and humanities researchers' output: The case of the érudit journal platform. Journal of the American Society for Information Science & Technology, 62(12), 2437–2442.

Lavik, G.A., & Sivertsen, G. (2017). ERIH PLUS—Making the SSH Visible, Searchable and Available. Procedia Computer Science, 106, 61–65. https://doi.org/10.1016/j.procs.2017.03.035

Mañana-Rodríguez, J., & Giménez-Toledo, E. (2013). Scholarly publishing in social sciences and humanities, associated probabilities of belonging and its spectrum: A quantitative approach for the Spanish case. Scientometrics, 94(3), 893–910.

Marchitelli, A., Galimberti, P., Bollini, A., & Mitchell, D. (2017). Improvement of editorial quality of journals indexed in DOAJ: A data analysis. JLIS.it, 8(1), 1–21. https://doi.org/10.4403/jlis.it-12052

McKiernan, E.C., Schimanski, L.A., Muñoz Nieves, C., Matthias, L., Niles, M.T., & Alperin, J.P. (2019). Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. eLife, 8, e47338. https://doi.org/10.7554/eLife.47338

Molas-Gallart, J. (2012). Research Governance and the Role of Evaluation: A Comparative Study. American Journal of Evaluation, 33(4), 583–598. https://doi.org/10.1177/1098214012450938

Nederhof, A.J. (1989). Books and chapters are not to be neglected in measuring research productivity. American Psychologist, 44(4), 734–735.

Nederhof, A.J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. Scientometrics, 66(1), 81–100.

Neuhaus, C., Neuhaus, E., Asher, A., & Wrede, C. (2006). The depth and breadth of Google Scholar: an empirical study. Portal: Libraries and the Academy, 6(2), 127–141.

Nisonger, T. (1998). Management of Serials in Libraries. Englewood: Libraries Unlimited.

Norwegian Association of Higher Education Institutions. (2004). A Bibliometric Model for Performance-based Budgeting of Research Institutions. https://npi.nsd.no/dok/Vekt_pa_forskning_2004_in_english.pdf

Olijhoek, T., Mitchell, D., & Bjornshauge, L. (2016). Criteria for Open Access and Publishing. ScienceOpen Research, January. doi:10.14293/s2199-1006.1.sor-edu.amhuhv.v1.

Ossenblok, T.L.B., Engels, T.C.E., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science—a comparison of publication patterns and incentive structures in Flanders and Norway (2005–9). Research Evaluation, 21(4), 280–290.

Pontille, D., & Torny, D. (2010a). The controversial policies of journal ratings: Evaluating social sciences and humanities. Research Evaluation, 19(5), 347–360. https://doi.org/10.3152/095820210X12809191250889

Pontille, D., & Torny, D. (2010b). Revues qui comptent, revues qu'on compte: produire des classements en économie et gestion. Revue de la regulation: Capitalisme, institutions, pouvoirs, 8. doi:10.4000/regulation.8881

Pontille, D., & Torny, D. (2012). Rendre publique l'évaluation des SHS: les controverses sur les listes de revues de l'AERES. Quaderni, 77, 11–24. doi:10.4000/quaderni.542

Porter, T.M. (1995). Trust in Numbers: The Pursuit of Objectivity in Science and Public Life. Princeton: Princeton University Press.

Pölönen, J. (2012). Comparison of Nordic Publication Channel Ratings with special regard to SSH, Nordforsk Workshop on Bibliometrics for the Social Sciences and Humanities, Helsinki, 10.10.2012. https://www.academia.edu/34516798/Comparison_of_Nordic_publication_channel_ratings_with_special_regard_to_SSH

Pölönen, J. (2018). Applications of, and Experiences with, the Norwegian Model in Finland. Journal of Data and Information Science, 3(4), 31–44.

Pölönen, J., Engels, T., & Guns, R. (2020). Ambiguity in identification of peer-reviewed publications in the Finnish and Flemish performance-based research funding systems. Science and Public Policy, scz041, https://doi.org/10.1093/scipol/scz041.

Pölönen, J., Leino, O., & Auranen, O. (2011). Coverage and Ranking of Journals: Comparison of six data sources. European Network of Indicator Designers (ENID) Conference in Rome, 7th–9th September 2011.

Pölönen, J., & Ruth, A.-S. (2015). Final report on 2014 review of ratings in Publication Forum, Federation of Finnish Learned Societies 2015. http://www.julkaisufoorumi.fi/sites/julkaisufoorumi.fi/files/publication_forum_final_report_on_2014_review_of_ratings.pdf.

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

Pölönen, J., & Sivertsen, G. (2017). Experiences with the rating of publication channels for the Norwegian Model: With a response to a proposal for automated ratings from Saarela et al. (2016). 22nd Nordic Workshop on Bibliometrics And Research Policy, 9.-10.11.2017 Helsinki. https://figshare.com/articles/Experiences_with_the_rating_of_publication_channels_for_the_Nordic_Model_With_a_response_to_a_proposal_for_automated_ratings_from_Saarela_et_al_2016_/5624731

Pölönen, J., & Wahlfors, L. (2016). Local use of a national rating of publication channels in Finnish universities (poster presentation). 21st Nordic Workshop on Bibliometrics and Research Policy, Copenhagen, 3.-4.11.2016. https://figshare.com/articles/Local_Use_of_a_National_Rating_of_Publication_Channels_in_Finnish_Universities_NWB_2016_poster_/4246541

Publication Forum. (2020). User guide for the Publication Forum classification 2019. The Committee for Public Information (TJNK) and Federation of Finnish Learned Societies (TSV): Helsinki. doi: https://doi.org/10.23847/isbn.9789525995312

Rey, O. (2009). Productivité et qualité scientifique: avec quelles publications compter? Dossier d'actualité de la VST, 46. http://www.inrp.fr/vst/LettreVST/46-juin-2009.php.

Román Román, A. (2010). Cómo valorar la internacionalidad de las revistas de Ciencias Humanas y su categorización en ERIH. Revista española de Documentación Científica, 33(3), 341–377. http://dx.doi.org/10.3989/redc.2010.3.735.

Saarela, M., Kärkkäinen, T., Lahtonen, T., & Rossi, T. (2016). Expert-based versus citation-based ranking of scholarly and scientific publication channels. Journal of Informetrics, 10(3), 693–718.

Saarela, M., & Kärkkäinen, T. (2020). Can we automate expert-based journal rankings? Analysis of the Finnish publication indicator. Journal of Informetrics, 14(2). doi: https://doi.org/10.1016/j.joi.2020.101008

Saenen, B., Morais, R., Gaillard, V., & Borrell-Damián, L. (2019). Research Assessment in the Transition to Open Science: 2019 EUA Open Science and Access Survey Results. https://eua.eu/downloads/publications/research%20assessment%20in%20the%20transition%20to%20open%20science.pdf

Sandström, U., & Sandström, E. (2009). The field factor: towards a metric for academic institutions. Research Evaluation, 18(3), 243–250. https://doi.org/10.3152/095820209X466892.

Schneider, J.W. (2009). An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway. European Political Science, 8(3), 364–378. doi:10.1057/eps.2009.19.

Schneider, J.W., Aagaard, K., & Bloch, C.W. (2015). What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models. Research Evaluation, 25(2), 1–13.

Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. BMJ, 314(7079), 498–502.

Serenko, A., & Dohan, M. (2011). Comparing the expert survey and citation impact journal ranking methods: Example from the field of artificial intelligence. Journal of Informetrics, 5(4), 629–648.

Sīle, L., & Vanderstraeten, R. (2018). Measuring changes in publication patterns in a context of performance-based research funding systems: the case of educational research in the University of Gothenburg (2005–2014). Scientometrics, 118, 71–91.

Research Paper

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. ISSI Newsletter, 6(1), 22–28.

Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. Scientometrics, 107(2), 357–368.

Sivertsen, G. (2016a). Publication-Based Funding: The Norwegian Model. In: M. Ochsner et al. (eds.), Research Assessment in the Humanities: Towards Criteria and Procedures, Springer International Publishing, 71–90.

Sivertsen, G. (2016b). Data integration in Scandinavia. Scientometrics, 106, 849–855. doi: 10.1007/s11192-015-1817-x

Sivertsen, G. (2017). Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective, Palgrave Communications, 3, 17078.

Sivertsen, G. (2018a). Balanced multilingualism in science. BiD: textos universitaris de biblioteconomia i documentació, 40.

Sivertsen, G. (2018b). The Norwegian Model in Norway. Journal of Data and Information Science, 3(4), 2–18.

Sivertsen, G. (2019). Developing Current Research Information Systems (CRIS) as data sources for studies of research. In Glänzel, W., Moed, H.F., Schmoch, U., Thelwall, M. (Eds.), Springer Handbook of Science and Technology Indicators. Cham: Springer, 667–683.

Sivertsen, G., Rousseau, R., & Zhang, L. (2019). Measuring Scientific Production with Modified Fractional Counting. Journal of Informetrics, 13(2), 679–694.

Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. Scientometrics, 91(2), 567–575.

Sivertsen, G., & Schneider, J. (2012). Evaluering av den bibliometriske forskningsindikator, Nordisk institutt for studier av innovasjon, forskning og utdanning. Rapport 17/2012. URL: http://ufm.dk/forskning-og-innovation/statistik-og-analyser/den-bibliometriske-forsknings indikator/endelig-rapport-august-2012.pdf.

Torres-Salinas, D., Bordons, M., Giménez-Toledo, E., Delgado-Lopez-Cozar, E., Jiménez-Contreras, E., & Sanz-Casado, E. (2010). Clasificación integrada de revistas científicas (CIRC): Propuesta de categorización de las revistas en ciencias sociales y humanas. El profesional de la información, 19(6), 675–683.

Verleysen, F.T., Ghesquière, P., & Engels, T.C.E. (2014). The objectives, design and selection process of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW). In W. Blockmans et al. (eds.) The use and abuse of bibliometrics. Academiae Europaea; Portland Press, 115–125.

van den Besselaar, P., Heyman, U., & Sandström, U. (2017). Perverse effects of output-based research funding? Butler's Australian case revisited. Journal of Informetrics, 11(3), 905–918. https://doi.org/10.1016/j.joi.2017.05.016.

van Eck, N.J., Waltman, L., van Raan, A.F.J., Klautz, R.J.M., & Peul, W.C. (2013). Citation Analysis May Severely Underestimate the Impact of Clinical Research as Compared to Basic Research. PLoSONE. https://doi.org/10.1371/journal.pone.0062395

Verleysen, F., & Rousseau, R. (2017). How the Existence of a Regional Bibliographic Information System can Help Evaluators to Conform to the Principles of the Leiden Manifesto. Journal of Educational Media and Library Science, 54(1), 97–109. https://doi.org/10.6120/JoEMLS.2017.541/0011.BC.AC

National Lists of Scholarly Publication Channels: An Overview and Recommendations
for Their Construction and Maintenance

Janne Pölönen et al.

Research Paper

Wahlfors, L., & Pölönen, J. (2018). Julkaisufoorumi-luokituksen käyttö yliopistoissa. Hallinnon Tutkimus, 37(1), 7–21.

Walters, W. (2017). Do subjective journal ratings represent whole journals or typical articles? Unweighted or weighted citation impact?. Journal of Informetrics, 11(3), 730–744.

Wilsdon, J., Allen, L., Belfiore, E. et al. (2015). The Metric Tide. Report of the Independent Review of the Role of Metrics in Research Assessment and Management, HEFCE. https://doi.org/10.13140/RG.2.1.4929.1363

Wouters, P., Sugimoto, C., Larivière, V. et al. (2019). Rethinking impact factors: Better ways to judge a journal. Nature, 569(7758), 621–623. doi:10.1038/d41586-019-01643-3

Zacharewicz, T., Lepori, B., Reale, E., & Jonkers, K. (2018). Performance-based research funding in EU Member States—A comparative assessment. Science and Public Policy, 46(1), 1–11.

Zhang, L., Rousseau, R., & Sivertsen, G. (2017). Science deserves to be judged by its contents, not by its wrapping: Revisiting Seglen's work on journal impact and research evaluation. PLoS ONE, 12(3), e0174205. doi:10.1371/journal.pone.0174205

Zhang, L., & Sivertsen, G., 2020. The new research assessment reform in China and its implementation. Scholarly Assessment Reports, 2(1), 3. doi:10.29024/sar.15

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