Conflicting Measures and Values: How Humanities Scholars in Australia and Sweden Use and React to Bibliometric Indicators

Björn Hammarfelt
Swedish School of Library and Information Science, University of Borås, Allégatan 1, Borås, 501 90, Sweden. E-mail: bjorn.hammarfelt@hb.se

Gaby Haddow
Department of Information Studies, School of Media, Culture & Creative Arts, Curtin University, GPO Box U1987, Perth, Western Australia 6845. E-mail: G.Haddow@curtin.edu.au

While bibliometric indicators, such as the journal impact factor, have long played an important role in many STEM disciplines it has been repeatedly shown that established bibliometric methods have limited use in the humanities. Using a questionnaire on metrics use and publication practices in Australia and Sweden, we tested the assumption that indicators play a minor role among humanities scholars. Our findings show that our respondents use indicators to a considerable degree, with a range of indicators and rankings being employed. The scholars use metrics as part of institutional policy, in CVs and applications, as well as for general promotion of their work. Notable in our results is that a much larger share of researchers (62%) in Australia used metrics compared to Sweden (14%). Scholar's attitudes regarding bibliometrics are mixed; many are critical of these measures, while at the same time feeling pressured to use them. One main tension described by our respondents is between intradisciplinary criteria of quality and formalized indicators, and negotiating these "orders of worth" is a challenging balancing act, especially for younger researchers.

Introduction

When it comes to bibliometric measures, the humanities have often been portrayed as the "other" that does not fit into the "bibliometric universe." Dissimilar publication patterns, the dependence on local languages and contexts, as well as referencing practices are distinctive features that make traditional citation analysis less applicable in the social sciences and humanities (Hicks, 2004; Hammarfelt, 2016). However, the differences between humanities disciplines in this regard should not be underestimated, and quite large variances in publication and citation patterns are also found within specific fields (Verleysen & Weeren, 2016). The inability of traditional bibliometric measures to assess research in the humanities has resulted in attempts to find alternative ways of measuring the impact of research. For example, it has been suggested that a publication-based model that tries to take the characteristics and heterogeneity of disciplines into account might, at least partially, be a solution (Sivertsen, 2010). Despite these efforts, many scholars in the humanities and the social sciences (SSH) remain skeptical of bibliometric measures (Hammarfelt & de Rijcke, 2015).

Bibliometric measurement continues to play an important role for assessing research, and the number of available indicators grows constantly. At the same time, the critical literature on the negative, or rather "constitutive effects" (Dahler-Larsen, 2011), of bibliometric measures has developed (de Rijcke, Wouters, Rushforth, Franssen, & Hammarfelt, 2016). The current literature focuses not only on effects on the macrolevel of publication patterns across disciplines or nations, but also tries to depict how measures come to be ingrained in local evaluation procedures and in the actual production of knowledge (Rushforth & de Rijcke, 2015). In this article we take a similar stance when proposing that the informal, and often individual, use and perception of bibliometric indicators is important to consider when understanding the influence that these measures might have. It is possible, for example, that the use of indicators in research proposals or job applications is more apparent for the

Received July 7, 2017; revised January 5, 2018; accepted March 12, 2018

 • Published online May 2, 2018 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.24043

individual researcher than national systems for resource allocation.

While the limitations of using bibliometrics in the humanities are well established, we still know very little about their role within the humanities and few studies have so far investigated how scholars use indicators in these fields. The focus of research in this area has predominately looked at institutional and individual responses to national evaluation systems (Aagaard, Bloch, & Schneider, 2015), while less formalized use of metrics, instigated by researchers themselves, has mostly gone unnoticed. Moreover, studies into the influence of bibliometric measurement on humanities scholars have tended to focus on either the publications produced or on researchers' perceptions of research evaluation measures. The first set of literature uses publication data as the subject of analysis (Engels, Ossenblok, & Spruyt, 2012; Ossenblok, Engels, & Sivertsen, 2012) with findings that indicate that SSH researchers are publishing more in the English language. In the second area of study, surveys of academics have been used to gain insights into their perceptions and practice in relation to research metrics (Aksnes & Rip, 2009; Derrick & Gillespie, 2013; Rousseau & Rousseau, 2017). However, this latter literature is strongly centered on science fields, with very few studies involving researchers in the humanities (Lewis & Ross, 2011).

Hence, the purpose of this study was to examine how scholars in the humanities use metrics, and the effects that indicators and evaluation systems might have on publication practices in the humanities. Three main questions are asked: (i) To what extent do researchers in the humanities use bibliometric indicators? (ii) In which contexts are these indicators used? (iii) Which effects, on research and publication practices, are indicators deemed to have according to humanities scholars? While providing direct causality between measures and possible effects is beyond the ambition of our study, we do try to distinguish a few factors that could be of particular interest when trying to understand the relationship between measurement and effects. In line with Dahler-Larsen (2011), we propose that the consequences of single evaluation systems are hard to distinguish from other influences such as disciplinary conventions and broader international trends. Therefore, we suggest that it is more appropriate to discuss the possible effects of indicators using the concept of "metric culture" (Hammarfelt & de Rijcke, 2015). Our definition of metric culture includes both formal systems of evaluation, such as national allocation systems or models for distributing funds within a university, and more informal valuation practices within fields and disciplines. The concept "metric culture" takes into consideration how researchers themselves use metrics, for example in applications for research grants or when presenting themselves in CVs or online. Moreover, our definition of "metrics" is not limited to "bibliometric indicators" but also incorporates other types of numerical data, such as views, downloads, and social media mentions.

Our study is based on an online survey distributed to SSH scholars in Australia and Sweden. The selection of two national contexts provides an opportunity to explore the role of different evaluation systems, national academic cultures, and language on researchers' publication practices and use of research metrics. These two countries share some similarities—both have had, or still have, national evaluation systems with publication statistics or the number of citations as a component—but they differ considerably in the organization of research, academic traditions, and language. Many humanities fields in Sweden are international in how research is produced and disseminated, and disciplines like history, which formerly had more of a national profile, are now also targeting a broader international audience. Still, for scholars in countries where English is not the native language, tensions between demands for dissemination in international (read English language) channels and efforts to remain nationally relevant are prevalent (Salö, 2016). Considering the importance of national context for humanities disciplines, we suggest that it is useful to extend our study to a country where English is not the native language. In this case, we decided that Australia and Sweden would be two suitable contexts to compare. By surveying researchers in both countries we provide a comparative overview across disciplinary and national boundaries, which is complemented by free text answers that allows for a more detailed and contextualized approach.

The article is structured as follows: first, we describe the background to our study by reviewing some key strands in the research on effects of indicator use with a particular focus on humanities fields, then we introduce our methodology and study design, followed by a presentation of our findings. The article concludes with a discussion in which key findings are highlighted and possibilities for future research are outlined.

Background

The "humanities" is understood here as a set of related research fields rather than as an ideal or an idea. The term "humanities" is related to the German "Geistenwissenschaften" or the French "sciences humaines"; however, these are broader concepts in terms of the disciplines included, yet more narrow as they are limited to the scholarly (scientific) domain. In this study we adhere to the definition given by the Organisation for Economic Co-operation and Development (OECD, 2007). It should be noted that the definition of the humanities differs between contexts (see Leydesdorff, Hammarfelt, & Salah, 2011), and institutional arrangements may, at the local level, have considerable impact on how a field is defined. Overall, the humanities are a heterogeneous group of fields that are not easily delineated, and major differences exist between journal-based disciplines and more book-based fields.

Bibliometrics and the Humanities

Generally, bibliometric research on the humanities has been preoccupied with questions regarding the inadequate coverage of publications in citation databases. The scant

DOI: 10.1002/asi

coverage is usually explained by diverse publication channels, the importance of "local" languages, as well as the wide-ranging audience of humanities research (Nederhof, 2006; Hammarfelt, 2016). Considerable interest in this area has been directed to publication patterns, and especially the role and fate of the scholarly monograph (Thompson, 2002). Overall, studies show that monographs—although they amount to a reasonably small share of the total output (Kyvik, 2003)—play a central role in many disciplines. Recently, there has been a concern that the share of monographs is decreasing due to the implementation of evaluation systems that emphasize journal output, yet a distinct move from books to journal articles has not been found (Engels, Ossenblok, & Spruyt, 2012). Nevertheless, many scholars in book-oriented fields claim that a "devaluation" of the monograph has resulted in a move towards journal articles. Some findings also suggest that the pressure to publish in international journals might be more evident for younger scholars who are yet to retain permanent positions (Hammarfelt & de Rijcke, 2015).

Humanities Scholars and Their Use of Metrics

Studies of how scholars use and make sense of metrics are still relatively scarce, and the findings so far almost exclusively concern medicine and the natural sciences. Overall, previous studies show that researchers are ambivalent in their view of citations and other bibliometric indicators (Buela-Casal & Zych, 2012; Hargens & Schuman, 1990; Derrick & Gillspie, 2013). Aksnes and Rip (2009) found that researchers were knowledgeable about citations and their use in evaluating research, yet they claimed to not keep track of their own citations. Aksnes and Rip (2009) also found that researchers formed so-called "folk theories" about citations. For example, some respondents suggested that the citation rate of a publication is dependent on the status of the author rather than "quality," and having a fashionable topic was seen as an important factor for being cited (Aksnes & Rip, 2009).

Only a few studies have directly focused on humanities scholars and their attitudes to metrics, but the findings so far point to a fairly critical stance towards bibliometric indicators and metrics-based evaluation systems. An example of this view comes from a historian in the survey conducted by Hammarfelt and de Rijcke (2015, p. 74) who stated: "I know quite a lot about bibliometric evaluation but I ignore it. It is a crazy system developed for other disciplines than my own." Similar sentiments are not uncommon and the very idea of quantified assessment is for many scholars alien to their view of quality (Hug, Ochsner, & Daniel, 2013). While the critical perspective is striking in previous research on this topic, it is also important to point to a wide diversity across fields. Moreover, it is essential to distinguish between general attitudes to metrics and compare these to more detailed descriptions of how humanities scholars use indicators in their professional life.

Material and Methods

This study builds on data from a questionnaire that explored publication practices and the use of metrics by social scientists and humanities scholars at Australian and Swedish universities. A set of questions about country, academic position, time in academia (years since first publication), and gender opened the survey. For the purpose of comparison across fields we asked the respondents to identify their field using the *Field of Science and Technology Classification* published by the OECD (2007) and the Australian and New Zealand Standard Research Classification (Australian Bureau of Statistics, 2008). Then followed a range of questions regarding research output, publication practices, awareness of indicators (including a set of well-known indicators, such as the journal impact factor), and use of metrics (for the complete survey, see Appendix 1). Our data were drawn from multiple-choice questions as well as more open-ended questions, and this approach allowed us to combine a quantitative overview with more qualitative insights. Similar to Hammarfelt and de Rijcke (2015), we posit that this ". . .allows the individual scholar to emerge, not only as a number in a graph but also as a voice in an ongoing discourse" (p. 67). Moreover, text-box entries tend to increase the quality of responses (van Selm & Jankowski, 2006). In total, the questionnaire consisted of 17 questions and took between 15–20 minutes to complete.

We used the *Qualtrics* software to distribute the survey, which was tested on a small sample of SSH researchers before being disseminated. Ethical approval to conduct the study was gained from Curtin University. Our strategy for distributing the survey was initially to target all senior academic line-managers at relevant departments, schools, and faculties in Australia and Sweden and ask them to distribute the survey to all employees. Hence, we first addressed scholars who had "legitimate authority" in their organization, which has been suggested as a strategy for increasing response rates (Dilman, 2007). The complete population of researchers in the selected fields was targeted in order to avoid "sampling errors" (Sills & Song, 2002), and we were careful not to exclude scholars. However, the diversity of organizational units, as well as the time required to undertake this work, meant that some inclusion judgments were made. After the deadline, which was set to 4 weeks, it was obvious that some managers had not distributed the survey, as there were zero responses from eight Australian and four Swedish universities. We then sent out a second round of invitations to scholars at the nonresponding universities; this time by manually collecting e-mail addresses of all scholars in SSH at these institutions. The universities were representative of the sector, as they included research-intensive universities, as well as more teaching-oriented and regional universities. In total, over 6,000 e-mail addresses for scholars in Australia and Sweden were located; one nonresponding Australian university was excluded because e-mail addresses were not readily available. A second round of invitations was sent to large groups of scholars at

DOI: 10.1002/asi

TABLE 1. Overview of data.

Country	Australia	Sweden	Total				
	157	263	420				
Gender	Female	Male	Do not identify as either	Do not wish to specify	Total		
	197	214	1	8	420		
Age	<5 years	5–10 years	11–15 years	16–20 years	20 >	Total	
	54	100	77	57	130	418	
OECD-Field	Humanities (general)	History and archaeology	Languages and literature	Philosophy, ethics and religion	Art	Humanities (other)	Total
	67	126	131	59	36	48	467

TABLE 2. Use of metrics across research fields.

	Have you used any of these evaluation tools in your CV, promotion applications, grant applications, etc.		
Field	Yes	No	Total
Humanities general	20	39	59
	34%	66%	
History and archaeology	32	77	109
	29%	71%	
Languages and literature	40	73	113
	35%	65%	
Philosophy, ethics and religion	19	33	52
	37%	63%	
Art (arts, history of arts, performing arts, music)	8	25	33
	24%	76%	
Other humanities	12	28	40
	30%	70%	
Total	131	275	406
	32%	68%	100%

each university in early June 2016, with a closing date 4 weeks later.

In all, 1,029 responses were recorded, and for the present study 420 responses from scholars identifying themselves as being in the "humanities" were used. This can be compared to the total number of researchers employed in the humanities at universities in Sweden (2,563) and Australia (6,790) (Swedish Higher Education Authority, 2017; Turner & Brass, 2014). The number of respondents is relatively small considering the total population and similar surveys targeting academics—for example, Mohammadi, Thelwall, and Kousha (2016)—illustrates the difficulty in attracting responses. Given that our survey is relatively extensive (17 questions) and demanding (several open-ended questions) we find the number of responses to be satisfactory. However, we had a substantially higher response rate among Swedish scholars, and we have not been able to find any viable explanation to why this is the case. We also have a much larger share of women responding in Australia (56%) compared to Sweden (41%). Finally, there is always a risk of self-selection bias when respondents are included based on willingness to participate. Generally, this would mean that scholars with a strong opinion on the use of metrics would be more inclined to answer our survey.

For the field analysis we used the OECD classification in full, which lists. "humanities (general)" as the main category, with "history and archaeology," "literature and languages," "philosophy, ethics, and religion," "art (arts, history of arts, performing arts, and music)" and "humanities other" as subcategories. This choice resulted in some respondents choosing to define themselves as "humanities general" (the overarching category) instead of "humanities other." Thus, some respondents might have chosen the broader category instead of using the subcategories. An overview of response rates in specific fields is presented in Table 1. As respondents were given the opportunity to select several fields, the data might include researchers who define themselves as bordering on the social sciences, and it also means that the total number of selected fields is higher than the total number of respondents. It should also be noted that we deliberately did not include institutions with a more direct focus on creative arts (such as Stockholm University of the Arts).

Qualtrics provides functions, such as filters and cross-tabulations, which were used for most of the quantitative data analysis, including testing for statistical significance for variables such as country, academic age, gender, and field. Additional analysis was performed using Excel. The qualitative data were separately coded and analyzed by both researchers. Key terms and relevant quotes that reflected a particular attitude/use were independently extracted and analyzed. In a few cases quotes were reread in order to reach an agreement on definitions and interpretations.

Findings

Overall, one third (32%) of all respondents stated that they had used metrics for evaluation or self-promotion in applications and CVs. However, the extent of use varies between fields, from almost 37% in "philosophy, ethics, and religion" to 24% in "art" (Table 2). While not being able to pinpoint precise explanations for these differences, we deem it likely that publication practices, that is, the extent to which researchers are publishing in journals rather than books and book chapters, is one possible reason.

DOI: 10.1002/asi

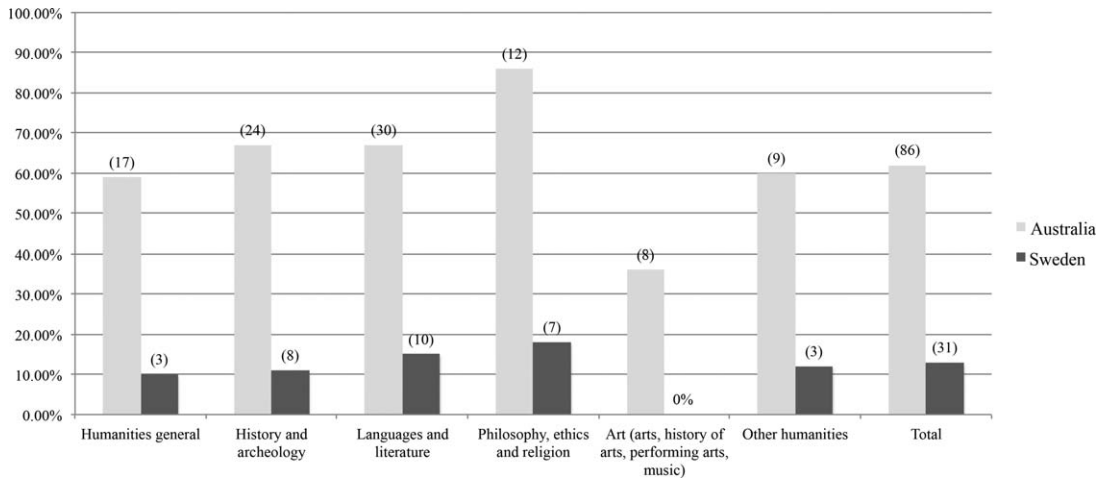

FIG. 1. Percentage of Australian and Swedish researchers using metrics across fields (total numbers in parentheses).

In the case of Sweden, the language of publication is a further factor to consider, as bibliometric databases generally have a much greater coverage of English sources (van Leeuwen, 2013). The quite remarkable difference in the uptake of metrics by Australian and Swedish researchers—only 14% of the Swedish respondents indicated that they had used metrics, with two thirds (62%) of the Australians stating that they used metrics of some kind—seems to support this interpretation. The large difference between Australian and Swedish scholars' use of indicators was also found to be statistically significant (p-value $< .001$ and chi-square 92.15). In comparison, large differences in the use of evaluation tools between Australia (68%) and Sweden (31%) where found also in the social sciences (Haddow & Hammarfelt, forthcoming). Again, it is difficult to pinpoint and separate specific factors that contribute to a greater reliance on metrics in particular fields and national context, and, as we reason in the Conclusion, it might very well be a range of factors that advance the use of indicators and bibliometric measures.

When contextualizing our findings regarding field differences, it is important to consider that the OECD classifications cover broad domains rather than specific fields. For example, when looking at the category "languages and literature" we note that it comprises two disciplines that to a large degree differ in publication patterns, with literature mainly being oriented towards books (monographs and chapters), while the journal article is the primary publication channel in linguistics (Thompson, 2002; Georgas & Cullars, 2005). In total, 35% of all respondents in "languages and literature" used metrics in their work, but a more detailed inquiry would probably reveal distinctive differences in indicator use between the two fields of linguistics and literary studies. Importantly, the differences between OECD fields could not be tested for statistical significance and the variance is quite small overall (Figure 1).

Scholars in the domain of "philosophy, ethics, and religion" used metrics to the greatest extent, with almost 37% respondents stating that they had used indicators or rankings. The importance of (international) journal publishing in these fields, especially in philosophy and ethics, is one possible explanation for the relatively frequent use of metrics. The use of metrics was lower among scholars in "history and archaeology" (29%), but like the results for "languages and literature" we should be mindful that archaeology has a high degree of journal publications indexed in the Web of Science (WoS), and coauthorship is much more common among archaeologists compared to historians (Verleysen & Weeren, 2016). Hence, our choice of using the OECD classification does to a certain extent hide variances within these rather broad domains and more detailed delineation of fields might provide further insights into how metrics are used in specific field and subfields. However, such studies would require substantially more data than the 420 responses that we had in our sample. Overall, we find that many humanities scholars use metrics, differences in metrics use across disciplines seems, at least tentatively, to relate to publication practices, and the uptake of indicators is substantially higher in Australia compared to Sweden.

A Range of Measures

Scholars in the humanities stated that during their career they have used a range of measures, including the h-index, impact factors, and various journal rankings (Figure 2). Specific providers such as Google Scholar, WoS, Scopus, and rankings such as, Scimago Journal rank (SCImago, 2007), the ERA list (Genoni & Haddow, 2009), and the Norwegian list (Sivertsen, 2016) are mentioned by the 112 researchers who provided detailed comments on specific tools. In addition, they made use of social networking sites such as Academia.edu, ResearchGate, and the digital author identifier ORCID. Some also utilized user data; for instance, the number of downloads or the number of views of their publications.

The data that scholars make use of come mainly from two types of providers: citation databases and social networking sites. Google Scholar together with WoS were the most frequently mentioned providers by our respondents;

DOI: 10.1002/asi

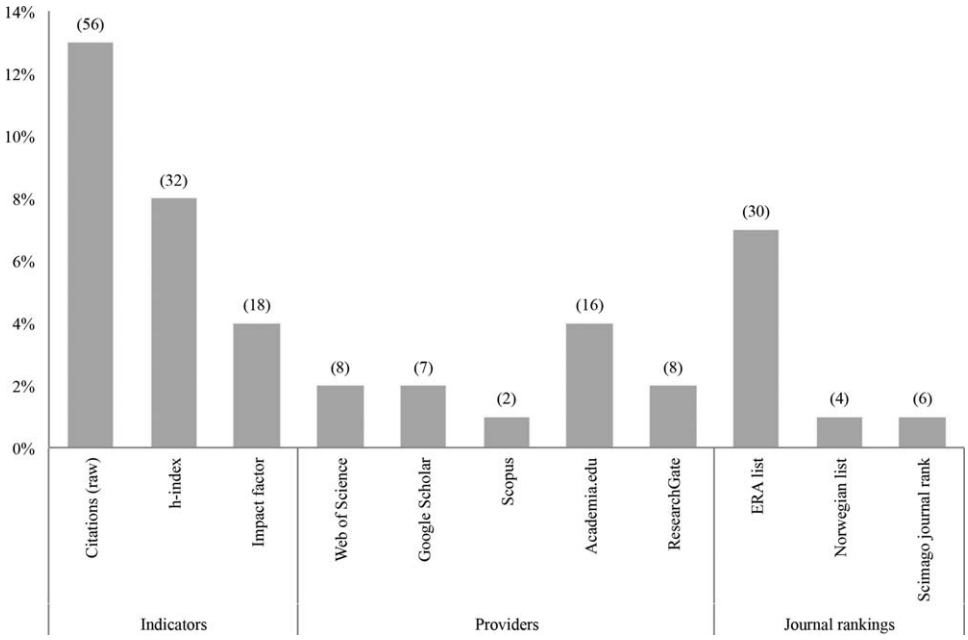

FIG. 2. Types of measures and services used across all fields. Percentage of respondents using a specific service or measure (total numbers in parentheses).

however, as indicated in our results (with 56 mentions of citations), many didn't specify where they get their data.

Social networking sites, such as ResearchGate and Academia.edu, provide scholars with information on how other users have viewed and downloaded their publications. Our respondents quite frequently mentioned Academia.edu particularly (16 times), and one explanation to why usage data are popular among humanities scholars could be that traditional measures, like citations, are generally less applicable in these fields. The stronger focus on articles in ResearchGate, as well as Thelwall's and Kousha's (2017) finding that this platform has low coverage in the humanities, might be an explanation to why Academia.edu has more users than its competitor. For some, the kind of monitoring that these platforms offer is enthusiastically received:

> . . .academia.edu allows me to see that my work is being widely looked at or read. Academia.edu also has the advantage I can see the country of origin of those who have viewed my publications (Australian, associate professor, History and Archaeology).

Several researchers also pointed to Academia.edu being an effective way of showcasing and disseminating their work to a wide audience, and some were using it on a daily basis. However, there was also criticism of the limitations of commercial platforms, and one respondent chose to use the local publication database instead, as it is ". . . not restricted to as academia.edu to 'members' only [sic]" (Swedish, postdoc, History and Archaeology).

Journal rankings, and especially the Excellence in Research for Australia (ERA) ranking, is a third influential tool for evaluating research. Rankings, in contrast to the Web of Science, for example, were from the beginning developed with the purpose of assigning value by listing items hierarchically. This means that most rankings have been developed for a specific purpose in a particular context and their development is highly politicized (Pontille & Torny, 2010). The ERA ranking is one such case, as it was developed as part of the larger evaluation initiative (ERA) in 2008. In the process of developing indicators for the ERA, possible methods for constructing a ranking of humanities journals were discussed, and concerns were raised that such an exercise would particularly harm fields that rely on local journals (Genoni & Haddow, 2009). The ERA ranking was abandoned in 2012, and it is not used in the current evaluation system, yet it still seems to play an important role when evaluating research. Sweden has never used journal rankings to evaluate research on a national level, but a dozen universities use a two-tier classification of journals developed in Norway (Hammarfelt et al., 2016). This, the so-called Norwegian list, is mentioned by some of our respondents, although its visibility is quite limited compared to the ERA ranking.

Use of Metrics

The most common use of metrics is in applications for funding and 23 respondents mentioned this, while 20 stated that they had used indicators when applying for a job or for promotion. While indicators are primarily used for job applications or when applying for research grants, the newer types of services are also used for dissemination and self-monitoring, reflected in the comment: "ResearchGate and Academia to promote my research publications DIRECTLY to my research peers" (Australian, lecturer, History and

DOI: 10.1002/asi

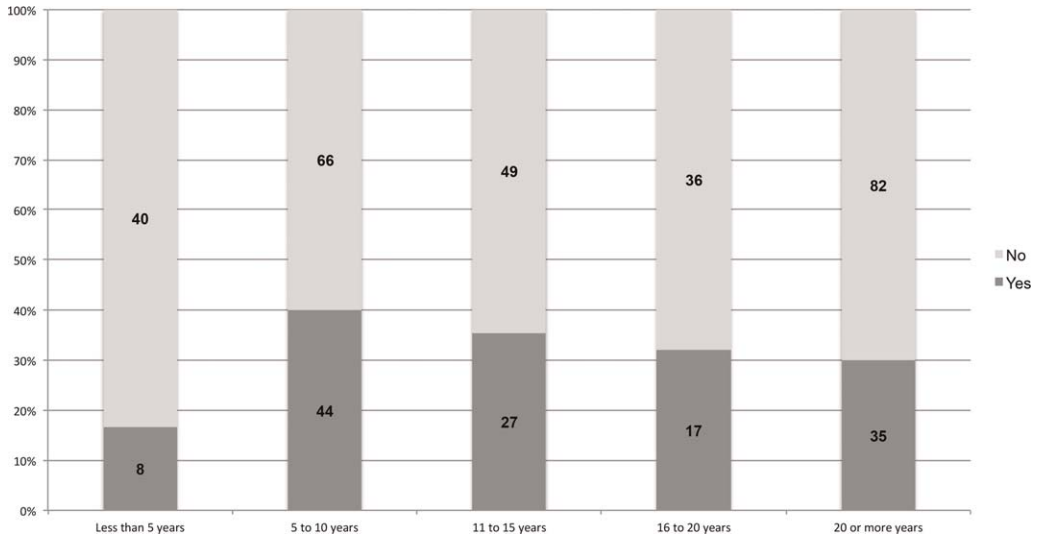

FIG. 3. Academic age and use of metrics (total number in bold).

Archaeology). To some degree it appears that this use, which is instigated by scholars themselves to monitor their own research, is more positively viewed.

Often, however, indicators are perceived as something that is imposed by the system and mandated from above as part of "department policy," as in this example; "More journal articles instead of monographs to increase number of publications. Thinking about classification of journal (Norwegian list) since this is very much stressed" (Swedish, assistant professor, Arts, History of Arts, Performing Arts, Music). Institutional demands are indeed mentioned by many and was the main reason for using bibliometrics, yet at times indicators are also employed to counteract institutional policies, as in this case where a researcher argued for the use of citations and the h-index: "Only reason is to ward off any criticism that journal articles are not in A* or A ranked journals. High citations of articles in decent B-level journals should count for something, even though department policy is to discourage and/or ignore them" (Australian, associate professor, History and Archaeology). Here a nonmandated indicator has been used to counterbalance institutional policies. Thus, an effective, as well as empowering method, of resisting a specific indicator is to bring in additional measures that provide alternative interpretations.

Criticism of Indicators

Respondents questioned the use of indicators and many are frustrated with bibliometric measurement. One scholar framed the problem in terms of the difficulty of ". . .translating our sense of value to people who want quantitative data" (Australian, unspecified position, History and Archaeology). Some are even more outspoken critics of indicators, as in this rather pessimistic account: "We are also currently working in a system that one could define as an unholy marriage of market capitalism and Stakhanovist micro-management, where we are forced to produce a definite number of outputs each and every year, as if we are caged hens in a barn. Despite all the rhetoric, quantity continues to trump quality" (Australian, associate professor, History and Archaeology). For these researchers there is a mismatch between their "values" and the kind of measurement that indicators offer. The second quote also identifies the specific values systems, "capitalism" and "state bureaucracy," which according to this respondent underpins bibliometric assessment.

Despite the harsh criticism, many scholars felt that they need to use these indicators. One respondent expressed it clearly when responding to the question regarding what measures she used: "Citations, Web of Science, Google Scholar. Very much dislike these methods of evaluations as they don't work for the humanities" (Australian, assistant professor, Humanities general), and similar sentiments are voiced by another respondent "Again, the Aussie criteria (h-index, etc) are at odds with disciplinary practices, but being in that system I do have to use them" (Australian, professor, History and Archaeology). These quotes illustrate that humanities scholars still felt the need to use indicators, despite their unease when doing so and the fact that the indicators do not reflect the scholars' notion of quality.

Quite a few researchers claimed to be knowledgeable about bibliometric measures, but chose to ignore them. Among the more senior scholars especially this was a common position: "Fortunately, I am old and consequently I do not have to bother. Had I been young, I would have left academia rather than adjust to the moronic 'publishing strategies' forced upon us" (Swedish, assistant professor, History and Archaeology). While some enjoy the privilege of ignoring these measures, others may not have the same opportunity. It is indeed logical that younger scholars, with less secure positions, should be most concerned with bibliometric measurement, and this conclusion is supported by earlier findings (Hammarfelt & de Rijcke, 2015). Looking at respondents overall, there are some signs that young

DOI: 10.1002/asi

scholars, with the exception of PhD students, are the most active users of metrics (Figure 3). It makes sense, as this period is one of the more competitive periods in an academic career and our findings indicate that scholars who are in the phase of establishing themselves as independent researchers, having an academic age of 5–15 years, were the most frequent users of metrics. The low turnout in the youngest cohort—those having published their first publication less than 5 years ago—can partly be explained by a considerable share of this group being PhD students rather than established academics, and thus having less use of metrics for promotions or applications. Still, our findings regarding indicator use and academic age are not statistically significant ($p = .26$ and chi-square 5.25).

A further factor that we explored was whether gender affected the use of metrics. Among our respondents it was more common that women (37%) had used metrics compared to men (26%). However, this result is not statistically significant, and when correcting for the proportionally larger response rate among Australian women, we find little evidence for differences in metric use being related to the gender of the researcher.

Indicators and Publication Practices

Several themes can be discerned in the answers regarding changes in publication practices. The proportion of different types of publications, for instance, books versus journals, is a recurring leitmotif in these accounts. Several respondents point to a move from books and book chapters to journal articles, although there are others who see little change in this regard. Many Swedish scholars in "history and archaeology" noted that researchers in their fields tend to publish more in English and less in Swedish. Criticism of this development is expressed in quite a few responses, but it is also evident that some scholars welcomed the change: "Shift to publishing in English and emphasis on publishing in peer review journals (thank god). Hopefully this decreases the degree of laziness among researchers in my field" (Swedish, postdoc, History and Archaeology).

When discussing publication practices, other humanities scholars pointed to similar changes as found for "history and archaeology"; that is, a trend towards more journal articles and, in the case of Sweden, a further emphasis on English-language publications. There were also suggestions that there is a focus on shorter texts, again in relation to journal publications, and that it is getting harder to publish books. A few respondents described strained communication systems where quality control, through, for example, peer review, is compromised: "Very many journals are currently launched, due to what I perceive as a very stressed situation concerning getting as many publications out as possible. It is frustrating that quite a few of these 'new' journals seem to take peer review processes lightheartedly" (Swedish, associate professor, Languages and Literature).

Scholars also expressed greater awareness of indicators when they reflected on their own publication practices: "I've become more aware of the value attached to channels of publication and have adapted my publication practice accordingly, practically [sic] guidance given by the Norwegian list" (Swedish, postdoc, History and Archaeology).

Digital media (blogs and tweets) as well as open access were also mentioned as important changes in scholarly communication. However, there was a concern that blogs and twitter are not recognized or valued as legitimate outputs, although they can be used to promote "proper" publications. In a similar vein, it was recognized that although open access is encouraged, these publication venues are generally ranked lowly in formal evaluation systems. Many scholars in the humanities communicated with a wider audience, mentioning social media as additional arenas for communication; however, there were others who expressed a concern that this activity was not adequately rewarded: "... there is more expectation that we engage in alternative forms of research dissemination (for instance, blogs, articles in *The Conversation* etc.) but no corresponding change in the requirement to publish in peer-reviewed journals/with university presses. As a result, there is just an increase in workload for not much gain" (Australian, lecturer, Humanities general). Overall, there was a sense of frustration that only "traditional" outputs are valued and, as the quote below demonstrates, open-access publishing was seen as empowering in relation to both scholarly activity and bibliometrics:

> My own [publication practices] has been very influenced by the bibliometry race, but I have recently started to question this. Peer review is not what it promises to be, journal articles means that you have to hack up your work in publishable chunks rather than in pieces that make sense and the "dead time" factor is impossible to predict which means that it is impossible to plan your working life.... I have just decided to publish pretty much everything concerning a specific project as a digital monograph rather than as a string of articles. Many colleagues seem to consider this "an act of resistance," but to me it is simply an act to set my brain and my results free (Swedish, assistant professor, Languages and Literature).

While there were many examples of how individual researchers navigated intradisciplinary demands in our material, there were also examples of meta-reflections regarding the field in question. For example, this Swedish literature scholar suggested that the field is currently recalibrating its publication practices:

> Literary studies have traditionally been a nationally based discipline with nationally based publishing patterns and structures. The language used has been Swedish. Since a couple of years the discipline is moving towards a more international publishing pattern with English as preferred publishing language. At present, thus, the discipline is trying to combine these two publishing strategies, the national strategy and the international strategy (Swedish, professor, Languages and Literature).

DOI: 10.1002/asi

It should be acknowledged that some scholars reported that they had seen no distinct changes in publishing practice over the last 10 years and, looking at comprehensive data on publication patterns in the humanities in Sweden (Uppsala University) and Norway, this stance has some empirical support; the proportion of books seems to be quite stable (Hammarfelt & de Rijcke, 2015; Sivertsen, 2016). However, findings suggest that the proportion of English-language outputs in the humanities is increasing rather rapidly, as shown by studies of publication patterns in Norway and Flanders (Ossenblok, Engels, & Sivertsen, 2012), and locally in Sweden (Hammarfelt & de Rijcke, 2015). But the extent of this change varies considerably between disciplines, and large differences in publication patterns might also be evident within specific fields (Verleysen & Weeren, 2016).

Finally, while some respondents reported little or no change in publication patterns in their field, others suggested that a focus on indicators has had an effect. Clearly, there are other factors, such as increasing digitalization of research dissemination in the humanities as well as a push towards internationalization and open access, which has to be taken into consideration. Hence, based on our findings it is not possible to draw strong conclusions regarding the effects that bibliometric indicators have on publication patterns. The concerns voiced by many of our respondents, however, do indicate that these measures are highly visible and several scholars point to undesirable consequences for knowledge production and dissemination.

Discussion

The most noteworthy finding in our study is the rather extensive use of metrics in the humanities; about one-third of all researchers in our survey used metrics in some form. This finding is especially interesting, as major critique and warnings against using bibliometric indicators in the humanities has been voiced both from the research field of bibliometrics and from humanist scholars themselves. It appears, though, that many users of metrics are well aware that established indicators do not align with publication patterns and traditional evaluation practices in their own fields. Thus, our respondents showed considerable ambivalence regarding these measures; on one hand, there is skepticism or direct criticism of their use; on the other, they appeared to use bibliometric indicators for showcasing their achievements.

It is clear that there are large differences between fields in their use of indicators, and it might in many cases be too simplified to discuss the use of metrics in the context of the "humanities" as a whole. The OECD classification used here might hide field-specific characteristics, and detailed studies would undoubtedly shed more light on the uptake of bibliometrics in specific fields and subfields. Our limited response rate and the possibility of self-selection bias influencing our results also make it difficult to draw strong and definite conclusions regarding the differences between research fields. Similarly, it was difficult to draw definite conclusions regarding the influence that gender or academic age might have on the use of metrics among humanities scholars. However, the major difference in metric use found in this study is not between fields, but between national contexts. Almost two thirds (62%) of the Australian respondents stated that they had used metrics, compared to only 14% of their Swedish colleagues. These figures point to bibliometric indicators being an integrated part of academic practice among Australian scholars in the humanities, while bibliometrics are used by a quite limited group of humanist researchers in Sweden. Explanations for these findings can only be hinted at here and we suspect that many factors may contribute to this difference in metric use. A broader sociological and historical inquiry into the academic systems in Sweden and Australia would be needed to reveal some of these factors, and we consider three possible explanations to be of particular interest to study further. First, the fact that Swedish is the main language of publication among many humanist scholars may make it difficult or even impossible to use conventional bibliometric measures to evaluate research in Sweden; second, the early implementation and subsequent discussions regarding a national system for bibliometric evaluation, including a ranking of journals, in Australia could have led to greater awareness regarding these measures, and to them being integrated into local assessment procedures in universities and departments; and third, the career system is perhaps more competitive in Australian academia (for example, through greater mobility and inflow of scholars from abroad). Additionally, we suggest that comparisons with other national contexts are likely to uncover further factors that affect the uptake of indicators among humanist scholars.

Overall, we can identify a range of types of indicator use in our material. Metrics were used for strategic purposes (in CVs, applications, and promotions) and respondents to our survey also pointed to them being mandated in institutional policies and formal evaluation systems. Yet our findings also show that services such as Academia.edu or ResearchGate may be used informally, as tools for keeping track of research achievements and audience, while bibliometric measures, for example, citations, can serve the purpose of offsetting or criticizing other indicators and journal rankings.

The great variety in measures used and the quite rapid uptake of alternative indicators, such as usage counts and services like Academia.edu, is noteworthy in our findings. Several researchers pointed to these measures as being useful for showcasing research and in situations where citations and other more established measures are less applicable. In fact, the use of alternative indicators can actually be seen as a means for researchers to take control over how they are evaluated, and, in this sense, they might have an empowering function (Hammarfelt, de Rijcke, & Rushforth, 2016). It is also evident that indicators can be used strategically to counter other evaluative devices; for example, by showing that an article in a "low-ranked" journal has gathered a substantial number of citations. Overall, we find that the measures and indicators used by scholars in our sample rarely align with one specific system of evaluation, but rather we

DOI: 10.1002/asi

see a complex and rich ecology of measures. The most evident impact of national systems is the ERA ranking of journals in the Australian system, a list that is now abandoned but still has major influence. The use of the ERA list suggests that evaluation systems that have been discarded can survive and travel into new contexts; we might call this the "afterlife of indicators."

The effects that bibliometric measures have on publication practices are difficult to isolate from other external or intradisciplinary developments, and it is particularly difficult to determine the effects of specific indicators or systems. Looking at our results, we find that some respondents in our survey claimed to have changed their personal publication strategies based on, or under the "guidance of," indicators or journal rankings. Others, often the more senior researchers, suggested that the field or their colleagues have changed publication practices, and the move towards English-language publications is a recurring theme in the Swedish context (and one that is supported by previous studies of publication patterns). We would find it surprising if metrics did not have an influence on research practices and publication strategies in an Australian setting, as almost two thirds of the Australian respondents reported using metrics. However, this requires further study, not only in regard to publication types—for example, books versus journals—but also on the epistemological effects that indicators might have. Hypothetically, a further focus on international journal publications could lead to time-consuming projects being abandoned, and more "international" topics might be chosen over local matters.

Leaving the discussion on distinct "effects," it is evident that scholars in the humanities experience tensions between intradisciplinary criteria of research quality and indicators of various kinds; they use indicators and at the same time consider them incommensurable with their notions of quality. The situation can be tackled in different ways, and some respondents claimed to ignore bibliometric measures, while others voiced criticism. The level of agreement between what indicators measure and what is valued in a specific field varies, but for the researchers in our study there was a substantial mismatch, and similar sentiments have also been found among literature scholars and art historians (Hug, Ochsner, & Daniel, 2013). It is, as one respondent formulated it, a problem of translating one type of value into another. Following Boltanski and Thévenot (2006), we could argue that rivalling orders of worth meet when bibliometric measures are applied to the humanities. The values of bibliometric measures and journal rankings correspond foremost to the "world of the market" and "industry," in which competitiveness, free circulation (global), and statistics are key representations. On the other hand, traditional criteria for evaluation in many disciplines in the humanities are more aligned with "the world of the domestic" and the "inspired," which in Boltanski's and Thévenot's account relate to reputation, trustworthiness, the local, authority, singularity, and creativity. Competing "orders of worth" are at play in many contexts and the need to seek justification according to different values is not unique for the humanities. Still, it could be argued that the current trend towards quantification of research performance, what we also call "metric culture," becomes especially troublesome in the humanities, as these disciplines have traditionally relied on "orders of worth" that cannot easily be aligned with bibliometric indicators. Our findings indicate that scholars often need to navigate between these value systems and this balancing act seems to be particularly challenging for young researchers. Finally, compromises and negotiations between these systems of valuation have to be formulated and structured by researchers themselves, but studies of this kind can provide some guidance and overview of the problems and conflicts that are at hand.

Acknowledgment

Gaby Haddow would like to acknowledge Curtin University's study program support for enabling this study. We also thank Livia Haddow for valuable advice in statistical matters, and we are grateful for the constructive comments on our article by Sven Hug and the anonymous reviewer. Finally, we would like to express our gratitude to all scholars who generously responded to our survey. This research received support from the Swedish Research Council (Grant Number 2013-7368).

References

Aagaard, K., Bloch, C., & Schneider, J.W. (2015). Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. Research Evaluation, 24(2), 106–117.

Aksnes, D.W., & Rip, A. (2009). Researchers' perceptions of citations. Research Policy, 38(6), 895–905.

Australian Bureau of Statistics (2008). Australian and New Zealand Standard Research Classification (ANZSRC).

Boltanski, L., & Thévenot, L. (2006). On justification: Economies of worth. Princeton, NJ: Princeton University Press.

Buela-Casal, G., & Zych, I. (2012). What do the scientists think about the impact factor? Scientometrics, 92(2), 281–292.

Dahler-Larsen, P. (2011). The evaluation society. Palo Alto, CA: Stanford University Press.

de Rijcke, S., Wouters, P., Rushforth, A., Franssen, T., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. Research Evaluation, 25(2), 161–169.

Derrick, G.E., & Gillespie, J. (2013). "A number you just can't get away from": Characteristics of adoption and the social construction of metrics use by researchers. In Hinze, S. & Lottman, A. (Eds.) Berlin: Proceedings of the 18th International Conference on Science and Technology Indicators, pp. 104–116.

Dilman, D.A. (2007). Mail and internet surveys: The tailored design method. (2nd ed.). Hoboken, NJ: Wiley.

Engels, T.C., Ossenblok, T.L., & Spruyt, E.H. (2012). Changing publication patterns in the social sciences and humanities, 2000–2009. Scientometrics, 93(2), 373–390.

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

Georgas, H., & Cullars, J. (2005). A citation study of the characteristics of the linguistics literature. College & Research Libraries, 66(6), 496–515.

Haddow, G. & Hammarfelt, B. (forthcoming). Quality, impact and quantification: Indicators and metrics use by social scientists.

DOI: 10.1002/asi

Hammarfelt, B. (2016). Beyond coverage: Toward a bibliometrics for the humanities. In Ochsner, M., Hug, S.E., & Daniel, H.D. (Eds.), Research assessment in the humanities: Towards criteria and procedures (pp. 115–131). Berlin: Springer International Publishing.

Hammarfelt, B., & de Rijcke, S. (2015). Accountability in context: Effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of arts at Uppsala University. Research Evaluation, 24(1), 63–77.

Hammarfelt, B., de Rijcke, S.D., & Rushforth, A.D. (2016). Quantified academic selves: The gamification of research through social networking services. Information Research, 21(2), 21–22.

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(3), 292–305.

Hargens, L.L., & Schuman, H. (1990). Citation counts and social comparisons: Scientists' use and evaluation of citation index data. Social Science Research, 19(3), 205–221.

Hicks, D. (2004). The four literatures of social science. In Moed, H.F., Glänzel, W., & Schmoch, U. (Eds.) Handbook of quantitative science and technology research (pp. 473–496). Dordrecht, Netherlands: Springer.

Hug, S.E., Ochsner, M., & Daniel, H.D. (2013). Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history. Research Evaluation, 22(5), 369–383.

Kyvik, S. (2003). Changing trends in publishing behaviour among university faculty, 1980–2000. Scientometrics, 58(1), 35–48.

Lewis, J.M., & Ross, S. (2011). Research funding systems in Australia, New Zealand and the UK: Policy settings and perceived effects. Policy & Politics, 39(3), 379–398.

Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the Arts & Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. Journal of the American Society for Information Science and Technology, 62(12), 2414–2426.

Mohammadi, E., Thelwall, M., & Kousha, K. (2016). Can Mendeley bookmarks reflect readership? A survey of user motivations. Journal of the Association for Information Science and Technology, 67(5), 1198–1209.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. Scientometrics, 66(1), 81–100.

Organisation for Economic Co-operation and Development. (2007). Revised field of science and technology (FOS) classification in the Frascati manual. Paris: Working Party of National Experts on Science and Technology Indicators, Organisation for Economic Co-operation and Development (OECD).

Ossenblok, T. L., Engels, T. C., & 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. (2010). The controversial policies of journal ratings: Evaluating social sciences and humanities. Research Evaluation, 19(5), 347–360.

Rousseau, S. & Rousseau, R. (2017). Being metric-wise: Heterogenity in bibliometric knowledge. El profesional de la información, 26(3), 480–487

Rushforth, A., & de Rijcke, S. (2015). Accounting for impact? The Journal Impact Factor and the making of biomedical research in the Netherlands. Minerva, 53(2), 117–139.

Salö, L. (2016). Languages and linguistic exchanges in Swedish academia. Dissertation. Stockholm University.

SCImago. (2007). SJR — SCImago Journal & Country Rank. Retrieved October 4, 2017, from http://www.scimagojr.com

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

Sills, S.J., & Song, C. (2002). Innovations in survey research: An application of web-based surveys. Social Science Computer Review, 20(1), 22–30.

Sivertsen, G. (2016). Publication-based funding: The Norwegian Model. In Ochsner, M., Hug, S.E., & Daniel, H.D. (Eds.), Research assessment in the humanities: Towards criteria and procedures (pp. 79–90). Berlin: Springer International Publishing.

Swedish Higher Education Authority. (2017). Employees in higher education 2016. 1654–3475 (Online).

Thelwall, M., & Kousha, K. (2017). ResearchGate articles: Age, discipline, audience size, and impact. Journal of the Association for Information Science and Technology, 68(2), 468–479.

Thompson, J.W. (2002). The death of the scholarly monograph in the humanities? Citation patterns in literary scholarship. Libri, 52(3), 121–136.

Turner, G., & Brass, K. (2014). Mapping the humanities, arts and social sciences in Australia. Canberra, Australia: Academy of the Humanities.

van Leeuwen, T. (2013). Bibliometric research evaluations, Web of Science and the social sciences and humanities: A problematic relationship? Bibliometrie-Praxis und Forschung, 2, 1–18.

van Selm, M., & Jankowski, N. W. (2006). Conducting online surveys. Quality & Quantity, 40(3), 435–456.

Verleysen, F. T., & Weeren, A. (2016). Clustering by publication patterns of senior authors in the social sciences and humanities. Journal of Informetrics, 10(1), 254–272.

Appendix 1: Survey

Thank you for linking to our survey about publication patterns and consequences of research evaluation for social scientists and humanities scholars in Australia and Sweden. Your participation in this research is greatly appreciated.

The survey asks for information about your research outputs and your field's outputs. However, the data collected are not going to be used to analyse an individual's productivity, but rather to examine research output of fields as a consequence of research evaluation tools. Please note that a few questions include items specific to the Australian or Swedish academic environment. These have been marked AUS and SWE, respectively.

As outlined in the invitation e-mail, all responses are anonymous and linking to the survey will not identify participants in the software. Responses will be treated as confidential and used only for the purposes of this project, which will be written up for publication and conference presentations. We ask for the name of your university to allow us to explore potential links between types of universities and responses in our analysis, but institutions will not be identified in any materials that are produced.

Australian respondents will be invited to indicate if they are willing to participate in a follow-up interview. The e-mail contact details will be used only for that purpose and that information will be deleted from the stored survey results data file.

By clicking to the next page you are indicating your consent to participate in our survey.

—

1. In which country are you an academic?
Australia/Sweden

DOI: 10.1002/asi

2. What is the name of your university? (free text)

3. What is your academic position?

Professor/associate professor (reader)/assistant professor (lecturer)/lecturer/postdoc/doctoral student/research assistant/other

4. Gender

Female/male/do not identify with any of the above/do not wish to specify

5. Please select up to three research classification terms from both of the lists below that best describes your field of research.

OECD classification (whole list)

Field of research codes (ANZSRC) (AUS) (Whole list)

6. What is your preferred language when publishing?

English/Swedish/German/Dutch/French/Spanish/Other (please specify)

7. Researchers in your field most often publish: (Rank first three by selecting an item and using the arrow button)

Book chapters

Conference articles

Creative works

Journal articles

Monographs

8. On average, how many of the following types of research-related outputs do you publish each year? Please use whole numbers regardless of whether your outputs are sole or co-authored works.

Blog posts/Book chapters/Book reviews/Creative works/Conference articles (published in full)/Edited books/Monographs/Newspaper articles/Peer reviewed journal articles Peer review reports, for instance, article reviews/Policy reports/Research proposals/Trade/professional journal articles/Tweets Wikipages/Other, please specif

9. How important are the following factors to you when choosing a publication channel?

(Extremely important/Very important/Somewhat important/Not important/Not familiar with this)

Peer review processes/Quality of peer review/Reputation of publishing channel/Open access/Publication speed/Maximum outreach to target audience/International visibility/National visibility/Department policy/Institutional policy/National policy/Requirements of funding agencies/Indexing by international database (eg Scopus/Web of Science)/Journal Impact Factor/Scimago Journal Rank (SJR)/Journal ranking lists/Inclusion in Norwegian list (SWE)/Inclusion in ERA Journal List (AUS)/Field of Research code assignment (AUS)

10. Are there other factors that are important to your selection of publication channels? (please specify)

11. How familiar are you with the following tools for evaluating research?

Citations (Web of Science, Scopus, Google Scholar)/Impact factor/Scimago Journal Rank/h-index/Norwegian list (SWE)/ERA journal list (AUS)/Other journal rankings/ratings/Other journal/publisher rankings/Altmetric services (Impact Story/Altmetrics.com)/Social platforms (Research Gate, Academia.edu, etc)

12. Have you used any of these evaluation tools in your CV, promotion applications, grant applications, etc.?

Yes/No

13. If yes, please list the tools you have used and for what purpose

14. Please comment on any changes in publication practice in your field over the last ten years

15. Please comment on any changes in your own publication practice over the last ten years

DOI: 10.1002/asi