

## Towards an understanding of the relationship between disciplinary research cultures and open access repository behaviours

**Jenny Fry** (Corresponding author)

School of the Arts, English and Drama, Loughborough University, Epinal Way, Loughborough, Leicestershire, LE11 3TU, UK. Tel: +44 (0)1509 223074. Email: [J.Fry@lboro.ac.uk](mailto:J.Fry@lboro.ac.uk)

**Valérie Spezi**

Centre for Information Management, School of Business and Economics, Loughborough University, Epinal Way, Loughborough, Leicestershire, LE11 3TU, UK. Tel: +44 (0) 1509 635690. Email: [V.C.L.Spezi@lboro.ac.uk](mailto:V.C.L.Spezi@lboro.ac.uk)

**Steve Probets**

Department of Computer Science, Loughborough University, Epinal Way, Loughborough, Leicestershire, LE11 3TU, UK. Tel: +44 (0) 1509 228055. Email: [S.G.Probets@lboro.ac.uk](mailto:S.G.Probets@lboro.ac.uk)

**Claire Creaser**

Centre for Information Management, School of Business and Economics, Loughborough University, Epinal Way, Loughborough, Leicestershire, LE11 3TU, UK. Tel: +44 (0) 1509 635682. Email: [C.Creaser@lboro.ac.uk](mailto:C.Creaser@lboro.ac.uk)

**Abstract:** This paper explores the cultural characteristics of three ‘open access (OA) friendly’ disciplines (physics, economics and clinical medicine) and the ways in which those characteristics influence perceptions, motivations and behaviours towards green OA. The empirical data are taken from two online surveys of European authors. Taking a domain analytic approach, the analysis draws on Becher and Trowler’s (2001) and Whitley’s (2000) theories to gain a deeper understanding of why open access repositories (OAR) play a particularly important role in the chosen disciplines. The surveys provided a unique opportunity to compare perceptions, motivations and behaviours of researchers at the discipline level with the parent metadiscipline. Albeit, participants were not drawn from a stratified sample of all the different sub-disciplines that constitute each discipline and therefore the generalizability of the findings to the discipline is limited. The differential role of informal and formal communication in each of the three disciplines has shaped green OA practices. For physicists and economists, preprints are an essential feature of their respective OAR landscapes, whereas for clinical medics final published articles have a central role. In comparing the disciplines with their parent metadisciplines there were some notable similarities/differences, which have methodological implications for studying research cultures.

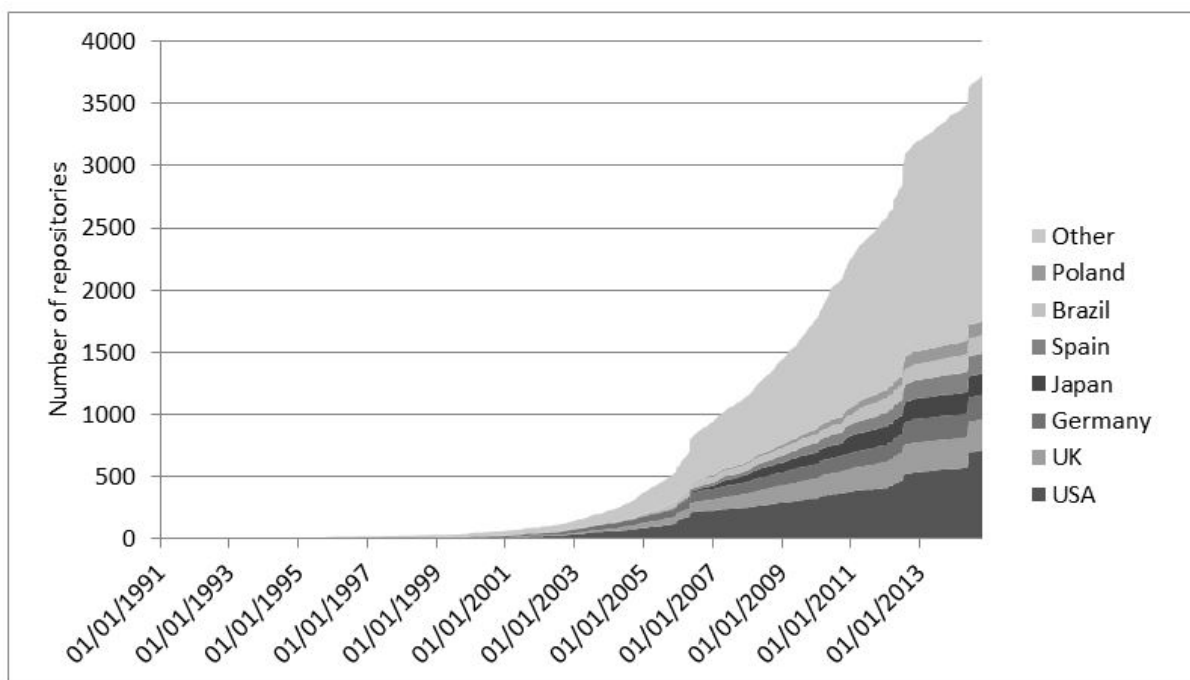
**Keywords:** Open Access, Open Access Repositories, researcher behaviours, domain analysis, research cultures, physics, economics and clinical medicine

## Introduction

In recent years there has been much debate about the most effective route to achieving Open Access (OA) to peer-reviewed published journal articles. The gold route, often referred to as the ‘author pays’ route, involves payment of an Article Processing Charge (APC) to publishers, enabling the article to be made available without subscription or access fee via an Open Access Journal (OAJ) or hybrid-OAJ<sup>1</sup>. The green route, often referred to as the ‘self-archiving’ route, entails authors submitting manuscripts to journals (subscription-based, OAJ, or hybrid-OAJ) but maintaining the right to upload a version of their work on a digital Open Access Repository (OAR). There have been various initiatives to foster both the green and gold route to OA with different countries, funding agencies, and research institutions implementing these in diverse ways.

During the past two decades there has been rapid growth in the number of OAR worldwide (Figure 1), with many of these repositories being classified as either Institutional Repositories (IR) or Subject-Based Repositories (SBR). An IR has been defined as “... a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members” (Lynch, 2003). By contrast SBR are not restricted to any single institution; they tend to focus on a specific discipline, or group of disciplines, and host outputs specific to the disciplinary focus of the repository.

Figure 1: Growth of OAR worldwide



(Source: ROAR)

Despite many initiatives supporting the ‘green’ or ‘self-archiving’ approach to Open Access, the slow rate of adoption of OAR by researchers, outside of specific ‘OA friendly’ disciplines,

such as *physics*, *economics* and *clinical medicine* where OA is well-established, has been acknowledged (Duranceau, 2008). Although Institutional Repositories have made a significant contribution to the number of OAR worldwide (Pinfield et al., 2014; Registry of Open Access Repositories, 2014) and according to the Ranking of Web Repositories<sup>ii</sup> some notable IR are ranked in the top twenty worldwide OAR, the deposit of material into OAR, and into IR in particular, remains problematic (Bernius, Hanauske, Dugall, Konig, 2013; Pinfield et al., 2014)<sup>iii</sup>. Xia (2008) argues that although there is a commonly held assumption that researchers familiar with depositing content into SBR are likely to adopt similar self-archiving practices in relation to IR there is no evidence to support such an assumption.

Metadisciplines such as the physical sciences, engineering, medical sciences, social sciences and arts & humanities, are often used as units of analysis to study scholarly communication and information behaviours, such as the adoption and use of OAR. Choice of unit of analysis, e.g. the sub-discipline, discipline or metadiscipline, creates a tension in analytic granularity (Fry, 2006, p. 313), with fine-grained analysis, e.g. the sub-discipline, revealing specificities, whereas coarse-grained analysis e.g. the metadiscipline, reveals generalities. Many disciplines are heterogeneous, constituted from multiple sub-discipline cultures which can mask specificities and differences within any single discipline. As Becher argues (1989, p.50), decisions about unit of analysis in studying academic research cultures involve a 'trade-off', but since few studies have compared findings at various different levels it is difficult to know what is the precise nature of that 'trade-off'. In using the discipline as a unit of analysis, as opposed to the sub-discipline, the findings reported in this paper also embody a trade-off, with sub-disciplines being potentially under or over represented, thus leading to a limitation in the generalizability of the conclusions to the discipline.

This paper is based on the data collected in the Behavioural strand of the PEER (Publishing and the Ecology of European Research) programme which spanned an 18-month period in 2009-11. The behavioural strand compared researchers' perceptions, motivations and behaviours in relation to OAR at the level of four metadisciplines. We consider the following questions:

1. What are the similarities / differences in perceptions, motivations and behaviours relating to green OA within 'OA friendly' disciplines such as physics, economics and clinical medicine?
2. Can Whitley's (2000) and Becher and Trowler's (2001) theories of the cultural organisation of disciplines be used to explain why the OAR phenomenon has played a particularly important role in physics, economics and clinical medicine?
3. To what extent is the metadiscipline, e.g. physical sciences, engineering, medical sciences, social sciences and arts & humanities, a meaningful unit of analysis? In other words, can perceptions, motivations and behaviours within specific disciplinary communities be inferred from findings at the metadiscipline level and vice versa?

This paper brings a greater understanding of researcher attitudes and behaviours towards green OA, and the use of OAR in particular, by focusing on similarities as much as differences in patterns of behaviour; and by comparing behaviours at the level of the discipline and metadiscipline.

## **Background**

Two key imperatives have driven OA initiatives forward during the past decade: lengthy publication lags, particularly in the sciences and medicine (Cullen and Chawner, 2011), and the so-called 'serials crisis'. Galin and Latchaw (2010) argue that OAR represent the most accomplished form of support for a 'free borrowing culture' (p.219), but they also recognise that the adoption and use of OAR as a central resource has been concentrated in specific disciplines. For example, it is widely recognised that economists and physicists have a long-standing tradition, that pre-dates digital networked resources, of sharing preprints and working papers amongst their respective research communities and that tradition has been embodied in the development and adoption of the SBR RePEc and arXiv respectively.

### *Disciplinary perspectives*

Related research has shown that, historically, patterns of dissemination, publication and information behaviours vary across disciplinary boundaries (Bates, 2002; Becher and Trowler, 2001; Crane, 1972; Garvey, 1979; Harley, Accord, Earl-Novell, Shannon & King, 2010; Hjørland and Albrechtsen, 1995; Kling and McKim, 2000; Mote, 1962). Recent research related specifically to the uptake and use of OAJ and OAR also illustrates the role that disciplinary differences play in patterns of adoption and use (Björk et al., 2010; Dallmeier-Tiessen et al., 2010; Fry et al., 2011).

Awareness and adoption of OA, and the differentiation between gold and green OA in particular, vary across disciplinary communities. Björk, Laakso, Welling, and Paetau, (2014) argue that differences in the adoption of green OA can be explained by the existence of preprint cultures, well-established SBR, high-quality OAJ, and mechanisms for funding APCs. Not only are there differences in levels of awareness of OA and OAR between disciplines, but the modalities and practicalities of achieving OA, particularly green OA, also differ from one discipline to another (Spezi, Fry, Creaser, Proberts, & White, 2013). Recent research into the uptake of OA in the UK showed that, out of a sample of 85,215 articles published in 2010, 35% were self-archived in an OAR and 5% were published in OAJ (or hybrid journals) (Van Noorden, 2012); of those articles available via OAJ a pre-publication version may also be available in an OAR. Table 1 presents a breakdown of Van Noorden's (2012) findings by discipline / metadiscipline.

Table 1: Material available on OA according to discipline/metadiscipline (Van Noorden, 2012, p.303)

Disciplines	Material <i>not</i> available on OA	Material available on OA	
		Self-archived in OAR (green OA)	Published in OAJ (gold OA)
Clinical medicine	62%	34%	4%
Biomedical research	43%	43%	14%
Physics	48%	48%	4%
Social sciences	65%	34%	1%
Arts & Humanities	86%	13%	1%

In contrast to Van Noorden's (2012) findings, Björk et al. (2010) argue that gold OA is widespread and well developed in the Life and Medical sciences, whereas self-archiving, or green OA, is much more developed in disciplines such as the Earth sciences, Physics and Astronomy. There may be various reasons underpinning this contrast in findings, e.g. sample size, actual vs. self-reported behaviours, national vs. international perspectives (where the impact of OA mandates can affect behaviours), etc. The fact that well-known OAJ, such as those from PLoS (Public Library of Science) or BioMed Central, have a high visibility amongst research communities in the Life and Medical sciences, may reinforce the close association of OA with OA publishing within those disciplines. In their study, Dallmeier-Tiessen et al. (2010) found that STM (Science, Technology and Medicine) represents 66% of the body of pure and hybrid OAJ, and contributes 77% of all OA articles. Within this, Biology and Life sciences represent 19% of OAJ and 21% of articles, while Medicine and Health sciences represent 28% of OAJ and 28% of articles (*ibid*).

#### *Organisation and control of research*

The related literature discussed above has shown that research communities constituting different disciplines will adopt and adapt innovation in the scholarly communication system, such as OAR, in different ways. Furthermore, within the metadisciplines often used as units of analysis to generalise and differentiate between researchers' behaviours in different knowledge domains (e.g. Abreu, Grinevich, Hughes & Kitson, 2009; Fry et al., 2011; The Royal Society, 2011), we might expect to find marked similarities or differences both within and across such broadly conceived boundaries.

What is lacking in the related literature is systematic explanation as to how patterns of adoption and adaptation are related to the fundamental cultural characteristics of disciplines. In terms of explaining the culture of disciplines and sub-disciplines Whitley (2000) identifies two interrelated elements: the degree of interdependence between researchers within and between disciplines in making a valid contribution to existing knowledge (mutual dependence); and the degree of uncertainty in producing and evaluating that knowledge (task uncertainty). Mutual dependence and task uncertainty each have two

analytically distinct elements that relate either to reputational control (strategic mutual dependence and strategic task uncertainty), or to the coordination of research techniques, strategies and outcomes (functional mutual dependence and technical task uncertainty). Whitley (2000) argues that the combination of feasible variations in the degree of mutual dependence and task uncertainty generates seven major types of discipline or sub-discipline. Of particular relevance to the three disciplines being analysed in this paper are what Whitley (*ibid*) describes as 'partitioned bureaucracies' (economics), 'professional adhocracies' (clinical medicine) and 'conceptually integrated bureaucracies' (physics).

'Partitioned bureaucracies', such as economics, are highly rule governed and hierarchically organised. The standardisation of training programmes and skills leads to theoretical coherence, which in turn; enables the reputational elite to control what constitutes legitimate research problems and strategies (typically through peer review). The lack of technical control over empirical phenomena, however, threatens theoretical coherence. There is a distinct difference in research organisation between the analytical theoretical core and applied peripheral areas (accounting for the relatively low-degree of functional mutual dependence).

In 'professional adhocracies' there is a variety of funding sources and organizations where research is conducted – e.g. medics may be employed by private firms, universities, hospitals or may have joint appointments - and this heterogeneity is an integral element of a research culture where there is no single reputational group that researchers must address when developing their research strategies. This may account for relatively high levels of strategic task uncertainty, since judgements about the relevance and importance of outcomes are likely to vary from one reputational group to another. On the other hand, skills and technical procedures are highly standardized and technical task uncertainty is relatively low.

For 'conceptually integrated bureaucracies', of which physics is archetypical, the limited diversity of funding sources means that there is a relative scarcity of facilities and resources resulting in a high-degree of competition for access to critical apparatus and funds. Competing claims to the significance of research problems and strategies within the overall theoretical framework requires adjudication by some central authority, e.g. national funding agencies. Theoretical coordination of research is highly valued and is used as a mechanism by which to integrate the goals of sub-groups into a unified cognitive order that distinguishes between the central concerns of sub-groups and those of the discipline. Results are relatively predictable and the theoretical implications of research outcomes are relatively easy to discern (hence the relatively low levels of technical and strategic task uncertainty and the relatively high levels of functional and strategic mutual dependence).

Giving consideration to the finer-level of granularity of sub-disciplines, Becher and Trowler (2001) draw an analogy with 'urban/rural' landscapes and 'tightly/loosely' knit communities. Urban landscapes are typified by a narrow set of intellectual concerns whereby research

activity is clustered around a limited number of discrete topics. Rural landscapes, on the other hand, span a broader set of intellectual concerns with research problems thinly scattered across the breadth of concerns. Tightly knit communities tend to demonstrate convergent characteristics whereby there is a high-degree of intellectual control over research problems and procedures using uniform standards. In contrast, loosely knit communities demonstrate divergent characteristics where intellectual plurality is more likely to be accommodated.

Becher and Trowler's (2001) notion of convergent / divergent characteristics seems to be particularly relevant in terms of augmenting Whitley's (2000) characterisation of 'partitioned bureaucracies', 'professional adhocracies' and 'conceptually integrated bureaucracies'. Whitley's 'partitioned bureaucracies' (economics) and 'conceptually integrated bureaucracies' (physics) exhibit characteristics consistent with Becher and Trowler's description of 'urban landscapes' / 'tightly knit' communities, which they argue results in convergent behaviours, such as a high-degree of intellectual control over research problems. As Whitley (2000) observes, in a number of aspects there are similarities in the ways in which research is organised and controlled in physics and economics (pg. 185), perhaps differing most in the way in which theories are empirically tested. Use of metadisciplines as a unit of analysis would place economics and physics in quite different categories and perhaps mask similarities in behaviours e.g. the adoption of OAR. In contrast, Whitley's 'professional adhocracies' (clinical medicine) exhibit characteristics similar to Becher and Trowler's (2001) description of 'rural landscapes' / 'loosely knit' communities and given the diversity of research problems, resources and audiences would place clinical medics at the divergent end of Becher and Trowler's (2001) divergent/convergent continuum. The strong element of 'professionalism' that characterises the culture of clinical medicine, however, results in convergent behaviours whereby diverse goals and intellectually plurality can be coordinated. This indicates that the behavioural patterns of clinical medics relating to OAR are likely to be more convergent than other 'rural landscapes' / 'loosely knit communities'.

It should be noted, however, that even at the finer-grained level of the discipline there can be wide-variation in researcher behaviours across different sub-disciplines and therefore the unit of analysis, e.g. sub-discipline as opposed to discipline, will influence how a knowledge domain is externally categorised by researchers. For example, not all sub-disciplines of physics demonstrate characteristics of a 'conceptually integrated bureaucracy' and likewise, as indicated above, applied sub-disciplines of economics are likely to have characteristics not in keeping with an archetypal 'partitioned bureaucracy'. Furthermore, as Case (2012) notes, categorisation of interdisciplinary knowledge domains into any single aggregation can be problematic. Such typologies do, however, provide a useful device for describing and, more importantly, explaining similarities and differences within and across knowledge domains.

## Method

PEER was a pioneering collaboration between publishers, repositories and researchers from 2008 to 2012, to investigate the effects of the large-scale, systematic deposit of authors' final peer-reviewed manuscripts on reader access, author visibility, and journal viability (see <http://www.peerproject.eu/>). Supported by the European Commission's eContentplus programme, three separate research projects were commissioned, focussing on author behaviours, journal usage, and economic aspects.

The data gathering exercise for the behavioural aspect of the PEER programme comprised two different electronic surveys conducted approximately one year apart, a series of focus groups and a workshop. This paper is based on the data gathered in the two surveys and where appropriate draws on the qualitative data in an anecdotal way to illustrate specific points. Details of the Phase 1 survey questions and focus group protocol are available in the baseline report (Fry et al, 2009); details of the Phase 2 survey questions are available in the final report (Fry et al 2011).

Invitations to complete the surveys were sent out to research-active authors via twelve publishers participating in the study<sup>iv</sup>. Invitations were sent to those authors who had had an article(s) published in one or more of the 240 journal titles selected to be part of the PEER programme. These journals had a Thomson Reuters JCR Impact Factor and more than 20% European Union (EU) content ('EU content' is defined as having at least one author from an EU country). Collectively the selected journal titles covered the life sciences, medicine, physical sciences and social sciences & humanities and were distributed across a range of JCR impact factors<sup>v</sup>. Details of the invitations sent out to researchers are provided in Table 2. The surveys were intended for EU-based corresponding authors; however, publishers were not always able to restrict the distribution to their European authors, and so the surveys were, on occasion, distributed more widely. A filter question was used to confine the analysis to European authors only.

*Table 2: Survey distribution, by discipline*

	Phase 1		Phase 2	
Medical sciences	8,901	(25%)	10,145	(29%)
Life sciences	7,089	(20%)	9,040	(26%)
Physical sciences & mathematics	16,077	(45%)	10,567	(30%)
Social sciences, humanities & arts	3,627	(10%)	4,998	(14%)
Total	35,694	(100%)	34,750	(100%)*

\* Rounding errors may occur.

A total of 3,136 valid responses were received for the Phase 1 survey. The response for the second survey was lower, with a total of 1,426 valid responses – this difference does not affect the validity of the results presented in this paper. The sampling strategy described above was the same for both surveys and some of the respondents to the Phase 2 survey



may also have completed the Phase 1 survey. In all subsequent figures in the paper, the number of respondents giving each answer is shown in italics.

A total of thirty-four disciplines were represented in the survey responses which were mapped onto four metadisciplines (Table 3) based on the PEER categorisation of those journal titles included in the PEER research programme. Respondents who indicated research interests in two or more metadisciplines were categorised as interdisciplinary. The focus of this paper is on three specific disciplines *Clinical medicine*, *Physics* and *Economics* within the Medical sciences, Physical sciences & mathematics, and the Social sciences, humanities & arts metadisciplines, respectively. Data for these three disciplines are analysed and presented alongside their parent metadiscipline in order to gauge the extent to which those disciplines vary within their grouping of origin. Table 4 shows the maximum numbers of respondents in each disciplinary category (some respondents did not answer all the questions). Pearson’s  $\chi^2$  statistic was used to compare the distribution of responses between discipline groups, and between the specific disciplines and the remainder of the respondents from each metadiscipline. Results have been reported as showing a difference where the probability of obtaining the observed value of  $\chi^2$  (p- value; given here to one decimal place) was less than 5% ( $p < 0.05$ ).

*Table 3: Discipline mapping*

<b>a. Medical Sciences</b>	Clinical medicine Clinical dentistry Anatomy & physiology Nursing & paramedical studies Health & community studies Pharmacy & pharmacology Other
<b>b. Life Sciences</b>	Biosciences Psychology & behavioural sciences Earth, marine & environmental sciences Veterinary science Agriculture & forestry Other
<b>c. Physical Sciences &amp; Mathematics</b>	Chemistry Physics Mathematics General engineering Chemical engineering Mineral, metallurgy & materials engineering Civil engineering Electrical, electronic & computer engineering Mechanical, aero & production engineering Information technology & systems sciences & computer software engineering Other

<b>d. Social Sciences, Humanities &amp; Arts</b>	Architecture, built environment & planning Catering & hospitality management Business & management studies Economics Geography Social studies Media studies Humanities & language based studies History Archaeology Modern languages Design & creative arts Education & Sports Other
--	---

*Table 4: Disciplines of survey respondents*

		Phase 1	Phase 2
Broad disciplinary grouping	Medical sciences	248	194
	Life sciences	416	311
	Physical sciences & mathematics	1773	454
	Social sciences, humanities & arts	259	167
	Interdisciplinary	440	300
	<b>Total</b>	<b>3136</b>	<b>1426</b>
Specific disciplines	Clinical medicine	210	185
	Physics	1148	194
	Economics	100	69

Note: the row totals in subsequent figures may not equate to the data presented in Table 4 as not all questions were answered by all respondents. Further, the total number of respondents included in the specific sub-disciplines includes a number of researchers who were classified as interdisciplinary in the broader categories as they reported research areas which spanned more than one of the broader groups. This particularly affects the medical sciences and clinical medics. All statistical testing was carried out on independent groups; i.e. clinical medicine was compared to medical sciences excluding clinical medicine; physics to physical sciences excluding physics; and economics to social sciences humanities and arts excluding economics.

One of the questions underlying this paper is the extent to which the metadiscipline is a meaningful unit of analysis compared to the discipline. We should note that in using the discipline as a unit of analysis this may obscure specific behaviours at the level of the sub-discipline level, which in any one sub-discipline may be quite different to that of neighbouring sub-disciplines. Participants from each of the disciplines were not drawn from

a stratified sample of all the different sub-disciplines that constitute each discipline, which is not uncommon in behavioural studies of researchers. It does mean, however, that some sub-disciplines may be under represented and others may be over represented, limiting the robustness of the disciplinary generalizations reported. In relation to OAR this could make a difference to the conclusions, given that anecdotally we know that some sub-fields have adopted OAR to a greater extent than others. For example, in physics arXiv is a de facto centralised resource for the sub-discipline of high-energy physics, whereas this is not the case for all sub-disciplines, such as experimental solid state physics—this means that repository awareness and use may be under / over represented in our findings.

## Findings

### *The importance of peer-reviewed scholarly journal articles*

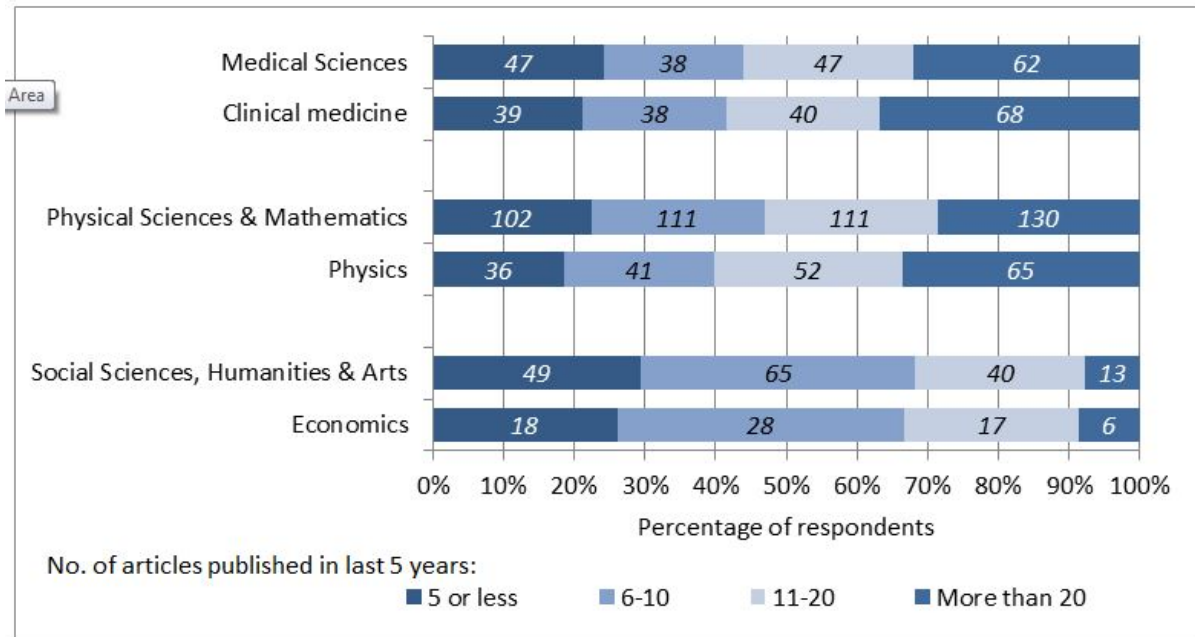
The majority of researchers from all three disciplines viewed peer-reviewed scholarly journal articles as important sources for their research. Respondents in Phase 1 were asked to rate the importance of different types of output as resources for their research from 1 to 5, where 1 = very important and 5 = not important at all. Table 5 illustrates that within all three disciplines peer-reviewed scholarly journals were rated as 1 or 2 (i.e. very important or important) by more researchers than other forms of output (including monographs, book chapters, conference proceedings, working papers<sup>vi</sup>).

*Table 5: Percentage of researchers rating selected types of output as important/very important resources for their research (Phase 1 survey)*

<i>Discipline</i>	<i>Scholarly peer-reviewed Journals</i>	<i>Monographs</i>	<i>Book Chapters</i>	<i>Conference proceedings</i>	<i>Working papers</i>
<i>Clinical medicine</i>	94%	29%	39%	48%	20%
<i>Physics</i>	97%	52%	59%	50%	18%
<i>Economics</i>	94%	43%	39%	24%	48%

Figure 2 gives details about the self-reported publishing rate of the three disciplines over a five-year period. Economists were less likely to have published more than 20 journal articles than clinical medics or physicists ( $\chi^2=25.4$ ,  $p<0.001$ ).

Figure 2: Approximately how many articles have you published in the last five years? (Phase 2 survey)



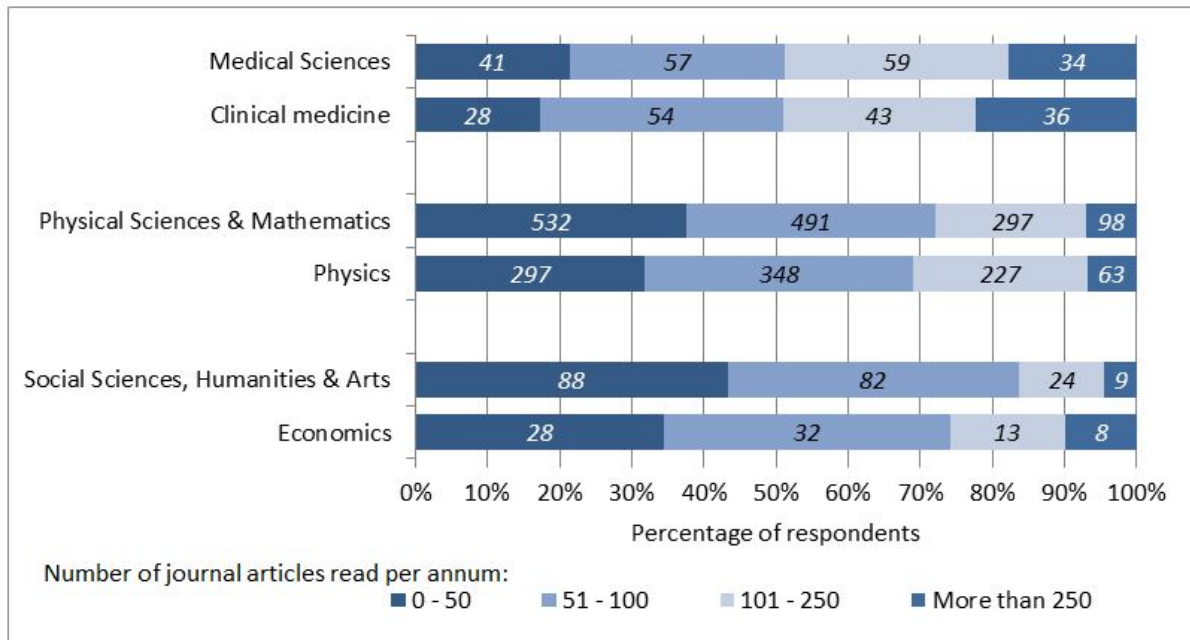
One factor that affects publication rates is multi-authorship (Gordon, 1980; Presser, 1980). A number of studies have noted a rise in levels of multi-authorship over time, although this varies by discipline. Based on a dataset taken from the *Web of Science* covering the time period 2000-2009, Gazni, Sugimoto, and Didegah (2012) found that in medicine on average 84% of articles published were multi-authored, with a similar average for the physical sciences (82%), but a substantially lower average for the social sciences (36%). Within each metadiscipline Gazni et al. (2012) found differences in the average percentage of multi-authored articles at the discipline level, e.g. clinical medicine 86%, physics 85%, and economics and business 48%.

Despite the reported rise in multi-authorship in economics (Ductor, 2015; Nowell & Grijalva, 2011; Piette & Ross, 1992; Schymura & Löschel, 2014) the frequency and extent is lower than in clinical medicine or physics. Another explanation for the lower self-reported publication rate amongst economists is the difficulty of getting articles published in economics, with frustrating iterations of revise and re-submit and lengthy publication lags once a manuscript has been accepted (Ellison, 2002). This is supported by Whitley's (2000) argument that the reputational elite within the analytical theoretical core of economics tightly control prestigious journal space.

The importance of peer-reviewed journal articles to researchers is further evidenced by Figure 3 which shows that almost 50% of researchers with a clinical medicine background read over 100 peer-reviewed journal articles per year, as do 31% of physicists and 26% of economists. Comparing the three disciplines with their parent metadiscipline, levels of journal article reading are about the same for clinical medics; the percentage of physicists

reading more than 100 papers per year is slightly higher than that of the parent discipline ( $\chi^2=25.6$ ,  $p<0.001$ ); and a higher percentage of economists are reading over 100 papers per year, compared to the parent discipline ( $\chi^2=9.98$ ,  $p<0.005$ ).

Figure 3: Approximately how many peer-reviewed journal articles do you read on average each year? (Phase 1 survey)

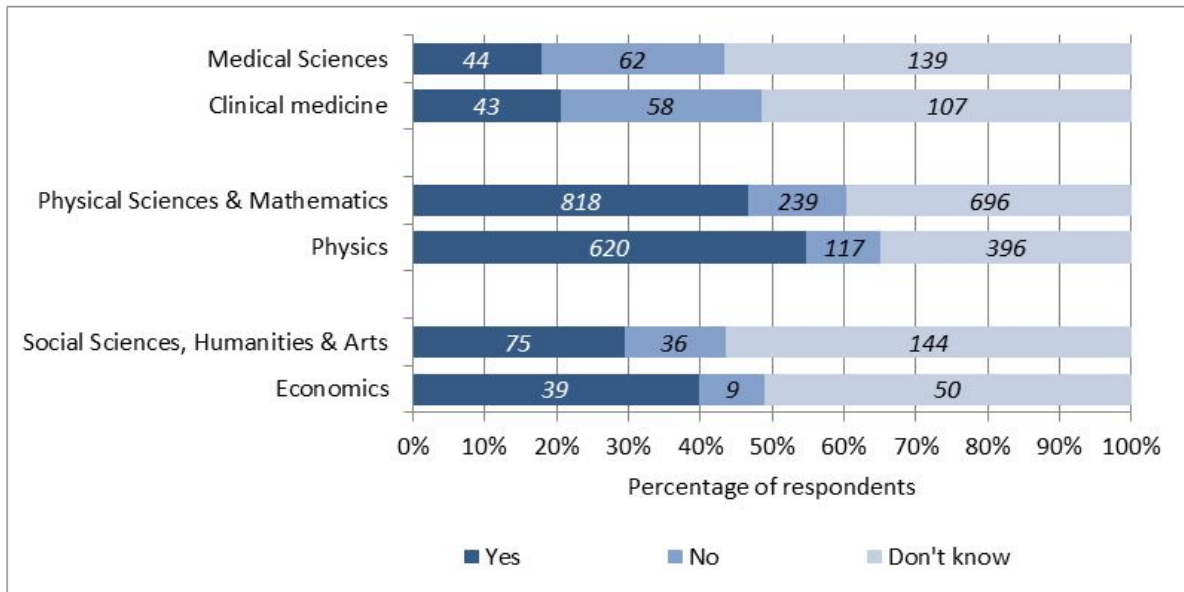


Although levels of reading are high across all disciplines, it is noticeable that economists and physicists read fewer peer-reviewed articles than researchers from clinical medicine ( $\chi^2=50.3$ ,  $p<0.001$ ). This corroborates the findings of Tenopir, Maysa and Wu (2011) and Tenopir, King and Bush (2004) that published journal articles are particularly important in the health sciences. In the case of economics, one possible reason for this could be related to the value that economists place on working papers. Table 5 (above) indicates that working papers are valued by a larger percentage of economists than by physicists and clinical medics ( $\chi^2=43.3$ ,  $p<0.001$ ). It would therefore seem understandable that amongst economists there are higher levels of readership of working papers.

#### Awareness of subject-based repositories

More physicists know of a suitable SBR for their research than clinical medics and economists, with clinical medics being the least aware of this kind of repository ( $\chi^2=103.7$ ,  $p<0.001$ ) (Figure 4). Researchers in physics were also more aware of a suitable SBR than researchers in the wider metadiscipline of Physical sciences & mathematics ( $\chi^2=31.3$ ,  $p<0.001$ ); apparent differences between the other disciplines and their parent metadiscipline were not statistically significant.

Figure 4: Are there any subject-based publicly available repository(ies) suitable for your research? (Phase 1 survey)

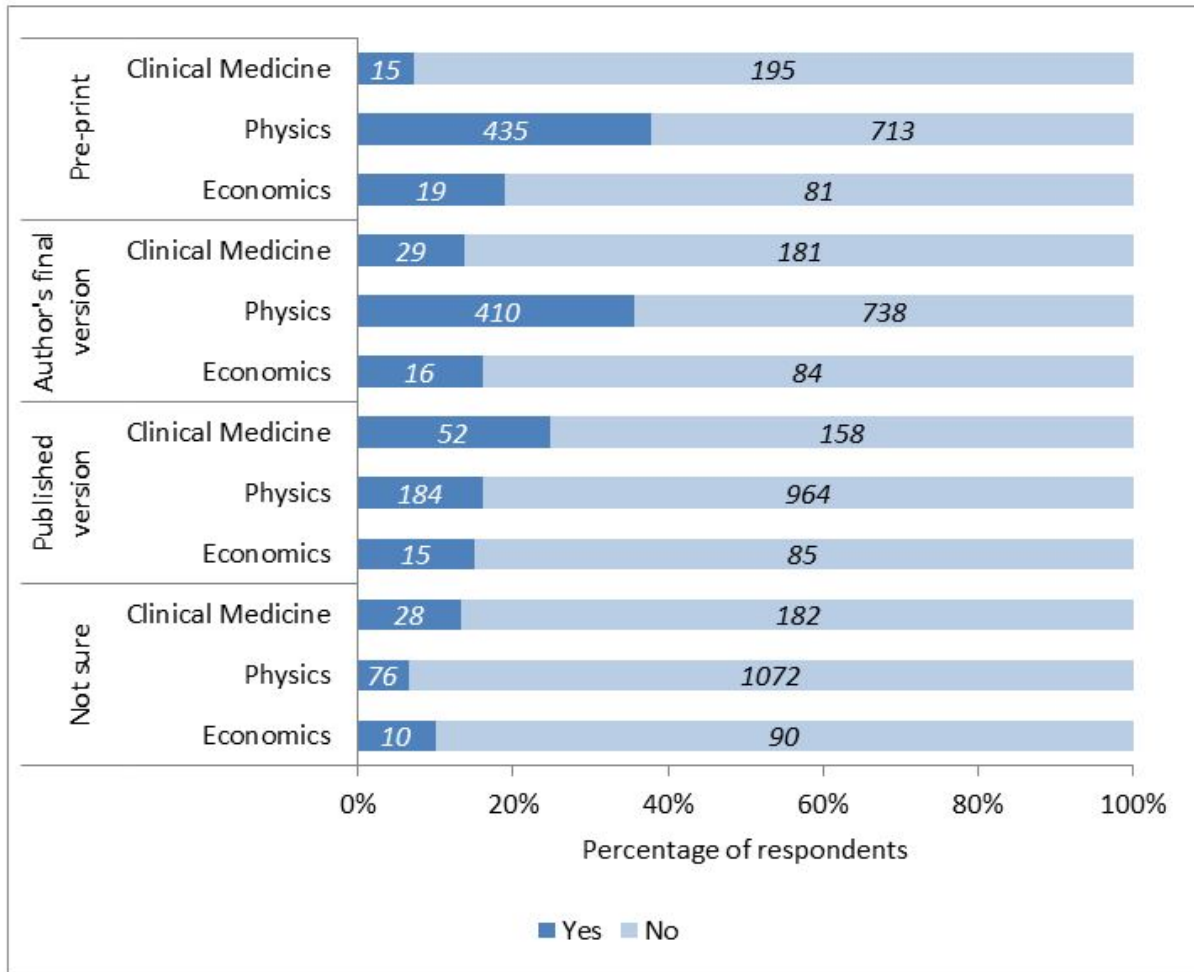


The low number of clinical medics who were aware of a suitable SBR for their articles (20%) is surprising given the success of PMC, but may be explained by differences in the deposit procedures between PMC and SBR in physics (arXiv) and economics (RePEc). With arXiv and RePEc, authors are often responsible for ‘self-archiving’ their material whereas, with PMC, it is primarily the participating publishers that upload material (potentially leading to lower levels of awareness) after appropriate embargo periods, although individual authors can also submit their final peer-reviewed journal manuscripts in compliance with the Public Access Policy mandated by the NIH<sup>vii</sup>.

*Self-archiving: version(s) of journal articles authors upload to OAR*

Repositories can contain various types of output. The three main types are working paper or preprint, i.e. the unrefereed version of a paper; the author’s final peer-reviewed manuscript, i.e. the accepted manuscript; and the published final version, i.e. the version of record. In terms of the versions of journal articles that researchers are depositing into either SBR or IR, figure 5 shows that physicists are more likely to submit a preprint than either economists or clinical medics ( $\chi^2=85.5$ ,  $p<0.001$ ). Conversely, in terms of the published final version, clinical medics were more likely to submit this version than either physicists or economics ( $\chi^2=9.9$ ,  $p<0.01$ ).

Figure 5: Which version(s) of these peer-reviewed journal article(s) did you deposit? (Phase 1 survey)



The reported submission behaviour of clinical medics is in contrast to physicists and economists, which may be a result of the different characteristics of the main SBR for each discipline in that PMC has a formal agreement with publishers as part of the submission process, whereas RePEc and arXiv are scholar controlled. On another level, however, variation across the three disciplines may reflect underpinning cultural differences in aspects of scholarly communication such as peer review and citation. Whitley (2000) argues that reputational control within 'professional adhocracies' (clinical medicine) is relatively diffuse, which is related to a diversity of problems and approaches. It might be surmised that, in the context of such cultural characteristics, a greater emphasis is placed on peer review as a 'stamp of authority' than in 'partitioned bureaucracies' or 'conceptually integrated bureaucracies' where the presence of a reputational elite means that 'time stamping' ideas is of more importance. The competitive nature of 'conceptually integrated bureaucracies' (physics), and to some extent the analytical theoretical core of 'partitioned bureaucracies' (economics), means that physicists and theoretical economists compete not only for funding and resources, but also for recognition. Anecdotal qualitative evidence

from our study indicates that for physicists, preprint archives play an important role in 'time stamping' ideas without incurring the penalty of time lags due to peer review:

“we all want to put [the paper] on arXiv as soon as possible, and in my view, we do it only after we get the first round of referees’ reports. In some cases, if that takes too long, if you send it to a journal and the journal takes, say, two months to let you know what happens, then you may not want to wait two months [...] this is a self-confidence level [...] you have to have a perception [that] what you have done is so interesting that people can’t wait to read it.” (Phase 1, Physics and Mathematics focus group).

It should also be highlighted that the high-degree of functional mutual dependence that characterises 'conceptually integrated bureaucracies' means that by the time a preprint has been submitted to an OAR it is likely to have already undergone a process of 'informal' internal peer review.

Looking at the citation preferences of clinical medics, physicists and economists (Figure 6) there are some statistically significant differences. Researchers in clinical medicine are more likely to agree that they will generally cite an author’s version of a journal article only if it has been peer-reviewed and accepted for publication ( $\chi^2=60.5$ ,  $p<0.001$ ) and less likely to be happy to cite a preprint ( $\chi^2=22.6$ ,  $p<0.005$ ). Dissemination and citation practices are closely intertwined. For example, economists are clearly engaged with dissemination of their work by means of working papers (Figure 7), with two-thirds having done so in the last five years, significantly more than in clinical medicine or physics ( $\chi^2=159.5$ ,  $p<0.001$ ). This corresponds with economists being more likely to agree that they were happy to cite preprints, or any version of a paper, in their own journal articles than either physicists or clinical medics.



Figure 6: In terms of your citation preferences for journal articles, please rate the following statements on a scale of 1-5 (1 = strongly agree and 5 = strongly disagree) (Phase 1 survey)

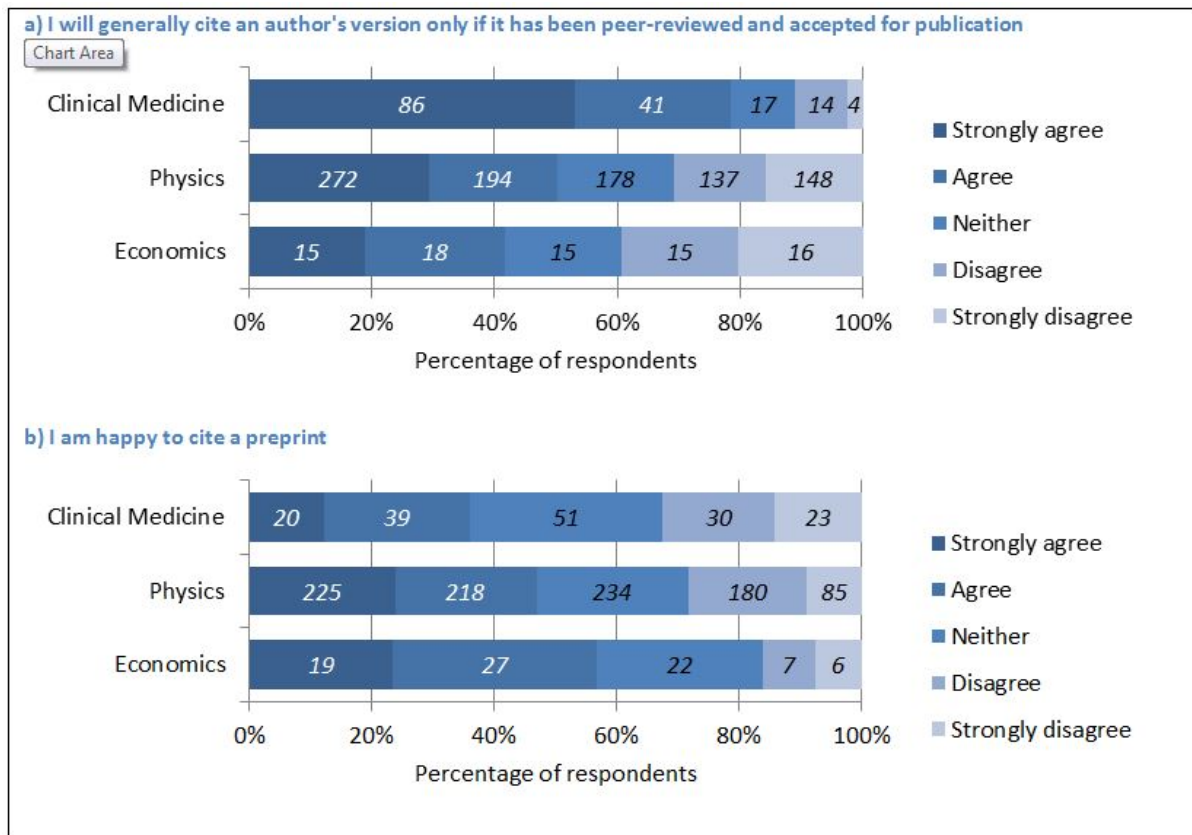
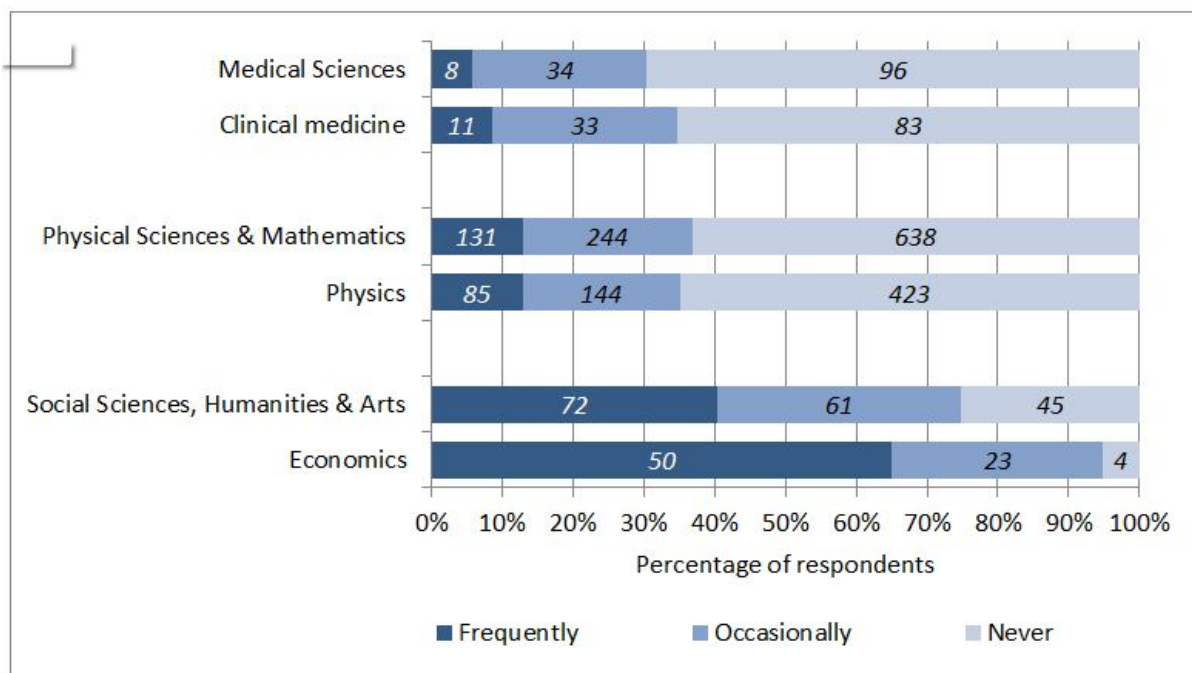


Figure 7: How have you published/disseminated your research in the last five years: working papers? (Phase 1 survey)



In terms of willingness to cite versions other than the final published version in journal articles, the clinical medics were more reluctant to cite an author's final peer-reviewed version or preprint than either physicists or economists.

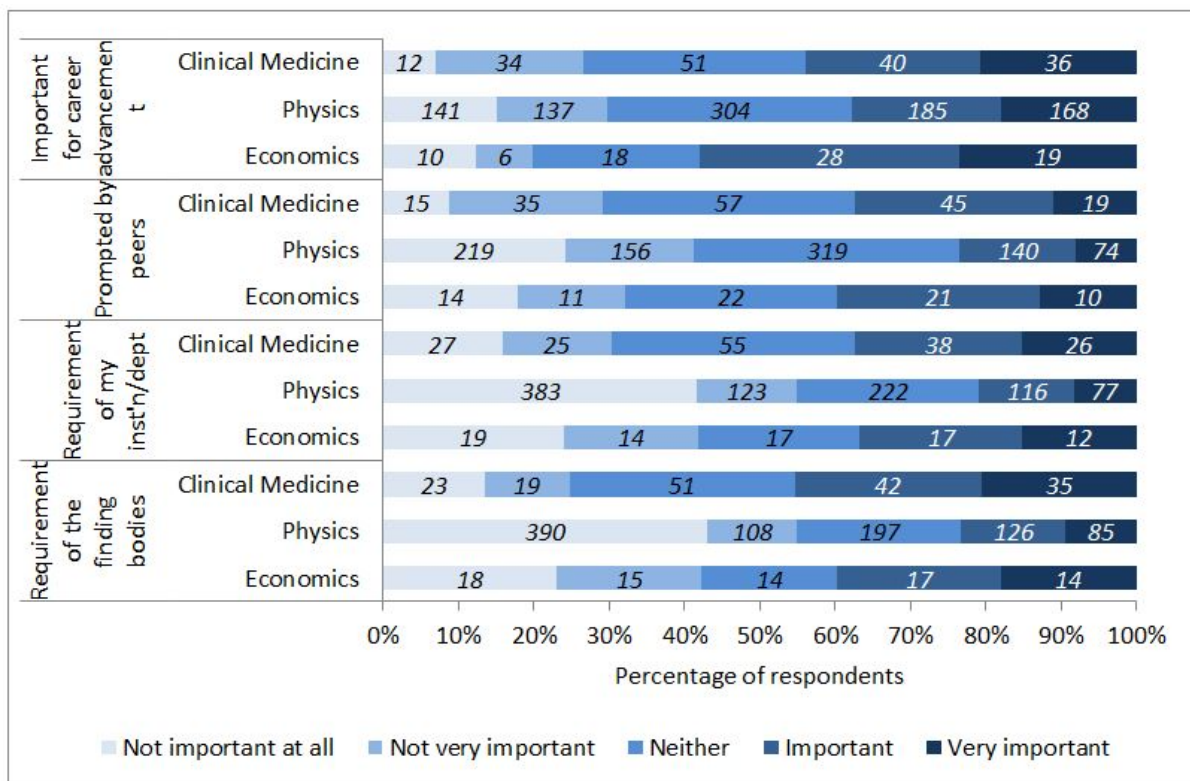
Of note, is that whilst economists indicated that they would be willing to cite preprints (or working papers) it has been suggested in the literature that actual citation practices in economic scholarly articles show otherwise; indeed, the use of working papers in writing journal articles is not always reflected in the final published version, as references to working papers are likely to be removed (or replaced by the published version, if available) in the process of turning a working paper into a scholarly article (Frandsen, 2009). This means that economists read working papers/preprints found in OAR when writing an article but this is not always evident in the references provided in the final published article.

Fry et al (2009) indicate that physicists may behave in a similar way to economists in terms of citation preferences in final published articles, with some physicists reporting that *"sometimes you may cite an article in the arXiv with the arXiv number of that article"* (pg. 64), whilst others feel that citation preferences are dependent on the target audience of the specific journal title *"If I have to quote, I would prefer to quote the published version, but in some cases I also prefer to put the arXiv [version], thinking of people who may have a difficulty to reach the published version, if it's in a journal they have no access to"* (pg. 65).

#### *Factors influencing self-archiving*

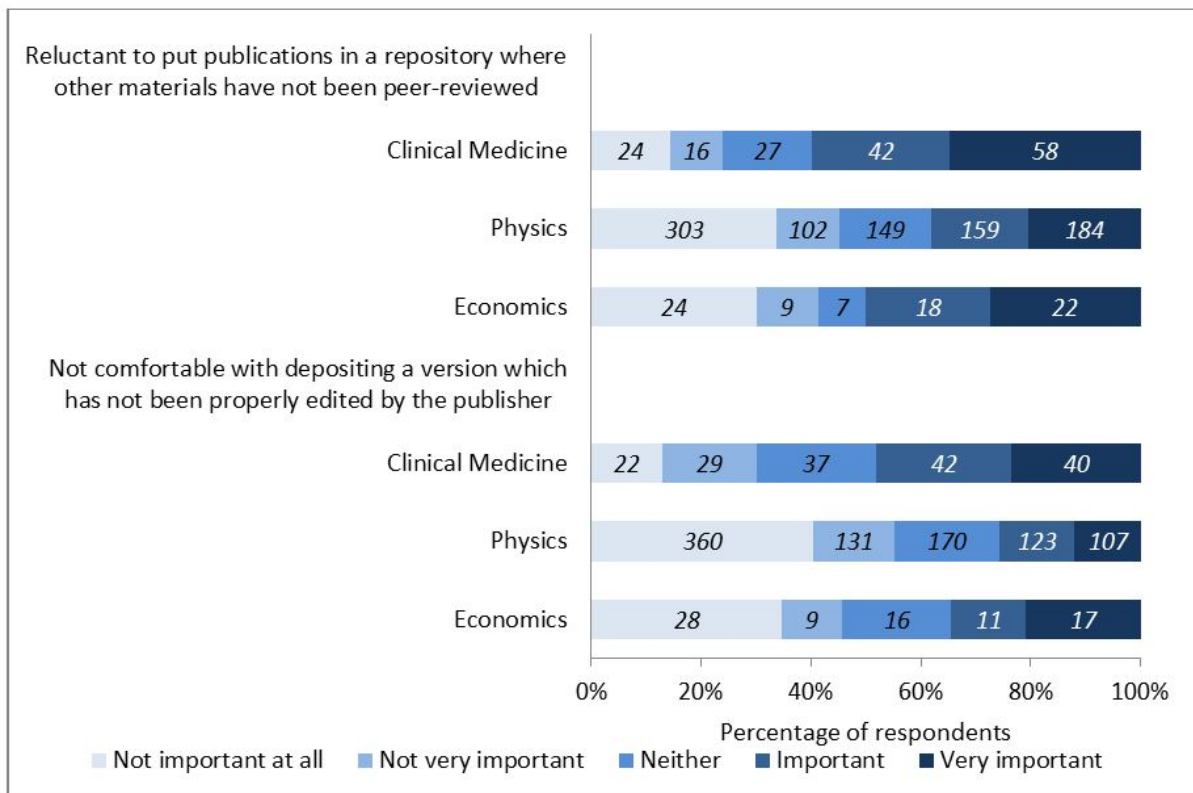
There are a variety of possible factors that might encourage researchers to submit versions of their articles in an OAR, some of which were examined in the survey (Figure 8). Career advancement was a less important factor for physicists than economists, with over half of economists rating it as 'important' or 'very important' ( $\chi^2=25.3$ ,  $p<0.005$ ). Being prompted by peers to deposit in an OAR was a more important factor for clinical medics and economists than physicists ( $\chi^2=35.1$ ,  $p<0.001$ ). Career advancement and funding body requirements were important factors for clinical medics. Respondents in physics seem less influenced by any OA requirements put in place by either their institution, department ( $\chi^2=56.4$ ,  $p<0.001$ ) or funding body ( $\chi^2=76.9$ ,  $p<0.001$ ), and were more likely to rate them as 'not at all important' than respondents in clinical medicine or economics.

Figure 8: How important are the following factors, or would the following factors be, in encouraging you to place peer-reviewed journal articles in a publicly available repository? (Please rate on a scale 1-5, 1 = very important and 5 = not important at all) (Phase 1 survey)



Some survey respondents indicated they had some reservations about placing versions of their articles in an OAR (Figure 9). A reluctance to place articles in an OAR where other materials have not been peer-reviewed was less of an issue for respondents in physics than those in clinical medicine ( $\chi^2=38.2$ ,  $p<0.001$ ). Similarly, not being comfortable depositing a version of their paper which has not been properly edited by the publisher was also less of an issue for respondents in physics than those in clinical medicine or economics ( $\chi^2=59.0$ ,  $p<0.001$ ). Other factors (principle of free access to all; widespread availability of research; availability to researchers with limited access to subscribed journals; journal subscription costs charged by publishers; speed of dissemination; and possibility of increased citations to the output) were included in the survey but did not show statistically significant differences in response patterns between the disciplines<sup>viii</sup>.

Figure 9: What reservations do you have about placing your peer-reviewed journal articles in publicly available repositories? (Please rate on a scale 1-5, 1 = very important and 5 = not important at all) (Phase 1 survey)



## Discussion

It is interesting to note that a common characteristic of Whitley's (2000) 'professional adhococracy' (*clinical medicine*), 'partitioned bureaucracy' (*economics*) and 'conceptually integrated bureaucracy' (*physics*) is the use of formal channels of communication for the control and coordination of research. Whitley (2000) argues that the intellectual elite in economics dominate reputations by controlling prestigious communication channels. It is likely that the establishment of SBR is often motivated by the desire to address perceived weaknesses, e.g. publication lag, in a discipline's formal communication system. In physics, for example, prestigious journal titles such as *Physical Review Letters* have long been established for the rapid dissemination of significant fundamental results, in the form of short reports. In 'conceptually integrated bureaucracies' the policing of intellectual boundaries and the notion of a 'pecking order' between theoretical and applied areas are similar to that of 'partitioned bureaucracies' (Whitley, 2000). In clinical medicine, there is a necessity to coordinate a diverse and frequently changing range of research problems. Such coordination is achieved, in part, by control over research methods and procedures, but also through a highly-controlled communication system that enables research outcomes to be compared, evaluated and coordinated.

As Becher and Trowler (2001) argue, communities that inhabit 'urban' landscapes rely on rapid and heavily used information networks; whereas, in 'rural' landscapes there is a reduced need for rapid and informal communication. This is corroborated by empirical studies investigating dissemination and publication behaviours within disciplines, such as Harley et al. (2010). The use of OA material, such as working papers by economists and preprints by physicists, to inform their research endeavour is well recognised and accepted as a disciplinary practice. Becher and Trowler (2001) also highlight that preprints are a characteristic of urban communities, since preprints establish the priority of an author's findings without the delay incurred by publication lags. It seems likely that the more hectic the pace of 'urban' research the greater the reliance on informal channels of communication: the adoption of SBR to disseminate research, with their mix of different types of output including published peer-reviewed journal articles and preprints, seems to support this assertion.

Both physicists and economists were more likely to be aware of an appropriate SBR than researchers at the metadiscipline level. It is interesting to note the low levels of awareness of SBR at the metadiscipline level given the worldwide rise of funders' and (cross-)institutional mandates encouraging (and often requiring) authors to deposit a version of their published articles in an OAR (Registry of Open Access Repository Mandates and Policies, n.d.).

Furthermore, comparison of the discipline with its parent metadiscipline revealed the extent to which behaviours described at the finer-grained level of the discipline could be considered to be 'atypical' or 'typical' of a coarser-grained level of aggregation. In this regard, researchers from the Social sciences, humanities & arts appeared to be the most heterogeneous in terms of their behaviours, although this might be expected given that it is a category that encompasses such wide-ranging intellectual territory. Furthermore, as a social science discipline, economics is unique in terms of its relatively high degree of technical task uncertainty combined with a high-degree of theoretical uniformity (dependent upon a high degree of strategic dependence), which is typical of Whitley's (2000) description of a 'partitioned bureaucracy'. Consequently, the research culture within economics, and resulting communication system, is likely to be quite different to that of other social science disciplines. Heterogeneity was also reflected in the comparison between physics and the Physical sciences & mathematics metadiscipline, which is perhaps less expected given the narrower spectrum of intellectual territory encompassed by this metadiscipline compared with, for example, the Social sciences, humanities & arts. By far the most homogenous metadiscipline was the Medical sciences with only limited variation in behaviours reported between the clinical medics and researchers at the aggregate level. An explanation for this may be that many of the disciplines that constitute the Medical sciences have a research culture that could be defined as a 'professional adhocracy'. Certainly, Whitley (2000) argues that the standardization of research procedures in 'professional adhocracies' implies a stronger 'organisational consciousness' and identity

with a corresponding emphasis on abiding by collective rules. Researchers in ‘professional adhocracies’ are more likely to be mobile across intellectual boundaries and research is more likely to be interdisciplinary in character than in ‘partitioned bureaucracies’ or ‘conceptually integrated bureaucracies’, so it is likely that there is a shared understanding of and ‘convergent’ behaviour relating to these collective rules across boundaries.

## Conclusions

Physicists, economists and clinical medics share particular cultural characteristics that may explain why these three disciplinary communities are ‘OA friendly’. These characteristics relate to Whitley’s notion of a reputational elite, the presence of which creates a high degree of competition for tightly controlled journal space. The existence of a reputational elite is particularly prevalent in ‘conceptually integrated bureaucracies’ (physics) and ‘partitioned bureaucracies’ (economics), which perhaps explains the development of pre-publication channels of dissemination that allow for ‘time stamping’ ideas or results prior to formal publication. Hence, the development of the scholarly-controlled SBRs arXiv and RePEc to support the long-standing preprint and working papers culture in physics and economics. Clinical medics were less consciously aware of participating in green OA, perhaps because in clinical medicine ‘self-archiving’ is predominantly achieved by publishers entering into an agreement for articles to be automatically deposited into PMC on behalf of authors. Arguably, PMC is not a scholarly-controlled innovation in the same way as arXiv and RePEc, emerging as it did in a specific OA policy context, but in focusing on published journal articles PMC reflects the need for the ‘stamp of authority’ characteristic of a ‘professional adhocracy’, whereby the lack of a reputational elite puts emphasis on the peer review process (Creaser et al., 2010, p.156) not just as a quality control mechanism, but also as a way to validate the significance of the research.

A caveat to these conclusions is that physics, economics and clinical medicine are discussed in general terms as if these disciplines are constituted by homogeneous disciplinary communities, whereas each of these three disciplines will have quite distinct sub-cultures, some that have more readily adopted OAR than others; this will impact on the generalizability of the conclusions.

Comparison between the two different levels of analysis, discipline and metadiscipline, has provided some interesting insights regarding methodological approaches when it comes to studying research cultures and the behaviour of those researchers that constitute them. As noted by Case (2012), whilst there are a number of rich in-depth studies of information behaviours within individual disciplines, large-scale comparisons at this level tend to be impractical. In practice, therefore, most comparative studies rely on aggregate results at the level of metadisciplines leading to something akin to a ‘bird’s-eye view’ of behaviours. At the level of individual disciplines, comparisons tend to lead to a crisper depiction than those at the aggregate level. There are pay-offs to both approaches, but the key point is that whilst it might be tempting to conflate disciplinary depictions to the parent metadiscipline,

our findings show that such conflation would be misleading outside of those knowledge domains where research communities have been empirically shown to be homogeneous in their behaviours. Conversely, the appeal of the 'bird's-eye view' is the scope for generalizability, but again if the metadiscipline in question is heterogeneous in character then empirical findings may be only marginally representative of the disciplines that constitute the aggregation.

The findings presented in this paper are important in that they contribute to the development of a greater understanding of disciplinary research cultures in relation to OAR and OA policy initiatives. The findings also inform methodological approaches for studying the information, dissemination and publication behaviours of researchers.

### **Acknowledgements**

The authors would like to acknowledge funding from PEER for the PEER Behavioural Research conducted between 2009-2011, which enabled the collection of the data used in this article. Our thanks to Sonya White who worked on the initial analysis of the data and Charles Oppenheim who was involved in Phase 1 of the project. We would also like to thank the anonymous referees for their recommended improvements to the paper.

### **References**

- Abreu, M., Grinevich, V., Hughes, A. & Kitson, M. (2009). *Knowledge exchange between academics and the business, public and third sectors*. Special report. Retrieved from [http://www.jbs.cam.ac.uk/fileadmin/user\\_upload/centre-for-business-research/downloads/special-reports/specialreport-knowledgeexchangeacademics.pdf](http://www.jbs.cam.ac.uk/fileadmin/user_upload/centre-for-business-research/downloads/special-reports/specialreport-knowledgeexchangeacademics.pdf)
- Bates, M. (2002). Speculations on browsing, directed searching, and linking in relation to the Bradford distribution. In H. Bruce, R. Fidel, P. Ingwersen, & P. Vakkari (Eds.), *Emerging frameworks and methods: Proceedings of the 4th International Conference on Conceptions of Library and Information Science (CoLIS4)* (pp. 137–149). Greenwood Village, CO: Libraries Unlimited.
- Becher, T. (1989). *Academic tribes and territories: Intellectual enquiry and the cultures of disciplines*. Milton Keynes: SRHE/ Open University Press.
- Becher, T. & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines*. Buckingham: SRHE & Open University Press.
- Bernius, S., Hanauske, M., Dugall, B., & Konig, W. (2013). Exploring the effects of a transition to Open Access: Insights from a simulation study. *Journal of the Association for Information Science and Technology*, 64(4), 701-726.
- Björk, B. C, Laakso, M., Welling, P., & Paetau, P. (2014). Anatomy of green Open Access. *Journal of the Association for Information Science and Technology*, 65(2), 237-250.

- Björk, B. C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Gudnason, G. (2010). Open Access to the scientific journal literature: Situation 2009. *PLoS ONE*, 5(6). doi:10.1371/journal.pone.0011273
- Case, Donald O. (2012), *Looking for information: A survey of research on information seeking, needs, and behaviour* (3<sup>rd</sup> ed.). Bingley: Emerald.
- Crane, D. (1972). *Invisible colleges: Diffusion of knowledge in scientific communities*. London: The University of Chicago Press.
- Cullen, R., & Chawner, B. (2011). Institutional repositories, Open Access, and scholarly communication: A study of conflicting paradigms. *Journal of Academic Librarianship*, 37(6), 460–470. doi:10.1016/j.acalib.2011.07.002
- Creaser, C., Fry, J., Greenwood, H., Oppenheim, C., Proberts, C., Spezi, V. & White, S. (2010). Authors' awareness and attitudes toward open access repositories. *New Review of Academic Librarianship*, 16 (sup1), 45-161. doi:10.1080/13614533.2010.518851
- Dallmeier-Tiessen, S., Darby, R., Goerner, B., Hyppoelae, J., Igo-Kemenes, P., Kahn, D., ... van der Stelt, W. (2010). First results of the SOAP project. Open access publishing in 2010. Digital Libraries. Retrieved from <http://arxiv.org/abs/1010.0506>
- Duranceau, E. F. (2008). The “wealth of networks” and institutional repositories: MIT, DSpace, and the Future of the Scholarly Commons. *Library Trends*, 57(2), 244–261. doi:10.1353/lib.0.0030
- Ductor, L. (2015). Does co-authorship lead to higher academic productivity? *Oxford Bulletin of Economics and Statistics*, 77(3), 0305-9049. Doi:10.1111/obes.12070.
- Ellison, G. (2002). The slowdown of the economics publishing process. *Journal of Political Economy*, 110(5), 947–993. doi:10.1086/341868
- Frandsen, T. F. (2009). The effects of open access on un-published documents: A case study of economics working papers. *Journal of Informetrics*, 3, 124–133. doi:10.1016/j.joi.2008.12.002
- Fry, J. (2006) Scholarly research and information practices: A domain analytic approach. *Information Processing and Management*, 42, 299-316. doi:10.1016/j.ipm.2004.09.004
- Fry, J., Oppenheim, C., Proberts, S., Creaser, C., Greenwood, H., Spezi, V., & White, S. (2009). *PEER behavioural research: Authors and users vis-à-vis journals and repositories*. Baseline report for *Publishing and the Ecology of European Research (PEER)*. Retrieved from <http://www.peerproject.eu/reports/>
- Fry, J., Proberts, S., Creaser, C., Greenwood, H., Spezi, V., & White, S. (2011). *PEER behavioural research: Authors and users vis-à-vis journals and repositories - Final Report*. Retrieved from <http://www.peerproject.eu/reports/>



- Galin, J. R., & Latchaw, J. (2010). From incentive to stewardship: the shifting discourse of academic publishing. *Computers and Composition*, 27(3), 211–224. doi:10.1016/j.compcom.2010.06.010
- Garvey, W.D., Lin, N., & Nelson, C.E. (1979). Communication in the physical and social sciences. In W.D. Garvey (Ed.), *Communication, the essence of science: facilitating information exchange among librarians, scientists, engineers, and students* (pp. 280–299). Oxford: Pergamon Press.
- Gazni, A., Sugimoto, C.R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the American Society for Information Science and Technology*, 63(2), 323–335. doi: 10.1002/asi.21688
- Gordon, M. D. (1980). A critical reassessment of inferred relations between multiple authorship, scientific collaboration, the production of papers and their acceptance for publication. *Scientometrics*, 2, 193–201. doi:10.1007/BF02016697
- Harley, D., Accord, S.K., Earl-Novell, S., Shannon, L. and King, C.J. (2010), Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines, Center for Studies in Higher Education, UC Berkeley, Berkeley, CA. Retrieved from <http://cshe.berkeley.edu/research/scholarlycommunication/>
- Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science*, 46, 400–425. doi: 10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y
- Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science and Technology*, 51(14), 1306–1320. doi:10.1002/1097-4571(2000)9999:9999<::AID-ASII047>3.0.CO;2-T
- Lynch, C.A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *Libraries and the Academy*, 3(2), 327–336.
- Mote, L.J.B. (1962). Reasons for the variation of information needs of scientists. *Journal of Documentation*, 18, 169–175.
- Nowell, C., & Grijalva, T. (2011). Trends in co-authorship in economics since 1985. *Applied Economics*, 43(28). doi: 10.1080/00036846.2010.491458
- Piette, M.J., & Ross, K.L. (1992). An analysis of the determinants of co-authorship in economics. *The Journal of Economic Education*, 23(3), 277–283. doi: 10.1080/00220485.1992.10844762.
- Pinfield, S., Salter, J., Bath, P.A, Hubbard, B., Millington, P., Anders, J.H.S., & Hussain, A. (2014) Open-access repositories worldwide, 2005-2012: Past growth, current characteristics, and future possibilities.

- Presser, S. (1980). Collaboration and the quality of research. *Social Studies of Science*, 10(1), 95–101. Retrieved from <http://sss.sagepub.com/content/10/1/95.abstract>
- Registry of Open Access Repositories (ROAR) (2014). Number of research institutional or departmental, research cross-institutional, and other types of repositories between 01/01/1991-21/10/2014]. Unpublished raw data.
- Registry of Open Access Repository Mandates and Policies (ROARMAP) (n.d.). <http://roarmap.eprints.org/> [Website]
- Schymura, M., & Löschel, A. (2014). Incidence and extent of co-authorship in environmental and resource economics: Evidence from the Journal of Environmental Economics and Management, *Scientometrics*, 99(3), 631-661.
- Spezi, V., Fry, J., Creaser, C., Proberts, S., & White, S. (2013). Researchers' green open access practice: a cross-disciplinary analysis. *Journal of Documentation*, 69(3), 334–359.
- Tenopir, C., King, D. W., & Bush, A. (2004). Medical faculty's use of print and electronic journals: Changes over time and in comparison with scientists. *Journal of the Medical Library Association : JMLA*, 92(2), 233–241.
- Tenopir, C., Maysa, R., & Wu, L. (2011). Journal article growth and reading patterns. *New Review of Information Networking*, 16(1), 4–22. doi:10.1080/13614576.2011.566796
- Thatcher, S. (2011). Copyediting's role in an open-access world. *Against the Grain*, 23(2), 30–34. Retrieved from <http://www.against-the-grain.com/2011/07/v23-2-copyeditings-role-in-an-open-access-world/>
- The Royal Society (2006). *Science communication: Survey of factors affecting science communication by scientists and engineers*. Report. Retrieved from [https://royalsociety.org/~media/Royal\\_Society\\_Content/policy/publications/2006/1111111395.pdf](https://royalsociety.org/~media/Royal_Society_Content/policy/publications/2006/1111111395.pdf)
- Van Noorden, R. (2012). Britain aims for broad open access. *Nature*, 486(7403), 302–303. doi:10.1038/486302a
- Wates, E., & Campbell, R. (2007). Author's version vs. publisher's version: An analysis of the copy-editing function. *Learned Publishing*, 20(2), 121–129. doi:10.1087/174148507X185090
- Whitley, R. (2<sup>nd</sup> Ed.). (2000). *The intellectual and social organization of the sciences*. Oxford: Clarendon Press.
- Xia, J. (2008). A comparison of subject and institutional repositories in self-archiving practices. *Journal of Academic Librarianship*, 34(6), 489–495. doi:10.1016/j.acalib.2008.09.016

## Appendix

### *Ingest procedures for subject-based repositories*

SBR often differ widely in terms of how material is ingested into the repository. For example, the arXiv<sup>ix</sup> repository which broadly operates across physics and mathematics disciplines typically allows authors to upload papers themselves. Within this repository there is a mix of refereed and unrefereed content, including preprints, as well as revised versions of journal articles resulting from referees' comments. In economics, Research Papers in Economics (RePEc)<sup>x</sup> operates slightly differently in that it is common for publishers and research institutions to provide metadata (which is often harvested automatically by RePEc) about their outputs, with the metadata linking to the IR or online journal of the participating institution. In contrast, PubMed Central (PMC)<sup>xi</sup> - the main SBR for the biomedical sciences - allows author deposit, but primarily uses a publisher deposit approach, whereby participating publishers submit metadata and full-text articles into the repository.

The process by which an author's work is made available through OAR is often referred to as 'self-archiving'. This term originates from an arXiv-type approach in which authors deposit their own articles directly into the archive, however, the term is now widely used to refer to the process by which material gets deposited into any OAR, regardless of whether it is deposited by an article's author, a repository administrator, or by a publisher, and independent of the motivation for doing so e.g. dissemination, preservation, or mandate compliance.

---

<sup>i</sup> A hybrid OAJ is a subscription-based journal that offers immediate OA if a fee is paid

<sup>ii</sup> According to the Ranking of Web Repositories as of January 2015 the following IR were ranked in the top 20 out of 1,983 indexed OAR: University of California e-Scholarship Repository (6), Universitat Autònoma de Barcelona depòsit digital de documents (11), The Digital Library of Theses and Dissertations of the University of São Paulo (12), Virginia Tech University Digital Library and Archives (15), Queensland University of Technology Institutional Repository (16), University of Southampton Institutional Repository (17), MIT Institutional Repository (18), University of North Texas Digital Library (19) and Lund University Publications (20), <http://repositories.webometrics.info/en/world>

<sup>iii</sup> The ingest procedures for subject-based repositories are explained in the appendix

<sup>iv</sup> STM publishers participating in the PEER project BMJ Publishing Group, Cambridge University Press, EDP Sciences, Elsevier, IOP Publishing, Nature Publishing Group, Oxford University Press, Portland Press, Sage Publications, Springer, Taylor & Francis Group and Wiley-Blackwell

<sup>v</sup> For more details of how the core journals were selected for the PEER programme please see D9.4 PEER Annual Report – Year 1. Retrieved from [http://www.peerproject.eu/fileadmin/media/reports/20090928\\_PEER\\_D9\\_4\\_annual\\_public\\_report\\_final.pdf](http://www.peerproject.eu/fileadmin/media/reports/20090928_PEER_D9_4_annual_public_report_final.pdf)

<sup>vi</sup> With regards to working papers and preprints, they will be often presented together in this paper. This is because they are similar in nature; they are both non-refereed mechanisms for disseminating findings and both may (or may not) be the precursor for peer-reviewed material. The proportion of physicists rating working papers as important / very important may be understated if respondents did not equate the term working papers with preprints

<sup>vii</sup> <http://publicaccess.nih.gov/>

<sup>viii</sup> For findings relating to authors behaviours in relation to accessing material in OA repositories please see supplementary material online

---

<sup>ix</sup> ArXiv is an online repository of electronic preprints of scientific papers in the fields of mathematics, physics, astronomy, computer science, quantitative biology, statistics, and quantitative finance, <http://arxiv.org/>

<sup>x</sup> Research Papers in Economics (RePEc) is a decentralized bibliographic database of working papers, journal articles, books, books chapters and software components, <http://repec.org/>

<sup>xi</sup> PubMed Central is an online free full-text archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health's National Library of Medicine, <http://www.ncbi.nlm.nih.gov/pmc/>