Key Perspectives

Consultants to the scholarly information industry

Open access self-archiving:
An author study

May 2005

Alma Swan and Sheridan Brown

Key Perspectives Limited
48 Old Coach Road, Playing Place, TRURO, Cornwall, TR3 6ET, UK
(Registered Office)
Tel. +44 (0)1392 879702
www.keyperspectives.co.uk www.keyperspectives.com

CONTENTS

Appendices:

Tables

Figures

EXECUTIVE SUMMARY

This, our second author study on open access, was carried out to determine the current state of play with respect to author self-archiving behaviour. The survey was carried out during the last quarter of 2004. There were 1296 respondents.

The survey also briefly explored author experiences and opinions on publishing in open access journals to follow up our previous study on this topic for JISC and the Open Society Institute. Many of the findings reported here match those of that previous study. For example, the main reasons for authors publishing their work in open access journals are the principle of free access for all and their perceptions that these journals reach larger audiences, publish more rapidly and are more prestigious that the toll-access (subscription-based) journals that they have traditionally published in. The principal reasons why authors have not published in open access journals are that they are unfamiliar with any in their field and that they cannot identify a suitable one in which to publish their work. These reasons, and their rank order, exactly match the findings from our survey that was specifically on open access publishing last year.

The purpose of this present study, however, was to move the focus onto self-archiving, the alternative means of providing open access to scholarly journal articles. Almost half (49%) of the respondent population have self-archived at least one article during the last three years in at least one of the three possible ways — by placing a copy of an article in an institutional (or departmental) repository, in a subject-based repository, or on a personal or institutional website. More people (27%) have so far opted for the last method — putting a copy on a website — than have used institutional (20%) or subject-based (12%) repositories, though the main growth in self-archiving activity over the last year has been in these latter two more structured, systematic methods for providing open access. Use of institutional repositories for this purpose has doubled and usage has increased by almost 60% for subject-based repositories.

Postprints (peer-reviewed articles) are deposited more frequently than preprints (articles prior to peer review) except in the longstanding self-archiving communities of physics and computer science. There are some differences between subject disciplines with respect to the level of self-archiving activity and the location of deposit (website, institutional or subject-based repositories). Self-archiving activity is greatest amongst the most prolific authors, that is, those who publish the largest number of papers.

There is still a substantial proportion of authors unaware of the possibility of providing open access to their work by self-archiving. Of the authors who have not yet self-archived any articles, 71% remain unaware of the option. With 49%

of the author population having self-archived in some way, this means that 36% of the total author population (71% of the remaining 51%), has not yet been appraised of this way of providing open access.

Authors have frequently expressed reluctance to self-archive because of the perceived time required and possible technical difficulties in carrying out this activity. The findings here show that 20% of authors found some degree of difficulty with the first act of depositing an article in a repository, but that this dropped to 9% for subsequent depositions. Similarly, 23% of authors took more than an hour to deposit their first article in a repository, but only 13% took this long subsequently, with most taking a few minutes. Another author worry regarding self-archiving is the danger of infringing agreed copyright agreements with publishers. Only 10% of authors currently know of the SHERPA/RoMEO list of publisher permissions policies with respect to self-archiving, where clear guidance as to what a publisher permits is provided. Where permission is understood by the author to be required, it seems it is being sought (this accounts for around 17% of self-archiving cases); where it is not known if permission is required, authors are not seeking it and are self-archiving without it.

Communicating their results to peers remains the primary reason for scholars publishing their work; in other words, they publish to have an impact on their field. Nonetheless, more than half still do not know what the citation rate is for their most recent articles. Almost all (98%) of authors use some form of bibliographic service to locate articles of interest in closed archives such as publisher websites, but only a much smaller proportion of people (up to 30%) are yet using the specialised OAI search engines to navigate the open access repositories. Nevertheless, at the time of this survey, 72% of authors were using Google to search the web for scholarly articles: the subsequent arrival of GoogleScholar, which indexes the content of open access repositories as well as general websites, and thus retrieves formally-archived open access material, can be expected have a bearing on the level to which open access archives are searched in future and consequently on the eventual impact of articles deposited therein.

The vast majority of authors (81%) would willingly comply with a mandate from their employer or research funder to deposit copies of their articles in an institutional or subject-based repository. A further 13% would comply reluctantly; 5% would not comply with such a mandate.

Alma Swan
Key Perspectives Ltd
Truro, UK
27 May 2005

1. INTRODUCTION

Twelve months ago we at Key Perspectives Ltd completed and reported on a study of authors who had published their work in open access journals, compared and contrasted with authors who had not done this[1,2]. The work was commissioned and funded by the Joint Information Systems Committee (JISC) in the UK and the Open Society Institute. Having thus learned about authors' experience of **open access publishing**, we embarked upon this current study of the alternative means to providing open access — by authors archiving copies of their articles in open access archives or repositories. This process is usually referred to as **'self-archiving'**.

The practice of self-archiving has its roots in the field of computer sciences, where researchers were depositing results in ftp archives some decades ago and, later, on websites. A preprint culture — that is, the distribution of drafts of research articles before they have been peer reviewed to colleagues around the world, to establish ownership of the piece of research, to move the subject along, and to invite critical commentary before final revision and submission of the articles to learned journals — had been in place for many years in print form in the computer science community, and as the digital age arrived the practice simply migrated from paper to electronic form. Today, there are more articles – preprint and postprint (peer-reviewed papers) - freely available through self-archiving in computer science than in any other subject. The computer science 'online library', Citeseer[3], currently has almost 723,000 articles that have been harvested from distributed sites around the world (websites, ftp archives) where authors have deposited their work. Not only does this indicate the size of the corpus of computer science research available on open access, but it clearly demonstrates the success of this mechanism (harvesting from distributed sites) for creating a subject-based open access archive.

There is another mechanism for creating a subject-based archive and that is for authors to deposit their work directly into a centralised repository. In 1991, the first centralised archive, for the high-energy physics community, was established at the Los Alamos National Laboratory. It is called arXiv[4] and today this houses some 300,000 documents, with around 42,000 being added each year. Its main areas of coverage are high energy physics, condensed matter physics and astrophysics: substantial numbers of articles in computer science and mathematics research reside there too, along with, latterly, quantitative biology. It was also, from the outset, the norm for postprints — the peer-reviewed version of each article — to be deposited in arXiv, too. In most cases these are in the form of the author's final version rather than the publisher's formatted file, though some publishers do permit the use of their own copyrighted version for

this purpose. The effect, then, was for research articles in the disciplines covered by arXiv to be available to anyone who wished to read them, even if their own institution could not afford to purchase the journals in which they were published. [On a point of terminology, the collective term for an electronic version of an article in draft (preprint) or final, peer-reviewed (postprint) form self-archived by the author is an 'e-print'].

That this practice could be spread to the rest of the scholarly community, freeing up the whole research literature from what he termed 'toll-access', that is, accessible only to those whose library could purchase the journals, was first mooted by Stevan Harnad in 1995[5,6]. Harnad has argued this case ever since, refining the model and rebutting[6] the arguments against the notion, which come not only from publishers, understandably nervous at what they see as a threat to their businesses, but also from the scholarly community itself — from researchers and librarians, both of whom are stakeholders in the developments in scholarly communications[7]. Their concerns have been debated extensively in public fora over the last decade (and continue to be), including the online American Scientist Open Access Forum set up and moderated by Harnad since 1998, the longest-running of all the open access discussion lists[8].

It is useful to lay out here the elements of this debate and the concerns that exercise the various parties. It should be noted that the focus of this present study is self-archiving, not open access publishing (in open access journals), which was extensively covered and discussed in our foregoing study[1,2]. The discussion here, therefore, concentrates on the issues around self-archiving that form the foci of resistance to the practice and which need to be overcome by proponents of open access if the whole research literature is to be 'made free'.

The first discussion point is the definition of what self-archiving is and what it is not. It is *not* an alternative to publishing in learned journals, but an adjunct, a complementary activity where an author publishes his or her article in whatever journal s/he chooses and then simply self-archives a copy. In practice, this means depositing the file, which is usually the author's final version of the article after peer review has been completed, in an open access archive or repository. There are two main types of such archives, which we will come to shortly. The articles are tagged in these archives as peer-reviewed postprints or as preprint drafts, so it is possible clearly to distinguish the two.

This brings us to the second point. Some researchers express a concern about the 'quality' of self-archived articles. Some disciplines use preprints much more extensively than others, but these pre-peer review articles are clearly tagged as such. It is true that some institutional archives may contain lots of other types of material as well (see Section 5.4.6 of this report) but the critical point here is that

with respect to the research literature, what is deposited as a postprint is a *copy* of a fully peer-reviewed article whose destiny was to be published in the traditional way in a conventional, quality-controlled journal. It has therefore been peer-reviewed in the usual way. Postprints are not some kind of self-published, second-rate alternative to conventional journal articles: they *are* those articles.

Authors have often cited the issue of copyright as a major stumbling block to self-archiving. They are anxious that, having signed over copyright to the publisher of the journal in which their article appears they will be contravening the agreement if they self-archive the article. To be sure, if they self-archive the publisher's own file (the PDF file supplied by the publisher to the author and containing the final formatting and layout assigned by the publisher) without permission, then this would in almost all cases be in contravention of copyright, if that resides with the publisher. The publisher has not copyrighted the author's final version, however, and in the vast majority of cases (over 90% is the latest estimate[9,10]) the publisher expressly permits an author to self-archive their own final draft — the version that was finally submitted to the publisher after peer-review revisions and recommendations have been incorporated.

The other main issue that is raised by authors[1,2] and, sometimes, by librarians, is how self-archiving might disrupt the present scholarly publishing model. Naturally, it is the perceived vulnerability of the journals published by learned societies, rather more than those of commercial publishers, that concerns authors. In this respect, it is worth examining what has happened to learned societies that have already had experience in this arena, those publishing in the areas covered by arXiv, alongside which they have had to live since 1991. It has already been said here that arXiv receives around 42,000 deposits per year. The ISI (Institute for Scientific Information) Science Citation Index covers around 420 physics journals, and to give a measure of the total volume of physics research, in 2003 these journals published a total of 116,721 articles: arXiv thus contains a substantial proportion (approximately one third) of the total physics research output and in the specialist areas mentioned earlier — condensed matter, astrophysics and high energy physics — the coverage of arXiv is pretty well complete.

In a separate exercise to this present study, we asked the American Physical Society (APS) and the Institute of Physics Publishing Ltd (IOPP) what their experiences have been over the 14 years that arXiv has been in existence. We asked how many subscriptions have been lost as a result of arXiv. Both societies said they could not identify any losses of subscriptions for this reason. Subscription movements for the journals they publish in the areas covered by arXiv are no different from those of their journals in other areas of physics over

the period. Moreover, both societies say that they do not view arXiv as a threat to their business (rather the opposite, in fact) and this is underlined by the fact that the APS helped establish an arXiv mirror site at the Brookhaven National Laboratory – hardly the action of a society with its back to the wall because of that repository. Now it is true that there are only a couple of experiments of this sort carried out so far (physics and computer science), where publishers have to co-exist with a successful open access archive, and so there is always the possibility that there is something of a 'special case' about this example. Quite what might make it such a special case has never been adequately argued, but it is a finite possibility. Nevertheless, the evidence there is to hand points to the likelihood that the peaceful — and perhaps mutually beneficial — co-existence of traditional journals and open access archives is entirely possible; in biological terms, mutualism, rather than parasitism or symbiosis, might best describe the relationship.

The final issue that is raised frequently is the cost to institutions that self-archiving might impose. This is much more in the area of responsibility of librarians and institutional administrators than of authors. Will setting up and running an open access archive in a research-based university, for example, cost a lot of money? How will it be paid for, whose budget will it fall under, can it be afforded, will it need an open cheque for the future? We collected together some actual figures from various archive managers for a study we undertook recently to develop a model for a national e-prints service for the United Kingdom. The figures varied wildly, as we meant them to for illustrative purposes, for we selected as our examples some of the largest and most ambitious, and some of the smallest and most modest, institutional archives in existence. For the whole range of costs, the reader is directed to the report of that study[11,12]. It is probably most helpful here to say that an average-sized research-based university can set up a functional archive for, say, ten thousand US dollars. Annual running costs vary according to the institution's existing levels of provision of IT services, what level of interventional support administrators are going to give the archive, and how much advocacy activity is to be included, but could amount to half or one FTE if ambitions do not run too high. For all the benefits such an archive brings to an institution (see below), this represents excellent value for money.

So much for the worries and concerns about self-archiving. Let's turn now to the arguments for it and the benefits that it can bring to the scholarly community, for there must be substantial benefits to be realised if the effort is to pay off. The benefits fall into two camps, those for the institution and those for the researchers (and some are shared, of course).

For the researcher, the most obvious benefit of making their work open access is the enhanced citations, and therefore impact, that result[13,14,15,16]. We know from

the work reported here and elsewhere[17,18] that authors publish primarily to communicate their research findings to their peers, so that they can be built upon in future research efforts. Depositing an article at the time of acceptance for publication also means that the inevitable delay at the publisher before the article finally appears in the journal is immaterial — the article is already available to anyone who wants to read it and use it for their work. The research cycle is thus shortened. And of course, the article is available to *all* interested parties, not just to readers in institutions that can afford the journal in which it is published.

There are other benefits, too. An institutional repository is a secure storage location for working documents or for research data; it becomes the mediator for a one-input, many-outputs scenario, where a researcher can retrieve whichever elements of his or her own research record are needed for a task-in-hand (perhaps writing a paper, a lecture, preparing teaching materials, preparing a CV). It can also provide the home for research data that cannot be published in traditional journal format but which supports research findings and which the author would like to make available to peers and colleagues, data such as very large datasets, video files, graphical files of various formats, audio files and mixed media output.

For the institution, the benefits are just as substantial. Research-based institutions share with the researcher the wish to enhance the visibility and impact of the research generated within that institution. Institutions also have administrative burdens that require access to, and organisation of, information about their employees' research records, research grant applications and fulfilment. They also need to carry out research performance evaluation (the Research Assessment Exercise in the UK being one such example,) and an institutional open access archive provides a permanent record of all the research output of that institution (provided that it has ensured all the researchers deposit copies of their articles, of course). An archive can also serve as a marketing tool for the institution, a shop window for potential students, staff and assessors on what is being generated by that institution. In a similar vein an institution can measure itself against other institutions that it sees as 'competitors' when all the outputs are openly visible in institutional archives. And, finally, a repository provides a place for *all* the digital output of that institution, so not just research articles but digital records of academic and cultural life in that institution can be stored there.

This gallop through the world of self-archiving brings us to the final discussion point here, which is the forms that self-archiving repositories might take. In this study we have distinguished the two main types, which are institutional and subject-based archives. Subject-based archives, such as arXiv discussed above, provide a location for the deposition of articles around a disciplinary theme. As

well as arXiv (which houses articles in physics, computer science and mathematics), there are other well-known examples, such as Cogprints[19] (cognitive sciences), also a centralised repository. RePEc[20] (economics) is similar but actually works by harvesting articles from distributed archives. Whilst there is the obvious attraction to the appropriate community of such subject-centred services, we have argued that the optimal system for encouraging and achieving self-archiving across the whole scholarly community is via a distributed system; in other words, a global network of institutional archives, all OAI-compliant and thus completely interoperable*, so that a user can locate and be directed to an original article wherever it resides and without having to know anything about its location[11,12]. Subject-based centralised archives have their devotees and can be extremely popular within their communities. They are few and far between, however, and apart from arXiv most have been filling extremely slowly; Cogprints, for example, despite its 8-year existence, still houses only around 2000 articles. Subject-based services can be very useful to researchers, but are probably most effectively created by service providers (search-and-retrieval services) that harvest relevant subject-focused information from *all* repositories and sort and organise it to form a subject-centred offering to the research community.

The reason for arguing for a distributed system is that it is institutions (employers) that can most effectively bring about an effective self-archiving practice across the board. To be sure, research funders can influence the researchers they fund. The Wellcome Foundation is just implementing a self-archiving mandate for its grantholders to self-archive their articles and is setting up a new European PubMed Central archive for this purpose[21]. But external research funds only benefit a fraction of the research carried out in universities, so research funders can only influence a fraction of researchers. The institutions themselves, however, can influence the whole body of scholars, in whatever disciplines they work, funded or not, and if all institutions provide an archive that is interoperable with every other archive then they are effectively contributing to a global database of freely accessible research — true open access.

*OAI-compliant means that the article metadata (the title, authors, keywords etc) are created in the format laid down by the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Search engines can then harvest the metadata from all archives making their metadata visible in this form, and present it to users in an appropriate way.

Acknowledgment

The Institute for Scientific Information in Philadelphia donated a mailing list of 25,000 names for this study. The Joint Information Systems Committee (JISC) funded the writing of this report and its publication. Both are gratefully acknowledged.

2. THE RESPONDENTS

The total respondent population is composed of four subpopulations as follows:

- Respondents to a call-to-respond posted on various open access and publishing-related online discussion lists (398 individuals; response rate unknown). This is termed the 'interested and informed' population.
- Respondents to an invitation sent to all individuals (851 people) whose email addresses could be collected by trawling open access repositories worldwide and scanning peer-reviewed, published articles (52 individuals, representing a 6% response rate). This is termed the 'archived' population.
- Respondents to an invitation sent out internally within the School of Electronics & Computer Sciences at Southampton University to 240 individuals (we are grateful for the help of Dr Leslie Carr with this task). This School made self-archiving mandatory in January 2003 so is an excellent test-bed for collecting the views of authors who have been required to make their work open access in this way (35 individuals, representing a 15% response rate). This is termed the 'Southampton' population.
- Respondents to an invitation sent to around 25,000 names randomly-selected from the Institute for Scientific Information's Science Citation Index and Arts & Humanities Citation Index (811 individuals, representing a 3% response rate). This is termed the 'randomly-selected' population. Southampton University researchers were specifically excluded by ISI when selecting names, as were the individuals invited as part of the 'archived' population, so as not to invite the same people twice. The proportions of invitations sent out in this campaign by ISI are shown in Table 1 below. All figures are percentages and are rounded:

Region/country	Arts & humanities	Social sciences	Sciences	Total
United Kingdom	1	4	9	**7**
Western Europe	11	11	15	**14**
Eastern Europe	7	10	12	**11**
North America	44	37	20	**27**
Japan	6	6	5	**6**
Asia (less Japan)	14	15	12	**13**
Central/South America	8	5	14	**11**
Africa	2	6	8	**6**
Australia/New Zealand	6	7	5	**5**
Total	**12**	**28**	**60**	**100**

Table 1: Breakdown of invitations sent out using c25,000 email addresses supplied by ISI

The total number of respondents was 1296.

74% of respondents work in universities, 13% in other non-commercial research institutions, 5% in the public sector and 5% in industry or business.

By geographical area the respondent pattern was as shown in Table 2. Figures are percentages and are rounded:

Region	Percentage of total respondents
Australia/New Zealand	7
Asia (except China and Japan)	4
China	3
Japan	1
Canada	4
USA	21
Central/South America	6
European union (except UK)	17
Other European countries (except EU or UK)	10
UK	18
Middle East	4
Africa	4

Table 2: Respondents by geographical area

By subject area the pattern of responses was as shown in Table 3. Figures are percentages and are rounded.

Subject	Percentage of total respondents
Agriculture & food science	5
Business & management	4
Chemistry	6
Computer sciences	12
Earth & geographical sciences	3
Engineering, materials science & technology	8
Humanities	8
Law & politics	1
Library & information science	6
Life sciences	17
Mathematics	6
Medical sciences	17
Physics	7
Psychology	9
Social sciences & education	10

Table 3: Respondents by subject area

Where appropriate and where there are significant differences between populations in the responses they gave, these differences are reported and

highlighted. Wherever only the whole population response is reported, this is because there are no significant differences between respondent subgroups.

3. OPEN ACCESS JOURNALS

Respondents were first asked whether they had submitted a manuscript to, or had a paper published in, an open access (OA) journal in the last three years. 66% had not done these things, 24% said they had and 9% don't know. It is important to understand the reasons why some people have chosen to publish their work on OA journals. The possible reasons why researchers might have elected to publish in open access journals were presented and respondents asked to indicate which of them were appropriate to their own case. The results are shown in Table 4. Reasons are presented in rank order. Figures are percentages of respondents and are rounded.

If you HAVE published in an Open Access journal, please indicate which of the following factors were reasons for doing so. Please select any that apply	
Reason	**% respondents**
The principle of free access for all readers	18
I perceive the readership to be larger than for a subscription-based journal	11
I perceive OA journals to have faster publication times than other types of journal	10
The OA journal(s) I have published in are prestigious in my field	9
I think my article will be more frequently cited	8
The OA journal(s) I have published in have a high impact in my field	6
I was attracted to the editor / editorial board	5
I am concerned about the cost to my institution of no OA journals	3
I object to publishing with a non-OA commercial publisher	3
My decision to publish in an OA journal was influenced by my co-publishing colleagues	2
The OA journal(s) I have published in are published from my own institution	1
My decision to publish in an OA journal was influenced by my institution	0.8
My decision to publish in an OA journal was influenced by my grant-awarding body	0.3

Table 4: Reasons for publishing in open access journals

Although there were some minor variations in these percentage scores across subject areas, the rank order of the reasons remains consistent. A complete analysis of responses broken down by subject area is shown in Appendix 1.

Some respondents offered additional comments which are reproduced in Appendix 2. A number of main themes arise from this list:

- The wish to support the principle of open access

- Several authors have been invited to submit their work to an OA journal
- Some authors value the increased visibility and other perceived advantages of publishing in an OA journal
- There is a fairly strong perception that OA journals publish work quickly

These responses – the rank ordering, the level of support for each reason, and the extra comments – all tally with the results obtained in a previous survey by KPL where the same questions were asked[1]. This previous survey was carried out on a much smaller sample population (around 150 individuals) and its findings have been called into question as a result of this, though we have asserted that the sampling methodology was carried out correctly and the findings were valid. This present larger-sample result almost exactly corroborates the previous findings which are that authors who choose to publish in an open access journal are primarily motivated by the open access principle, by their perceptions of large readerships and rapid publication times for such journals, and by the fact that the OA journals they have published in have a good reputation in their field.

The survey also explored the reasons why authors have *not* published in an open access journal (and this is the majority of authors). The results are shown in Table 5 in rank order. Figures are percentages of respondents and are rounded.

If you have NOT published in an Open Access journal, please indicate if any of the following factors were reasons	
Reason	**% respondents**
I am not familiar enough with OA journals in my field to feel confident about submitting my work	36
I could not identify any OA journals to publish in	22
I perceive the OA journals in my field to have low prestige	17
I perceive the OA journals in my field to have low impact	16
I cannot find the funds to pay any publication fees required by OA journals	14
I object in principle to paying a publication fee to OA journals that charge one	13
I perceive the readership to be smaller than for a subscription-based journal	11
I always publish my work in the same journals and am satisfied with this way of working	9
I am concerned about the archiving of work published in OA journals	7
I perceive the OA journals in my field to have poor peer review procedures in place	6
My decision was influenced by my co-publishing colleagues	4
My decision was influenced by my institution	4
My decision was influenced by my grant-awarding body	2
I perceive the OA journals in my field to have slower publication times than traditional journals	0.5

Table 5: Reasons for not publishing in open access journals

A complete analysis of responses broken down by subject area is found in Appendix 3. A large number of respondents appended comments, which are reproduced verbatim in Appendix 4.

Once again, the rank order and level of response for the options match those found in the previous study. The main reasons why authors have not published in an open access journal are that:

- they are not familiar with the concept or with OA journals in their field
- they could not identify a suitable one to publish in.

Interestingly, just as they did in the previous survey, authors who have not published in OA journals also say that they perceive them to have low prestige and impact, directly contradicting the perceptions of the authors who *have* published in an OA journal.

Forty nine percent of respondents say they are likely to publish at least one article in an open access journal in the next 3 years. 27% say this is *very likely* and 22% that it is *likely*. 15% are ambivalent (*neither likely nor unlikely*), 12% think it *somewhat unlikely*, 3% *will not do so* and 18% *don't know*.

4. THE USE OF RESEARCH INFORMATION

When gathering information to use in assessing how scholarly communication will develop, which is one of the main purposes of this study, it is helpful to understand how researchers produce and use scholarly information. This enables the assessment of how scholars in different disciplines work, how they produce and use information, their levels of awareness of advances in digital information services, and how much and how effectively they use them. Ten of the questions in the survey addressed these issues.

4.1 Ease of access to work-related information

With respect to research articles, more than half (54%) of respondents say they have easy access to *most* of the articles they need for their work. A further 10% say they have easy access to *all* the articles they need. Just over one quarter of respondents (26%) have easy access to *some* of the articles they need and 9% can only easily access a *few* of the articles they need. There are minor differences between subject areas here, the main findings being that computer scientists, engineers and mathematicians have better access to the research information they need than average, and humanities scholars have the most difficulty in accessing what they need, but in general there are not great differences between disciplines. The full breakdown is given in Appendix 5.

4.2 Age of articles most commonly used

There is, perhaps expectedly, more variation in the answers to the question which asked when most of the articles needed for respondents' work were published. The results for the whole population are shown in Table 6. Figures are percentages of respondents and are rounded.

For your own work, when were MOST of the articles you consult – or cite in your own articles – published?	
Period when most articles needed by respondents were published	**Percentage of total respondents**
Up to 2 years ago	8
Up to 5 years ago	35
Up to 10 years ago	34
Up to 20 years ago	17
Up to 50 years ago	4
More than 50 years ago	1

Table 6: How long ago most of the articles needed by respondents for their work were published

The full breakdown by subject area is in Table 7, and the data are presented in graphical form in Figure 1. The peak is in different places depending on subject area. For humanities the peak is around 20 years ago, for chemistry, engineering and medicine it is around 10 years ago and in computer science, life sciences and library & information science it is more recent at around 5 years ago.

For your own work, when were MOST of the articles you consult – or cite in your own articles – published?						
Subject	**Up to 2 years ago**	**Up to 5 years ago**	**Up to 10 years ago**	**Up to 20 years ago**	**Up to 50 years ago**	**More than 50 years ago**
Agriculture & food science	6	37	36	16	6	0
Business & management	11	49	28	11	2	0
Chemistry	7	28	42	21	2	0
Computer sciences	7	51	29	12	2	0
Earth & geographical sciences	15	33	33	12	6	0
Engineering, materials science & technology	4	30	37	23	5	1
Humanities	7	22	23	29	14	3
Law & politics	0	41	24	35	0	0
Library & information science	21	51	22	4	1	1
Life sciences	8	39	38	12	3	1
Mathematics	4	31	26	26	11	1
Medical sciences	8	37	39	15	1	1
Physics	5	35	31	22	6	0
Psychology	3	30	45	18	2	2
Social sciences & education	12	37	38	12	1	0

Table 7: Age of the most commonly used and cited articles by subject area of respondents

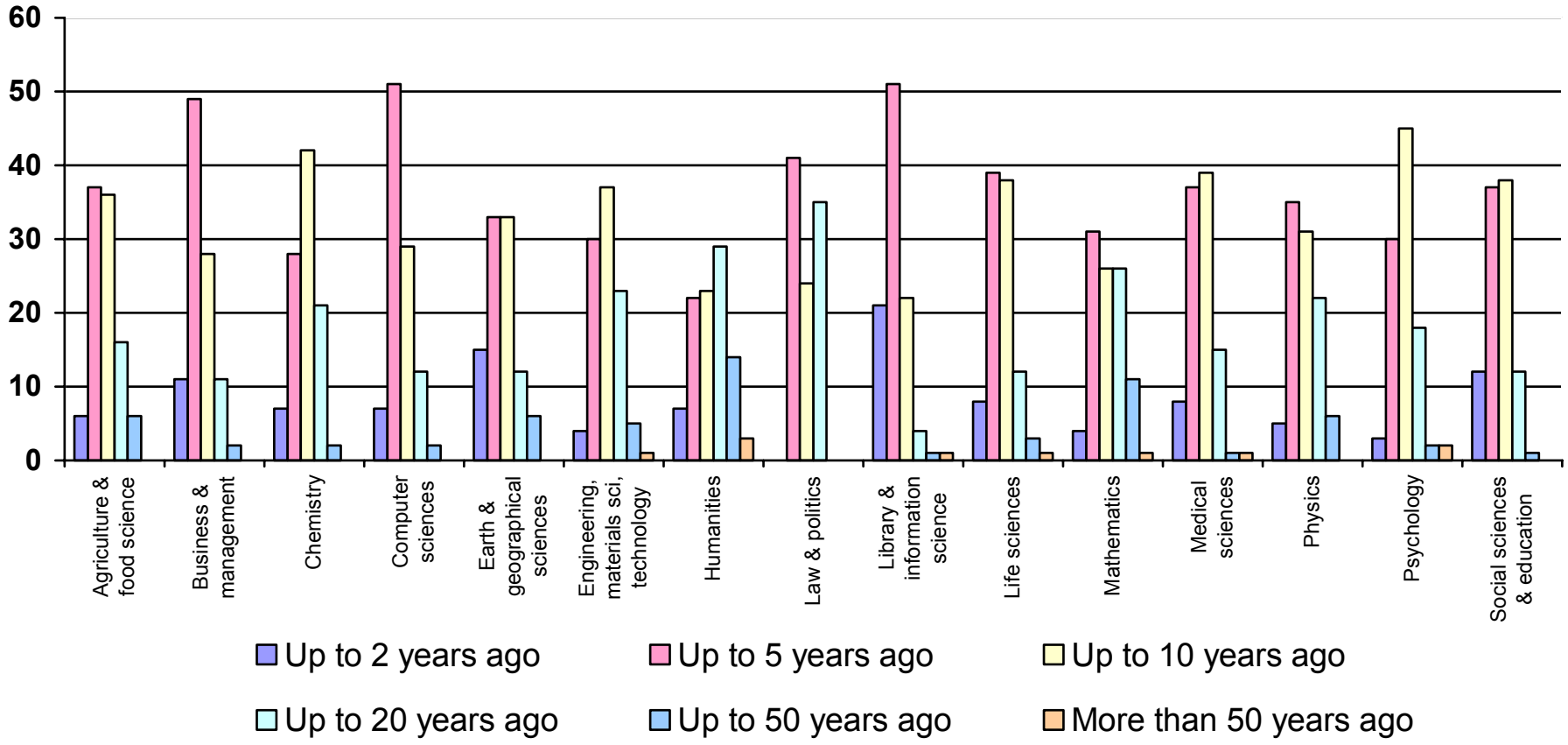

Figure 1: Age of the most commonly used and cited articles by subject area of respondents (clustered column chart)

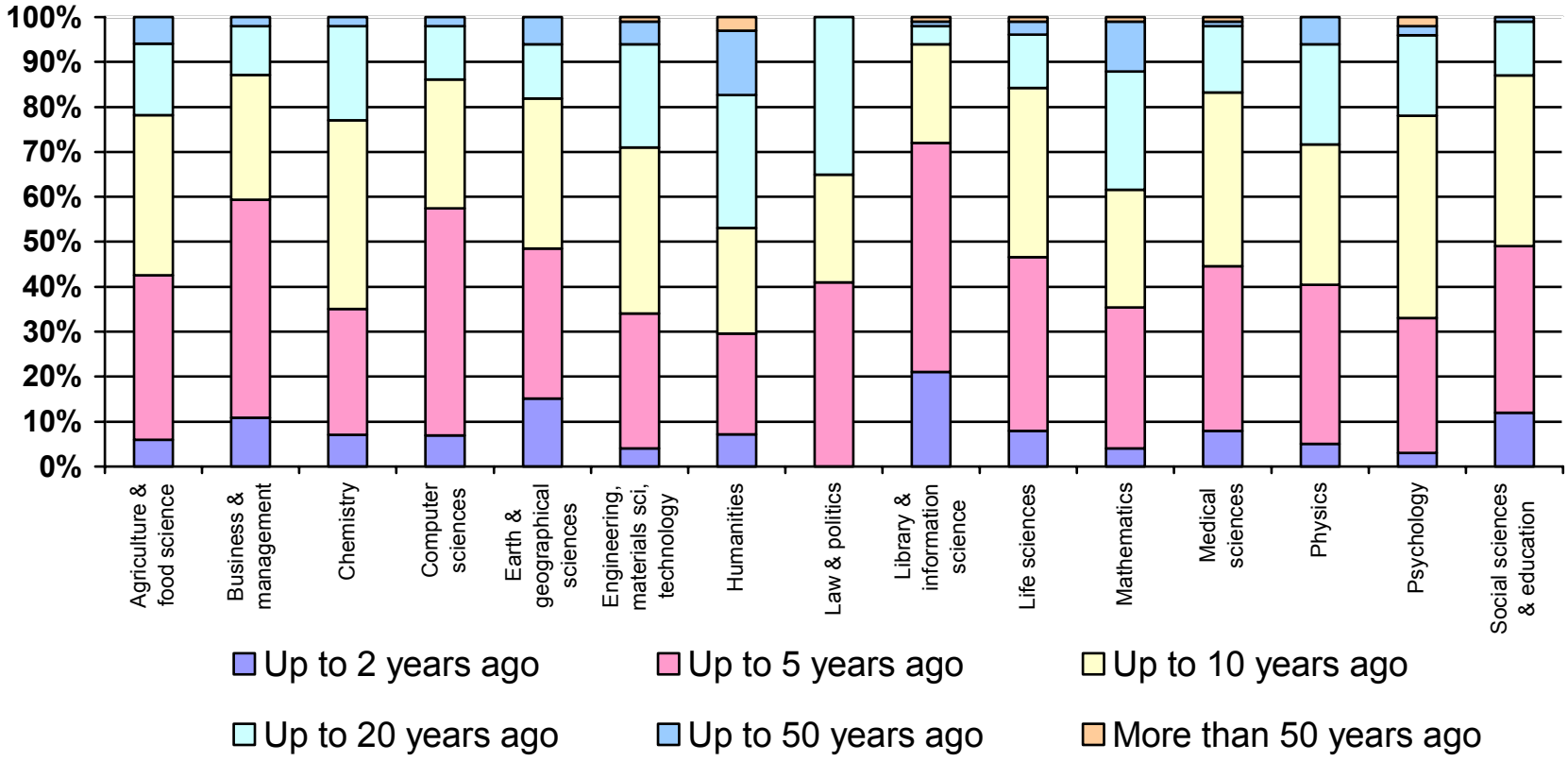

Figure 2: Age of the most commonly used and cited articles by subject area of respondents (stacked column chart)

4.3 Respondents' publishing activities

4.3.1 Number of articles published

Each year, 23% of respondents publish up to 1 article. Fifty percent publish 2-3 articles, 15% publish 4-5 and 11% publish more than 5 articles. The subject area where most papers are published is engineering, materials science & technology. Scholars in humanities and in library & information science publish the fewest. The full breakdown is given in Table 8 and Figures 3 and 4. Figures in the table are percentages of respondents and are rounded.

Approximately how many articles do you publish each year?				
Subject	**0-1**	**2-3**	**4-5**	**More than 5**
Agriculture & food science	19	57	16	7
Business & management	28	49	19	4
Chemistry	10	52	24	12
Computer sciences	27	47	11	15
Earth & geographical sciences	12	67	9	12
Engineering, materials science & technology	17	44	14	24
Humanities	35	43	12	8
Law & politics	24	59	6	12
Library & information science	44	44	3	8
Life sciences	18	52	20	11
Mathematics	28	44	18	10
Medical sciences	20	49	15	15
Physics	20	55	15	9
Psychology	24	49	17	10
Social sciences & education	25	55	15	5
Whole population	**23**	**50**	**15**	**11**

Table 8: Number of articles published each year, by subject area

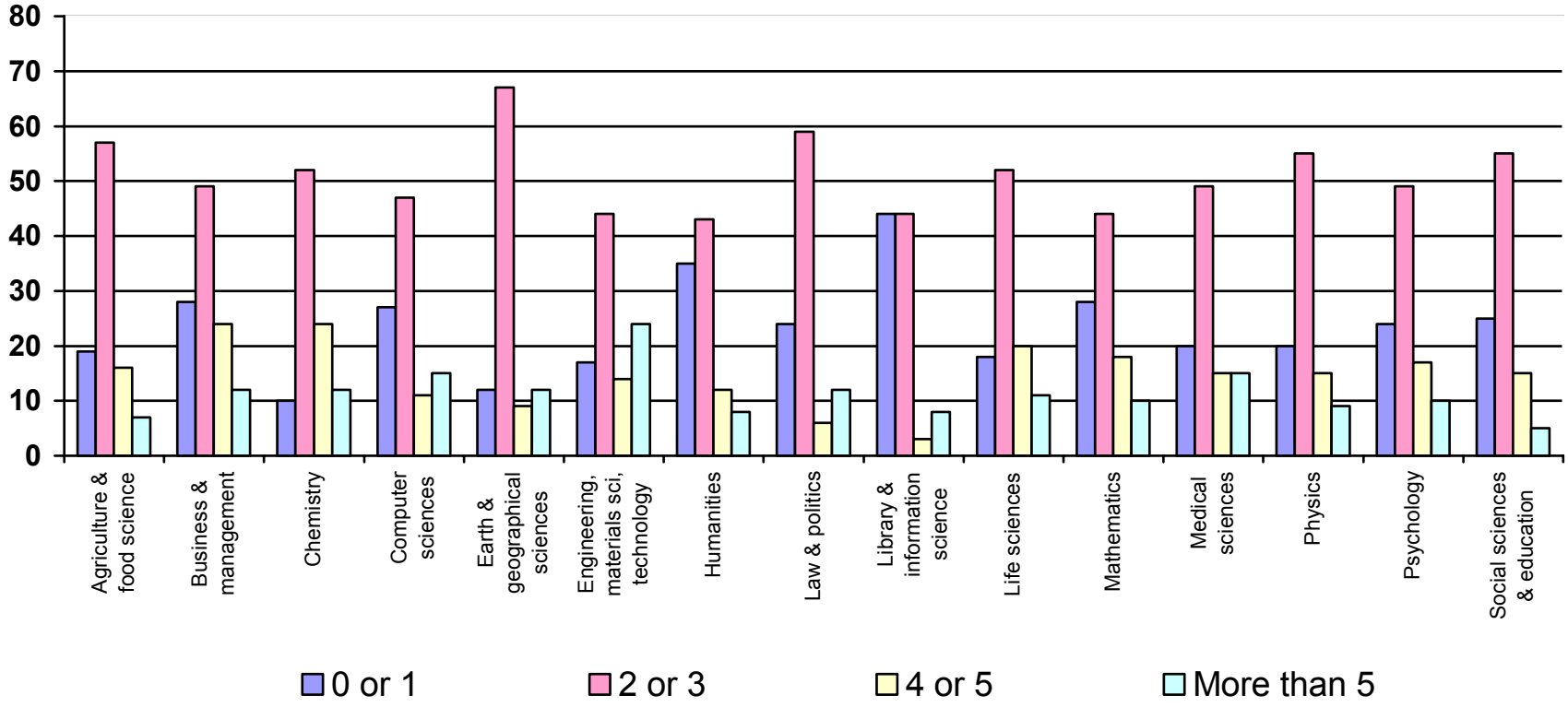

Figure 3: Number of articles published each year by subject area (clustered column chart)

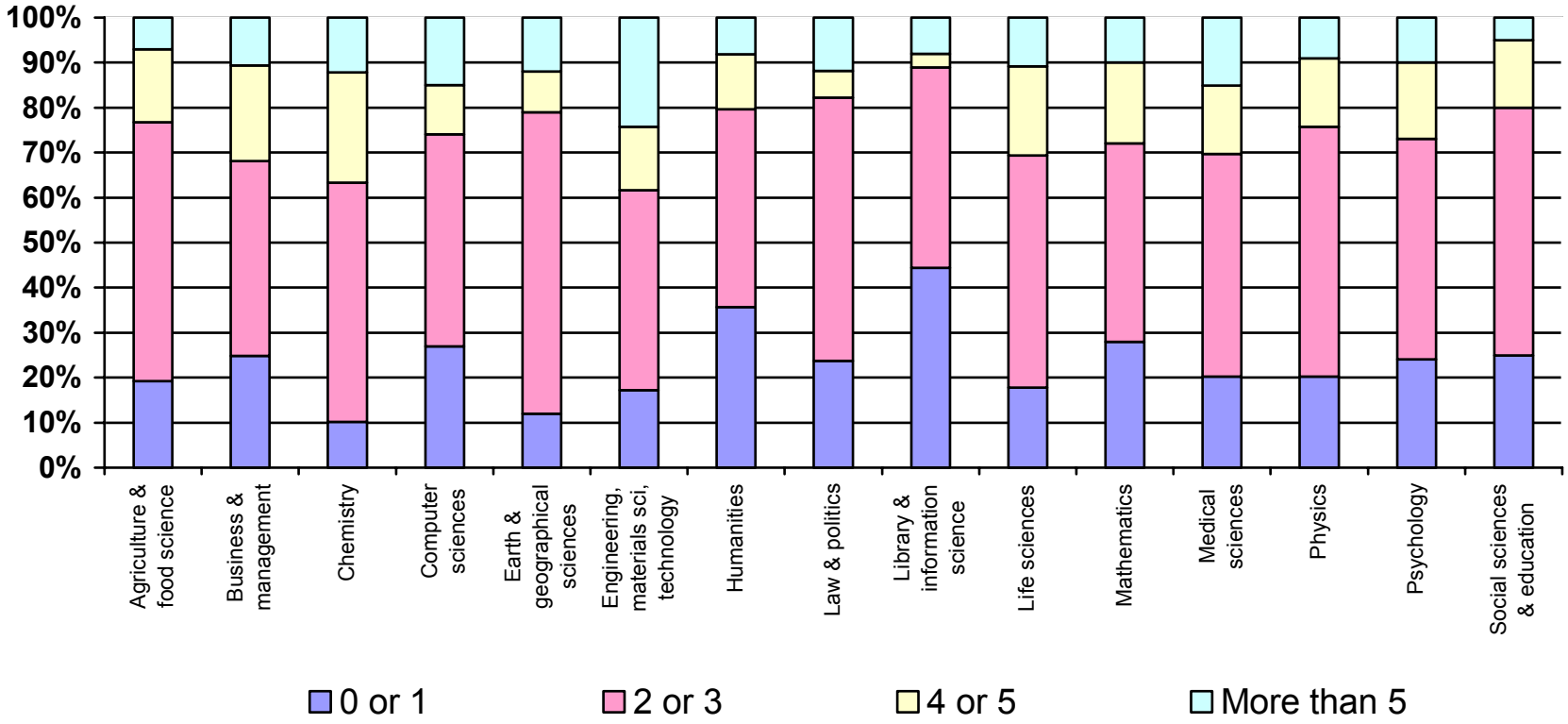

Figure 4: Number of articles published each year by subject area (stacked column chart)

4.3.2 Respondents' citation records

More than half (52%) of respondents don't know what their average citation count per article is two years after publication. Thirteen percent of respondents say their average citation count is 0-2 at this time point; 17% say it is 3-5, 10% say it is 6-10 and 7% say it is more than 10. Scholars in earth and geographical sciences claim the highest rate of citations. The full breakdown of answers to this question by subject area is given in Table 9 and Figures 5 and 6. Figures in the table are percentages of respondents and are rounded.

What is your average citation count per article?					
Subject	**0-2**	**3-5**	**6-10**	**More than 10**	**Don't know**
Agriculture & food science	21	19	10	11	39
Business & management	21	23	6	6	43
Chemistry	15	22	20	10	34
Computer sciences	21	21	7	5	49
Earth & geographical sciences	15	24	9	15	36
Engineering, materials science & technology	17	27	9	7	39
Humanities	12	9	6	6	65
Law & politics	6	0	0	0	94
Library & information science	21	14	5	5	52
Life sciences	7	15	12	8	58
Mathematics	21	15	6	4	53
Medical sciences	10	16	15	6	51
Physics	13	26	11	8	41
Psychology	12	17	11	6	53
Social sciences & education	12	9	8	7	63
Whole population	**13**	**17**	**10**	**7**	**52**

Table 9: Respondents' average citation count per article, by subject area

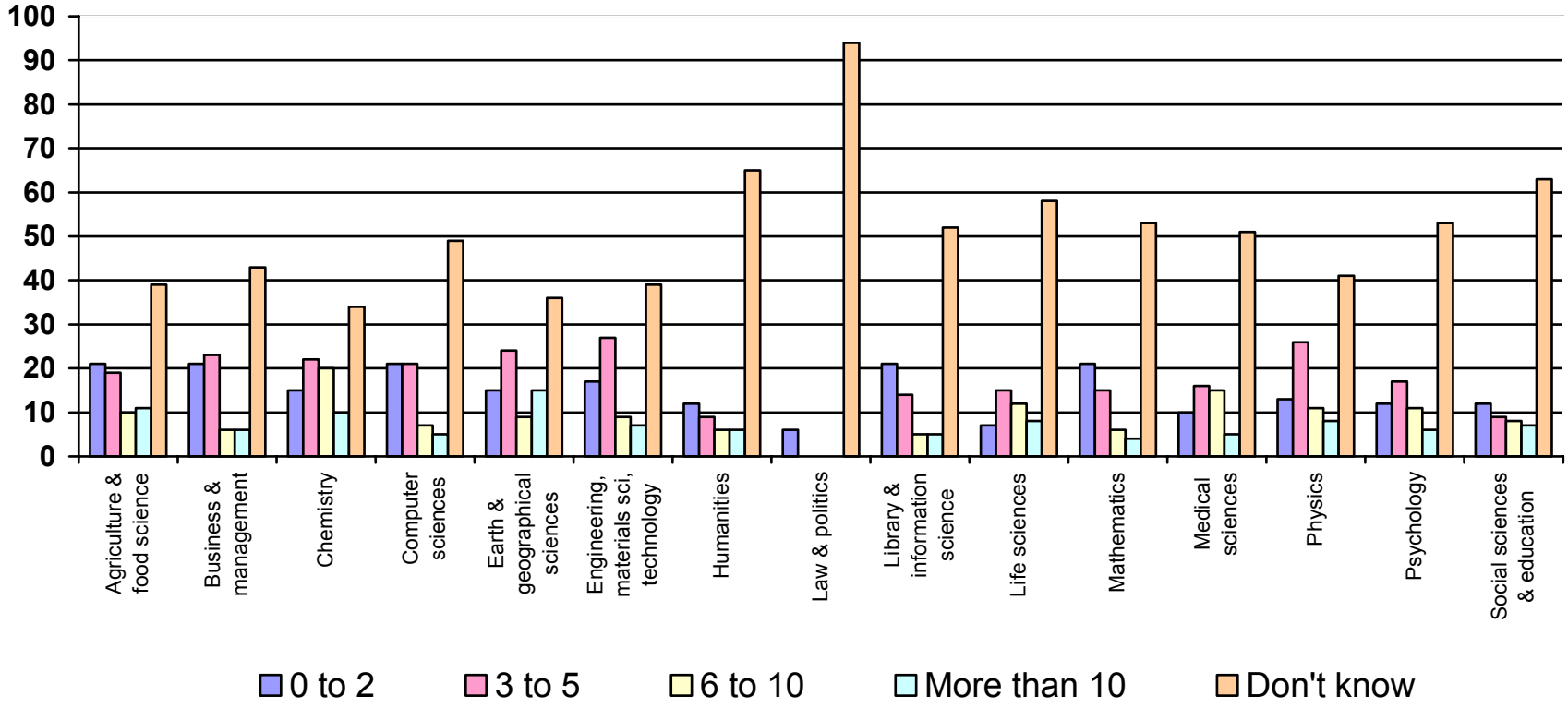

Figure 5: Respondents' average citation count per article, by subject area (clustered column chart)

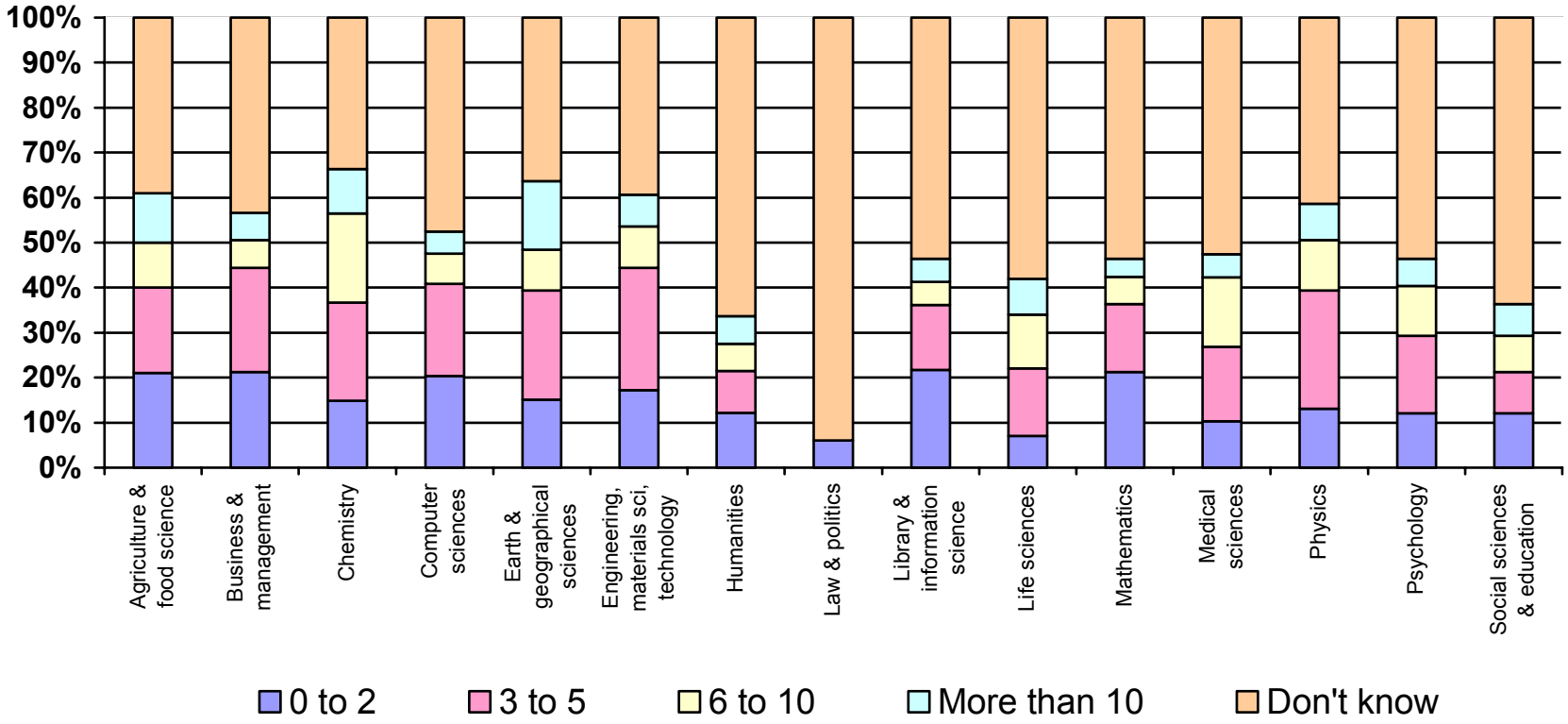

Figure 6: Respondents' average citation count per article, by subject area (stacked column chart)

4.3.3 Publishing objectives

Respondents were also asked to say what their objectives are in publishing their work. The results are shown in table 10. The shaded columns show the cumulated totals for the 'agree' and 'strongly agree' options and for the 'disagree' and 'strongly disagree' options, giving overall measures of agreement or disagreement with the statements. Figures are percentages of respondents and are rounded.

In general, what are your objectives when publishing scholarly work? Please indicate the degree to which you agree or disagree with the following statements						
Reason for publishing work	**Strongly agree**	**Agree**	**Strongly agree or agree**	**Disagree**	**Strongly disagree**	**Strongly disagree or disagree**
To communicate results to their peers	60	32	92	1	0	1
To advance their career	34	46	80	3	2	5
For personal prestige in their filed	24	52	76	5	1	6
To increase their chances of gaining funding	24	40	64	8	4	12
For direct financial reward	1	8	9	24	47	71

Table 10: Scholars' reasons for publishing their work

A number of respondents also offered additional comments after this question and these are reproduced verbatim in Appendix 6.

The answers here also corroborate findings from a previous study we carried out[1,2] and confirm that the primary reason that scholars publish is to make their research findings widely available for others to read and build upon. Career and personal reasons for publishing all come behind this foremost reason.

4.4 Searching for information

The study also explored how researchers look for research information – that is, in their role now as *users* of information. Bibliographic databases are quite heavily used. Nineteen percent of respondents search them daily and 42% do so at least once a week. A further 21% search them at least once per month. Only 16% do so less frequently than once per month. In all, 98% of respondents use this type of database as a search tool. This is in contrast to the proportion of people who use other means to locate research information they need for their work (see Sections 4.4.1 and 4.4.2).

4.4.1 Research articles in closed archives

When searching closed (or 'toll-access') archives, that is, collections where the user has access only to the full-text of articles in journals to which they or their institution have a subscription, most people use ScienceDirect (54%). Obviously it is the science-based disciplines that are the primary users of ScienceDirect. Scholars in the humanities and other non-science subject disciplines predominantly use other full-text services. The breakdown of these results by subject area is shown in Appendix 7. In total, 86% of people use this type of service.

The overall findings are shown in Table 11 below. Figures are percentages of respondents and are rounded.

When you search for articles online in CLOSED ARCHIVES, (i.e. where you are only allowed access to the full-text of articles in journals you or your library subscribe to), which services do you use regularly?	
Services used regularly to search closed archives	**Percentage of total respondents**
ScienceDirect	54
Individual publishers' websites	38
Subject-specific full-text services	25
Other cross-subject full-text services	22
CrossRef Search	7

Table 11: How users search for articles in closed archives

4.4.2 Research articles in open archives

Open access archives are reached by different means. Respondents were provided with a list of search engines that specifically harvest from these archives around the world and were asked to indicate which they use. The results are shown in Table 12. Figures are percentages of respondents and are rounded in most cases.

Which services(s) do you use to search the content of OPEN ACCESS ARCHIVES (i.e. where you can access the full-text of any research article)?	
Services used regularly to search open access archives	**Percentage of total respondents**
Scirus	14
Citebase Search	10
OAIster	3
Open Archives Initiative Information in Engineering, Computer Science & Physics (OAIIECSP)	3
Public Knowledge Project Open Archives harvester (PKP)	2
Perseus	1.6
Arc	1.5
CYCLADES	0.5
Callima	0.4
SAIL-eprints	0.2
TORII	0.1

Table 12: How users search for articles in open access archives

It should be noted that answers to this question were not mutually exclusive, so a respondent was able to give more than one answer. A total of 30% of respondents are represented here (391 individual respondents). This finding supports data presented later in this report that show that awareness of open access and how to use open access tools remains limited. Computer scientists are the most frequent users of these services, with life scientists and medical scientists following. Other disciplines show very low levels of usage in general.

Respondents who do search open access archives do so fairly infrequently compared to the findings for bibliographic databases, for example. Fifty percent search open access archives less frequently than once per month; 17% do so at least once per month, 12% at least once per week and 9% several times per week.

When it comes to searching the World Wide Web for research articles on scholars' websites, 72% use Google (note: this survey was carried out just before Google Scholar was launched late in 2004), and 23% use another search engine. 11% don't search websites for this kind of information.

5. SELF-ARCHIVING

There are three ways a researcher can provide open access to articles by self-archiving. S/he can deposit a copy of an article on a personal or institutional website, or place it in an institutional open access archive, or put it in a subject-based, centralised, open access archive (such as the physics archive, called arXiv, or Cogprints, the cognitive science archive). Articles may be in preprint (pre-peer review or pre-refereeing) or postprint (after peer review or refereed) form.

5.1 Self-archiving experience

5.1.1 The level of self-archiving activity

Respondents were asked to indicate how many times in the past three years they had deposited full copies of their research articles in these various ways. In each case the majority of researchers had not deposited any articles in that way. The full results for this question are shown in Table 13. Figures are percentages of respondents and are rounded.

In the past 3 years, how many times have you deposited full copies of the following version of research papers in the following ways?						
Open access provision method	**None**	**Once**	**2-3 times**	**4-5 times**	**>5 times**	**Total percentage self-archiving in this way**
Pre-refereeing draft on personal web page	79	5	7	3	7	21
Refereed, published research article on personal web page	69	6	9	4	13	31
Pre-refereeing draft in departmental or institutional OA archive	82	4	7	3	4	18
Refereed, published research article in departmental or institutional OA archive	75	5	8	4	8	25
Pre-refereeing draft in a centralised subject-based open archive	89	4	3	2	3	13
Refereed, published research article in a centralised subject-based open archive	85	4	3	3	5	15

Table 13: Patterns of self-archived open access provision by researchers

A breakdown of self-archiving activities by the subject area of respondents is shown in Table 14. Figures are the percentages of respondents in each subject area.

% who have self archived in the following ways						
Subject area of respondents	**Preprint on web page**	**Postprint on web page**	**Preprint in institutional archive**	**Postprint in institutional archive**	**Preprint in subject-based archive**	**Postprint in subject-based archive**
Agriculture & food science	8	28	18	29	10	15
Business & management	25	28	16	18	11	5
Chemistry	14	16	17	28	8	15
Computer sciences	45	60	37	42	20	24
Earth & geographical sciences	7	33	12	52	8	8
Engineering, materials science & technology	23	37	26	33	17	24
Humanities	43	45	15	24	12	16
Law & politics	38	38	27	20	13	0
Library & information science	31	32	21	20	10	16
Life sciences	12	27	10	25	6	8
Mathematics	33	44	27	25	38	33
Medical sciences	25	57	12	22	7	13
Physics	21	36	22	29	31	32
Psychology	20	33	7	16	9	12
Social sciences & education	31	30	28	20	8	13
Whole population	**21**	**31**	**18**	**25**	**13**	**15**

Table 14: Self-archiving activities by subject area of respondents

The data from this question are also presented in graphical form in Figures 7 and 8.

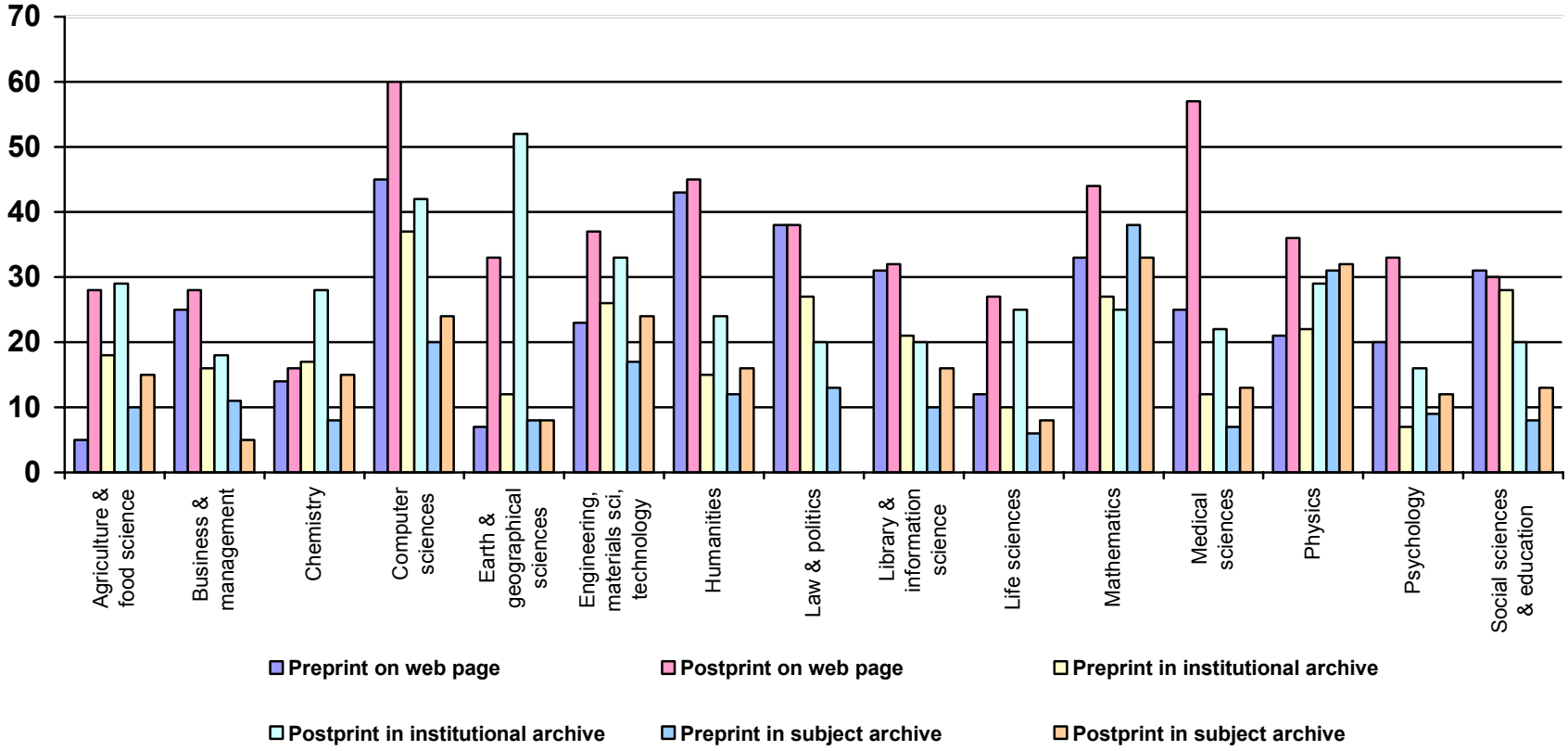

Figure 7: Self-archiving activity level by subject area. Bars show percentages of respondents in each subject area who have self-archived by each method (clustered column chart)

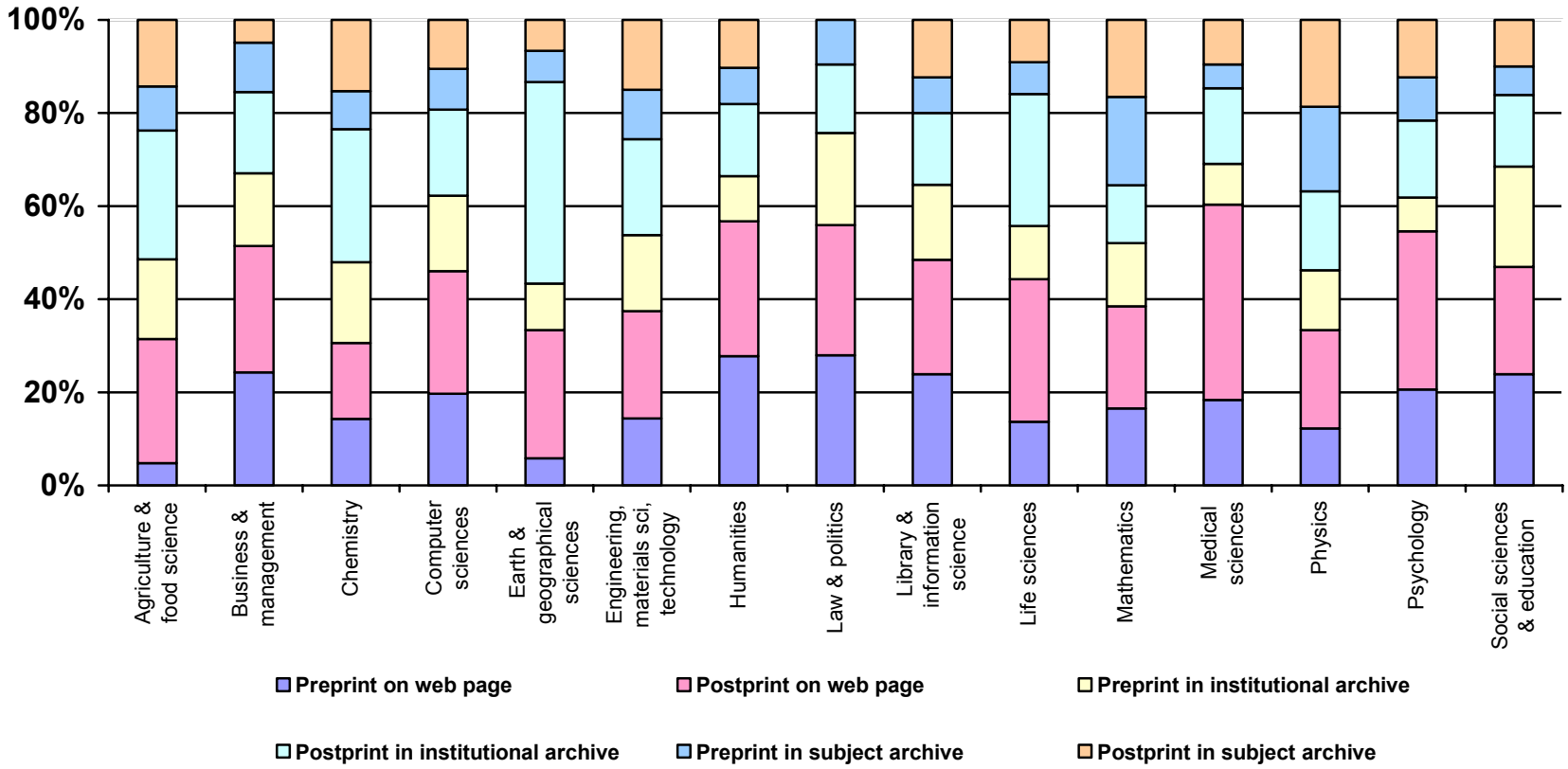

Figure 8: Self-archiving activity level by subject area. Bars show percentages of respondents in each subject area who have self-archived by each method (clustered column chart)

Finally, it is informative to report how many individuals are involved in self-archiving at all. The question permitted respondents to check any number of options, so that some of the people who have self-archived on their websites, for example, may also have self-archived in institutional or subject-based repositories. The overall findings presented above, therefore, do not reveal the true level of self-archiving within the population. To find out what this was, the results for this question were analysed on a case-by-case basis so that the number of individuals represented in the results was revealed.

The total number of people in the respondent population who have self-archived in any way is 631, that is, 49% of the total population. This total population of 1296 researchers has carried out 1303 individual acts of self-archiving; since 631 individuals have self-archived, the average number of self-archiving acts per self-archiver is 2.1. The last time we measured the level of self-archiving[1,2], 10 months earlier than this present survey, 23% of the population had self-archived in some way at least once. See below for more details of this earlier survey in comparison to the present one.

The proportions of self-archivers in each of the subpopulations (see Section 2 for definitions) surveyed is:

'Interested and informed' population:	35%
Archived population:	85%
Southampton population:	77%
Randomly-selected population:	42%

It can be noted that the 'interested and informed' population is not self-archiving as much as the average, or as much as the randomly-selected population, so the term 'interested' is something of a misnomer. The results for the 'archived' and 'Southampton' populations are also worthy of comment. The archived population comprised individuals whose names were obtained from papers stored in open access archives (see section 2). In practice, then, every one of these individuals should be a self-archiver, by definition. Nonetheless, 15% of them have not recorded any self-archiving activity in this survey, underlining the fact that an author may become a self-archiver by default, that is, by his or her papers being archived by co-authors or by proxy archivists (librarians, archive administrators, colleagues). The Southampton population was comprised of researchers within the School of Electronics & Computer Science at Southampton University, all of whom publish research papers every year and, since the School imposed a mandate on self-archiving article from January 2003, should have more than one article in the Southampton repository. In this case, though, 23%

do not record any self-archiving activity. Examination of the logs for the Southampton repository (courtesy of Dr Les Carr) suggests that the explanation for this is that colleagues or departmental administrators have actually deposited the articles in many cases, leaving individual researchers either unaware of the fact that their work resides in the repository or insufficiently familiar with the procedures to feel able to respond to the questions in this survey. Who is doing the depositing of self-archived articles is examined in Section 5.4.

There is more that can be learned from this question, too, as follows.

How do these different respondent populations compare in detail?
Table 15 shows the level of self-archiving for each of the respondent subpopulations. It shows the number of people in each subpopulation who have self-archived articles in one or more of the six possible ways (which are: putting a preprint on a web page; putting a postprint on a web page; putting a preprint in an institutional archive; putting a postprint in an institutional archive; putting a preprint in a subject archive; or putting a postprint in a subject archive). Some people have done all six, some none, and some have done some or many of these things. Figures in Table 15 are all percentages of the sub-populations concerned and are rounded.

Respondent sub-populations	**Number of ways in which respondents have self-archived their articles**					
	1	**2**	**3**	**4**	**5**	**6**
'Interested & informed' population	20	7	6	0	2	1
Archived population	27	31	14	8	4	2
Southampton population	29	26	11	11	0	0
Randomly-selected population	20	13	4	3	1	2

Table 15: Patterns of self-archiving by respondent sub-populations

Table 16 compares the findings from this present survey (last quarter 2004) with those from the previous survey (January 2004). Figures are percentages of the total respondent population and are rounded.

Open access provision method	**Earlier survey**	**Present survey**
Pre-refereeing draft on personal web page	23	18
Refereed, published research article on personal web page	22	27
Pre-refereeing draft in departmental or institutional OA archive	4	15
Refereed, published research article in departmental or institutional OA archive	10	20
Pre-refereeing draft in a centralised subject-based open archive	7	9
Refereed, published research article in a centralised subject-based open archive	7	12

Table 16: Comparison of earlier and present surveys with respect to self-archiving patterns

The data in Table 16 are shown graphically in Figure 9.

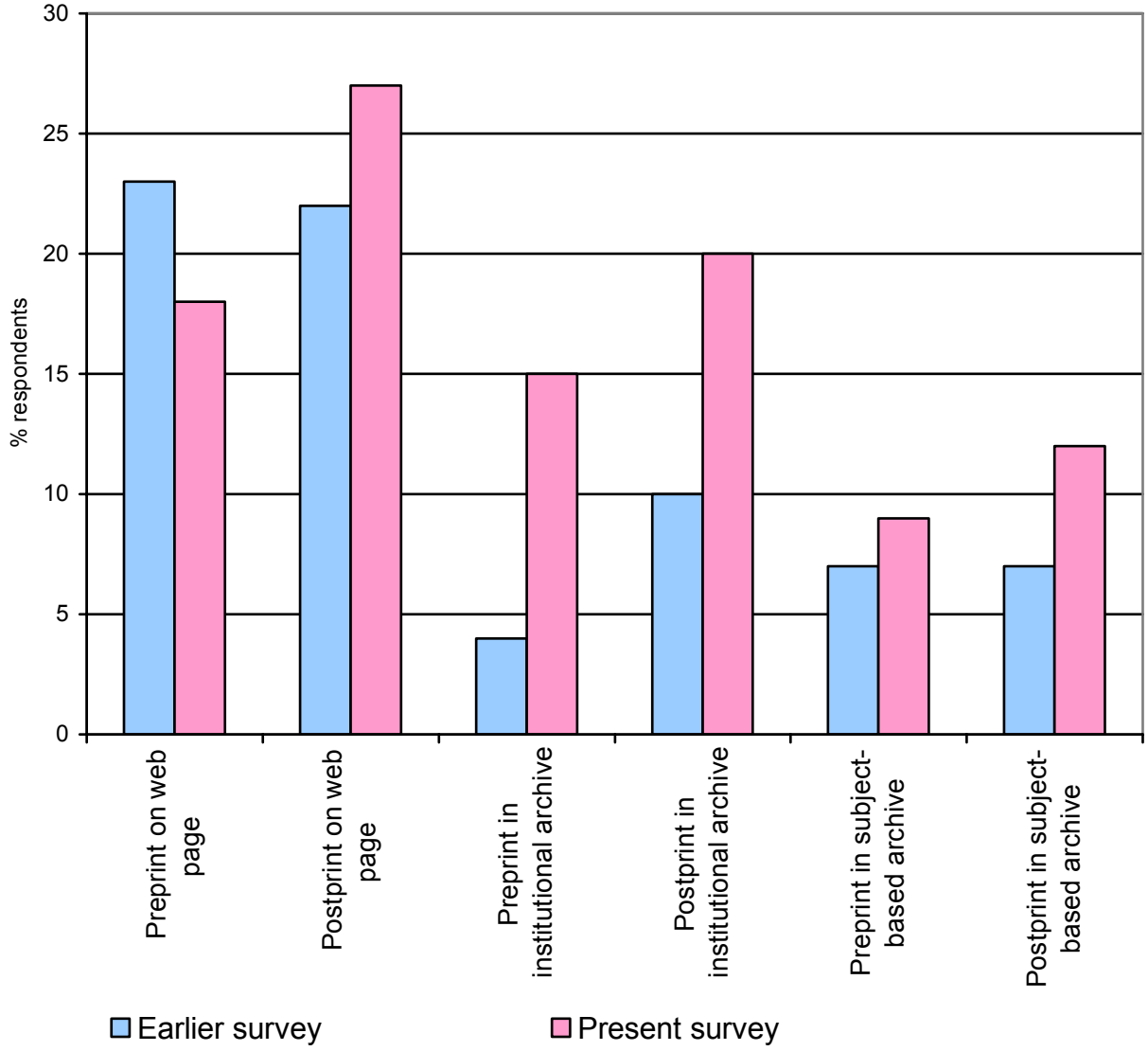

Figure 9: Comparison of earlier and present surveys with respect to self-archiving patterns

Table 17 presents the data on the number of individual acts of self-archiving that have been carried out by each of the subpopulations, along with a figure for the average number of self-archiving acts per person. The numbers in this table are absolute numbers.

Open access provision method	Whole population	'Interested and informed' population	Archived population	Randomly-selected population	Southamp-ton population
Pre-refereeing draft on personal web page	231	107	12	108	4
Refereed, published research article on personal web page	344	130	17	181	16
Pre-refereeing draft in departmental or institutional OA archive	192	55	19	102	16
Refereed, published research article in departmental or institutional OA archive	262	83	23	139	17
Pre-refereeing draft in a centralised subject-based open archive	119	32	14	62	1
Refereed, published research article in a centralised subject-based open archive	155	52	13	88	2
Total number of acts of self-archiving by this population	1303	469	98	680	56
Average number of acts of self-archiving per researcher	1.01	**1.15**	**1.75**	**0.83**	**1.60**
Number in population	1296	398	52	811	35

Table 17: Individual acts of self-archiving of different types by each respondent population

The correlation between self-archiving activity and number of articles published can also be examined. Table 18 shows the results of this exercise. The results are also shown graphically in Figures 10 and 11.

	Percentage of respondents			
Open access provision method	**Those publishing 0-1 papers per year**	**Those publishing 2-3 papers per year**	**Those publishing 4-5 papers per year**	**Those publishing more than 5 papers per year**
Pre-refereeing draft on personal web page	23	20	22	21
Refereed, published research article on personal web page	25	30	36	39
Pre-refereeing draft in departmental or institutional OA archive	12	19	14	20
Refereed, published research article in departmental or institutional OA archive	20	25	26	29
Pre-refereeing draft in a centralised subject-based open archive	15	10	12	11
Refereed, published research article in a centralised subject-based open archive	17	13	15	17

Table 18: Self-archiving activities by respondents publishing varying numbers of papers per year

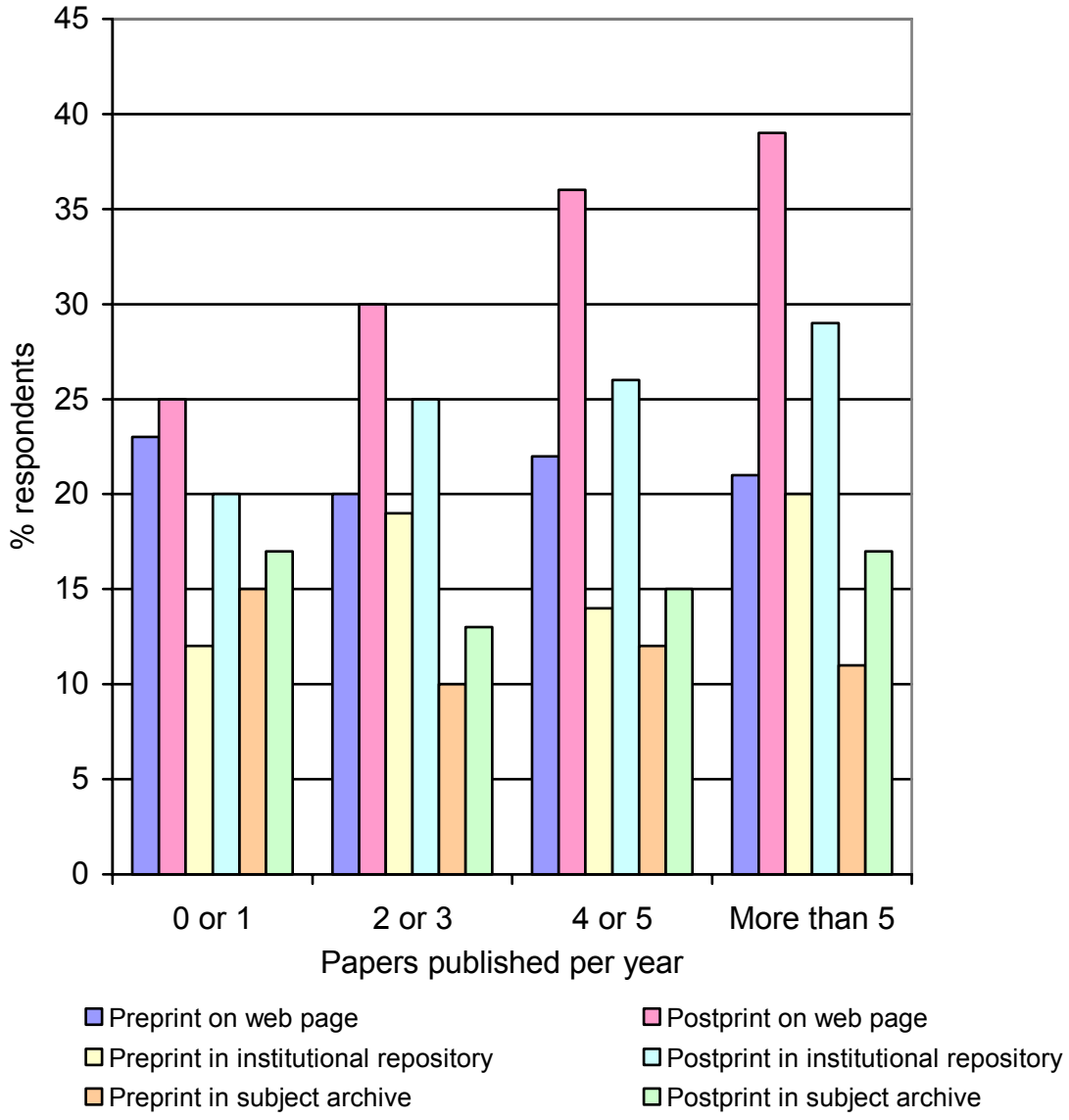

Figure 10: Self-archiving activities by respondents publishing varying numbers of papers per year (clustered column chart)

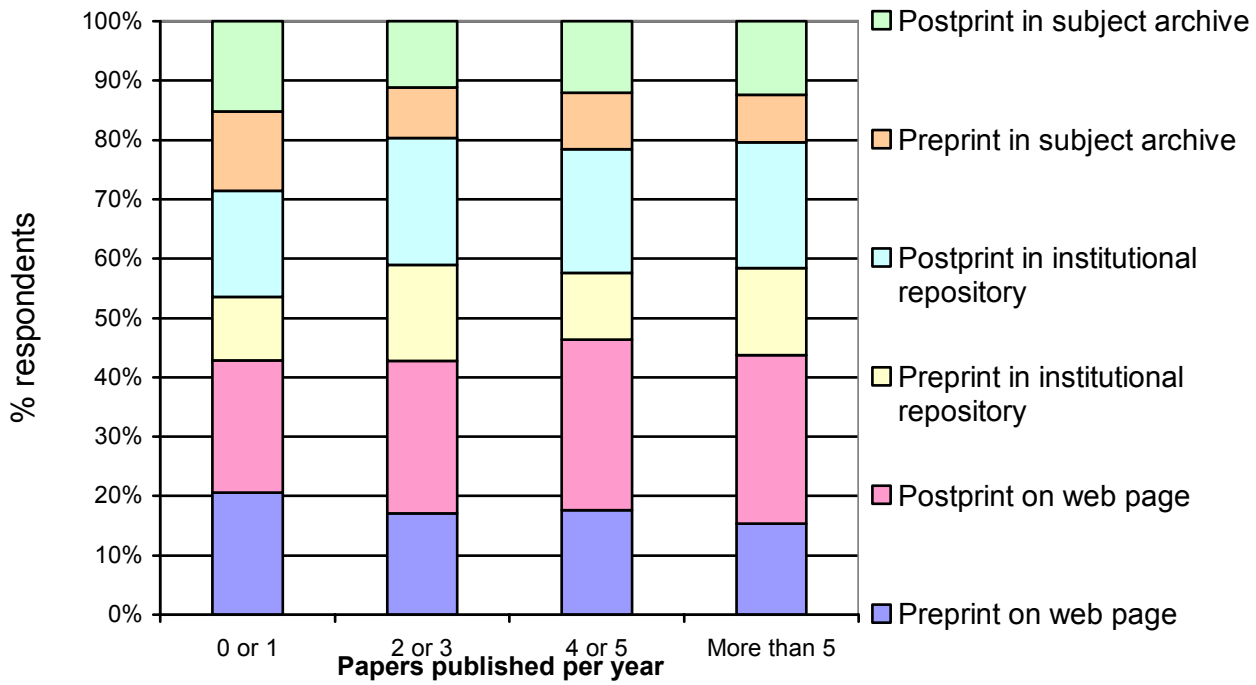

Figure 11: Self-archiving activities by respondents publishing varying numbers of papers per year (stacked column chart)

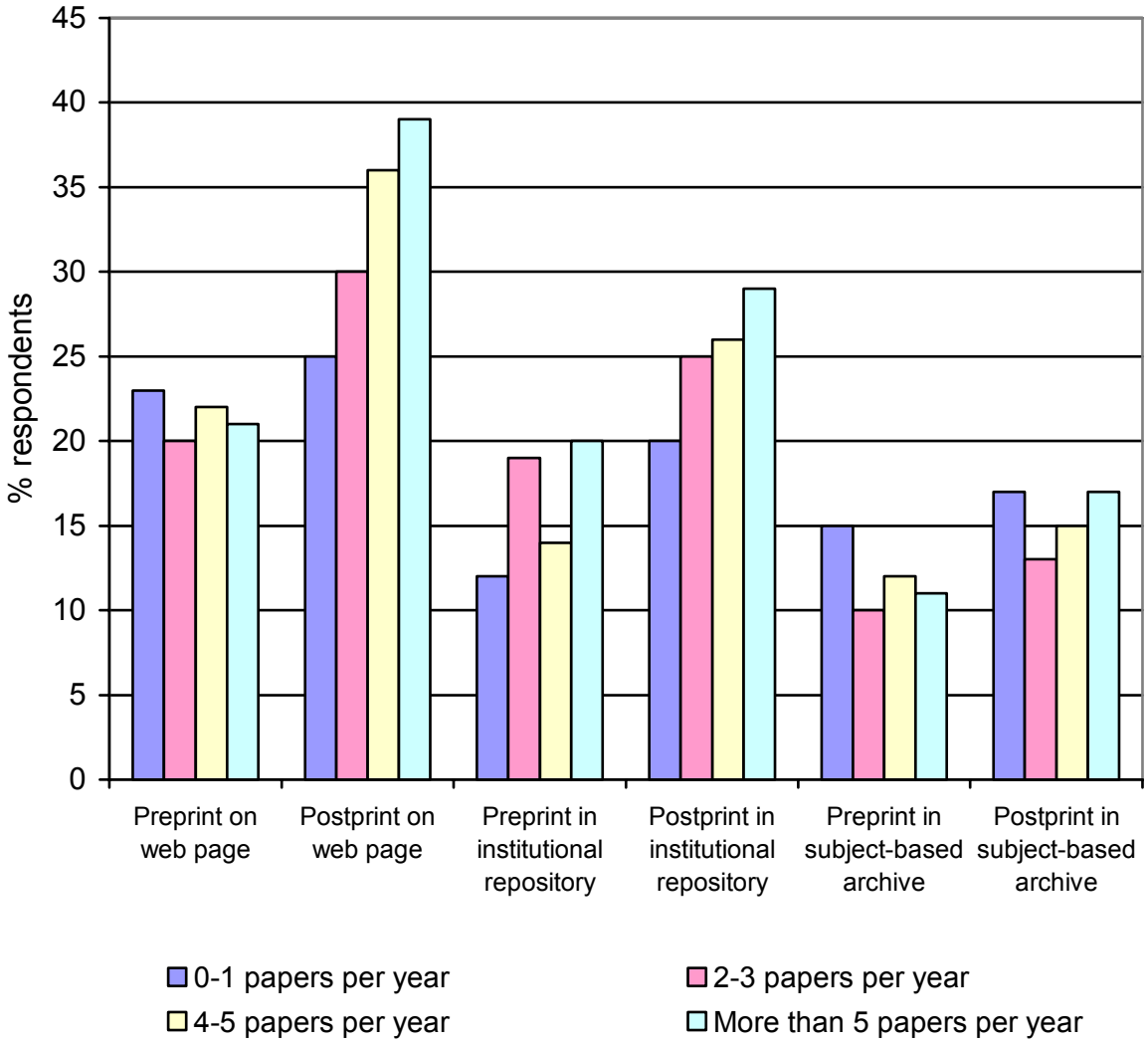

Figure 12: Self-archiving activities by respondents publishing varying numbers of papers per year. Data expressed by activity (clustered column chart)

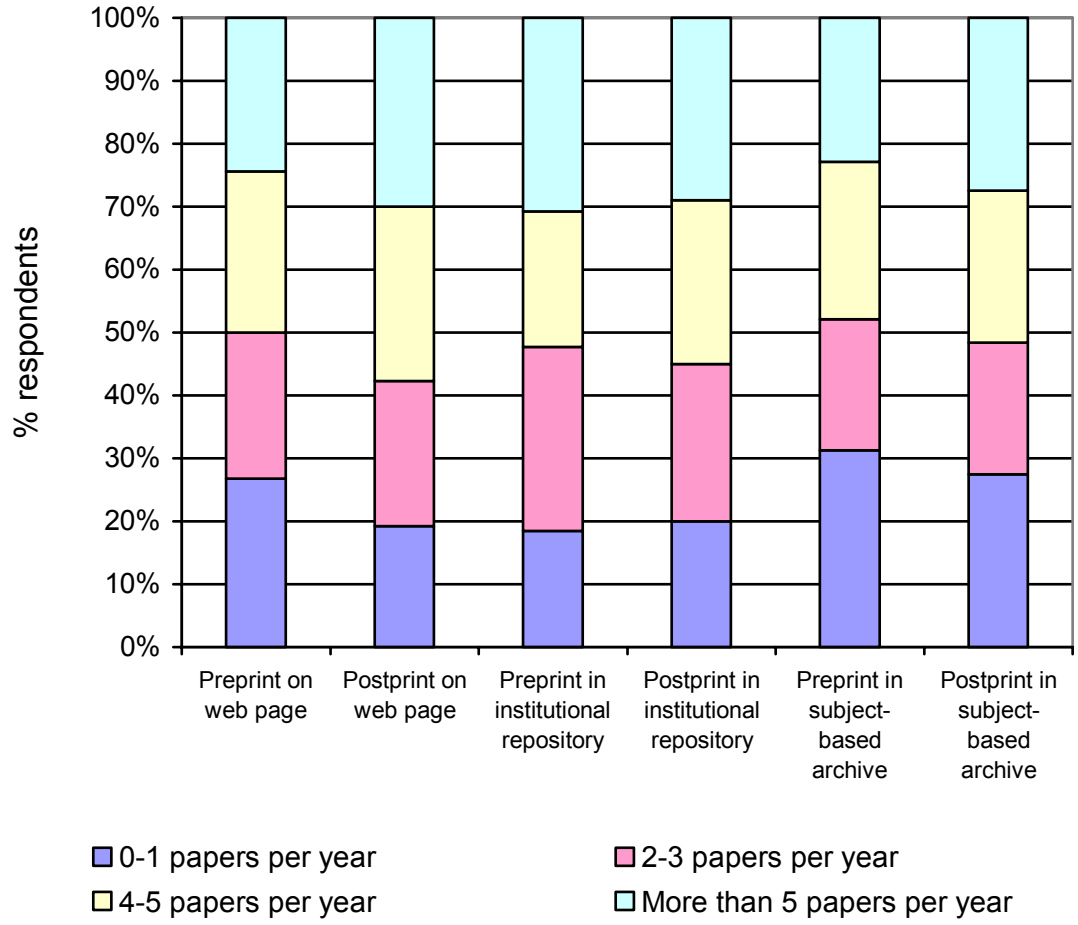

Figure 13: Self-archiving activities by respondents publishing varying numbers of papers per year. Data expressed by activity (stacked column chart)

5.1.2 Length of experience of self-archiving

Those respondents who had self-archived were asked for how long they had been doing this. The results for the whole population are shown in Table 19 and Figure 14. Figures in the table are percentages of respondents who have self-archived and are rounded.

If you have done any of the above, for how long have you been doing this?	
Experience of self-archiving	**Percentage of respondents**
Within the last year	21
2-3 years	32
3-5 years	21
Longer than 5 years	26

Table 19: Length of time for which researchers have been self-archiving

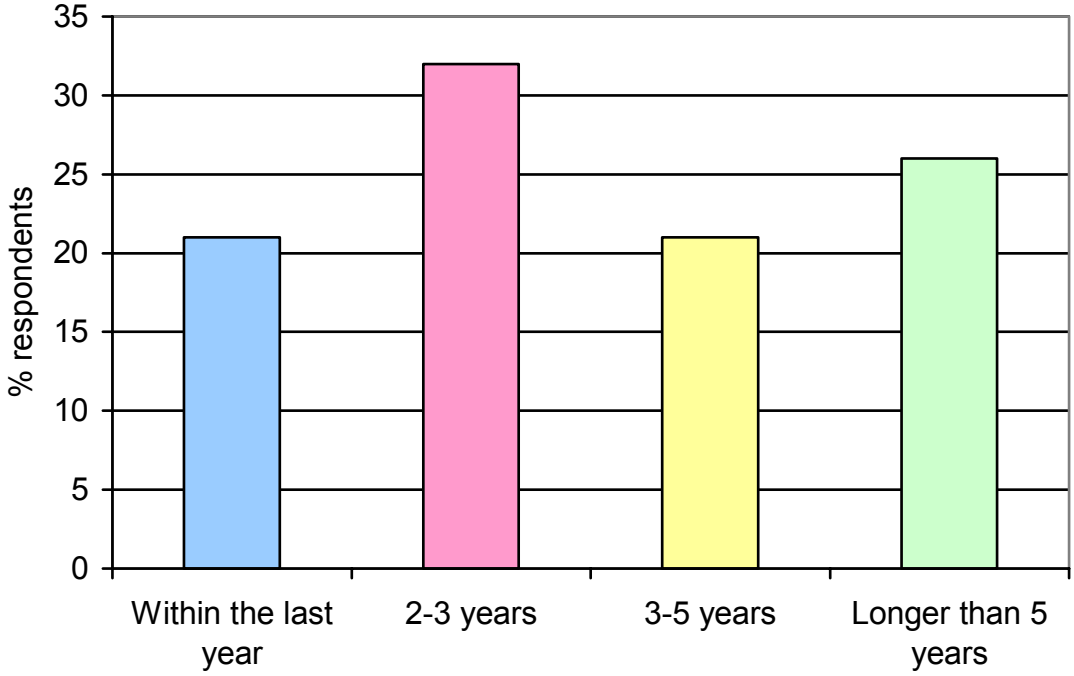

Figure 14: Length of time for which researchers have been self-archiving

It appears that the momentum is growing with respect to self-archiving activity. Just over a quarter of self-archivers began this activity over 5 years ago. Between three and five years ago roughly ten percent of them began to self-archive each year. This increased to 16% per year more recently and the last year accounted for 21% of the total taking up the self-archiving habit.

By subject area, the results break down as in Table 20 and the results are also shown graphically in Figures 15 and 16. Figures are percentages of respondents in each category and are rounded. As expected, because of the 14-year existence of the arXiv, the high-energy physics-based archive, the groups who have been self-archiving longest are physicists and computer scientists. With respect to the latter discipline, self-archiving has been going on longer and the numbers of articles archived are larger. (Law & politics scores highly in this regard but the overall numbers of respondents in this category is very small and this must be taken into account when interpreting these findings.)

If you have done any of the above, for how long have you been doing this?				
Subject area of respondents	**Longer than 5 years**	**3-5 years**	**2-3 years**	**Within the last year only**
Agriculture & food science	22	11	47	19
Business & management	25	25	20	30
Chemistry	19	19	30	32
Computer sciences	41	17	28	15
Earth & geographical sciences	12	12	38	38
Engineering, materials science & technology	24	25	32	19
Humanities	18	32	28	22
Law & politics	46	18	9	27
Library & information science	16	19	30	35
Life sciences	26	21	33	20
Mathematics	29	33	24	14
Medical sciences	15	20	43	23
Physics	35	30	25	11
Psychology	19	22	41	19
Social sciences & education	14	23	37	26
Whole population	**6**	**21**	**32**	**21**

Table 20: Length of self-archiving experience in different subject areas

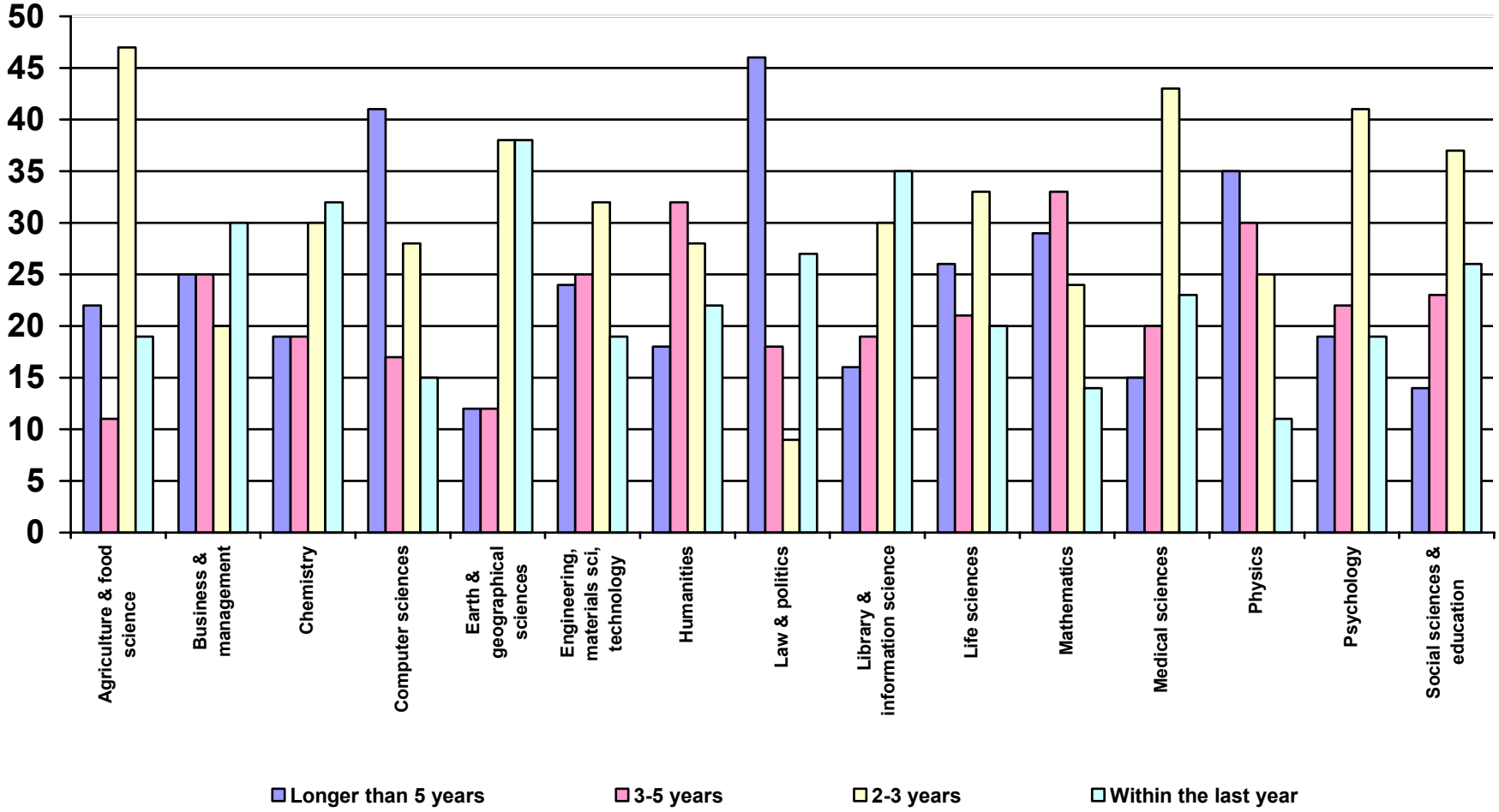

Figure 15: Length of self-archiving experience in different subject areas (clustered column chart)

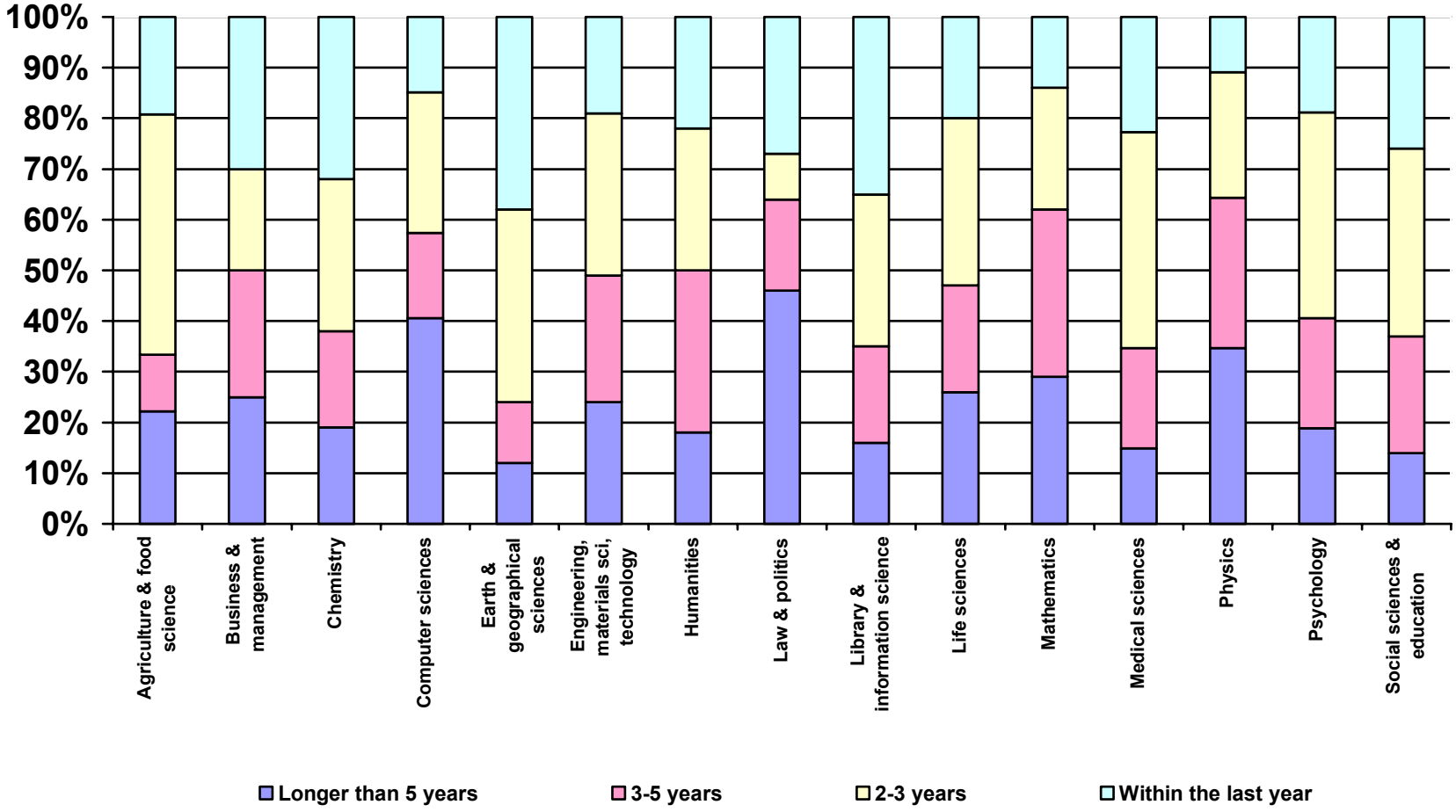

Figure 16: Length of self-archiving experience in different subject areas (stacked column chart)

5.2 Awareness of self-archiving as a means to providing open access

Previous surveys by KPL[1,2,17,18] and others[22] have indicated that there is a substantial level of ignorance within the scholarly community with respect to open access, both open access journals and self-archiving. Those respondents who had *not* self-archived their work by any means were asked whether they were aware of the possibility of providing open access to their work in this way. Twenty nine percent of them were aware of this and 71% were not. Since the non-archivers represent 51% of the whole population, this means that 39% of researchers overall are as yet unaware of self-archiving as a means to providing open access to their work.

Researchers in some subject areas have a greater level of awareness of self-archiving as a route to open access than others. Library/information science scholars are particularly notable in this respect, as are the communities that have self-archived for longest – physics, computer sciences and mathematics. Table 21 shows the figures (percentages of respondents, rounded): the data are also displayed graphically in Figures 17 and 18.

If you have not deposited drafts or refereed, published research articles in any of the ways listed, are you aware of the possibility of providing open access by self-archiving your work in open archives?		
Subject area	**Yes**	**No**
Agriculture & food science	21	79
Business & management	23	77
Chemistry	29	71
Computer sciences	53	47
Earth & geographical sciences	38	63
Engineering, materials science & technology	38	62
Humanities	34	66
Law & politics	33	67
Library & information science	60	40
Life sciences	20	80
Mathematics	47	53
Medical sciences	14	86
Physics	41	59
Psychology	26	72
Social sciences & education	27	73
All respondents	**29**	**71**

Table 21: Awareness of self-archiving as a means to providing open access: results by subject area

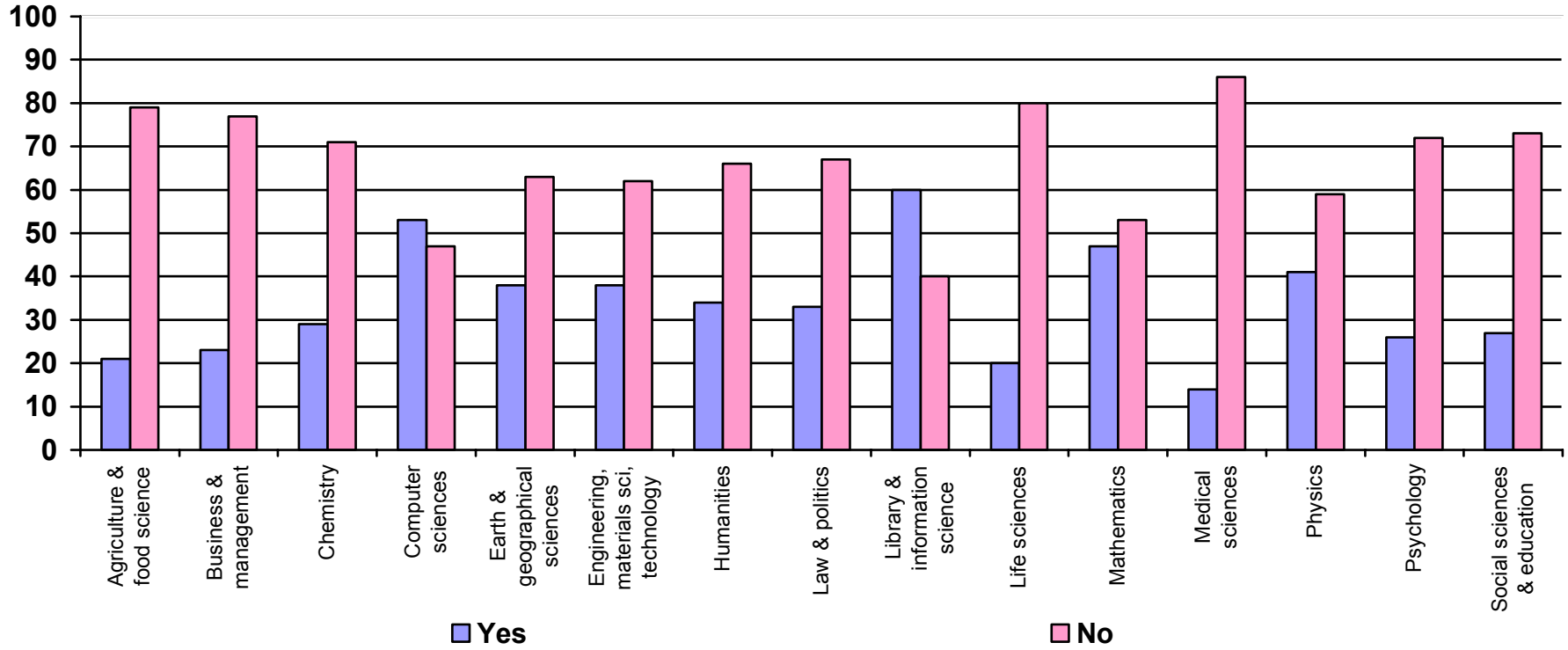

Figure 17: Awareness of self-archiving as a means to providing open access: results by subject area (clustered column chart)

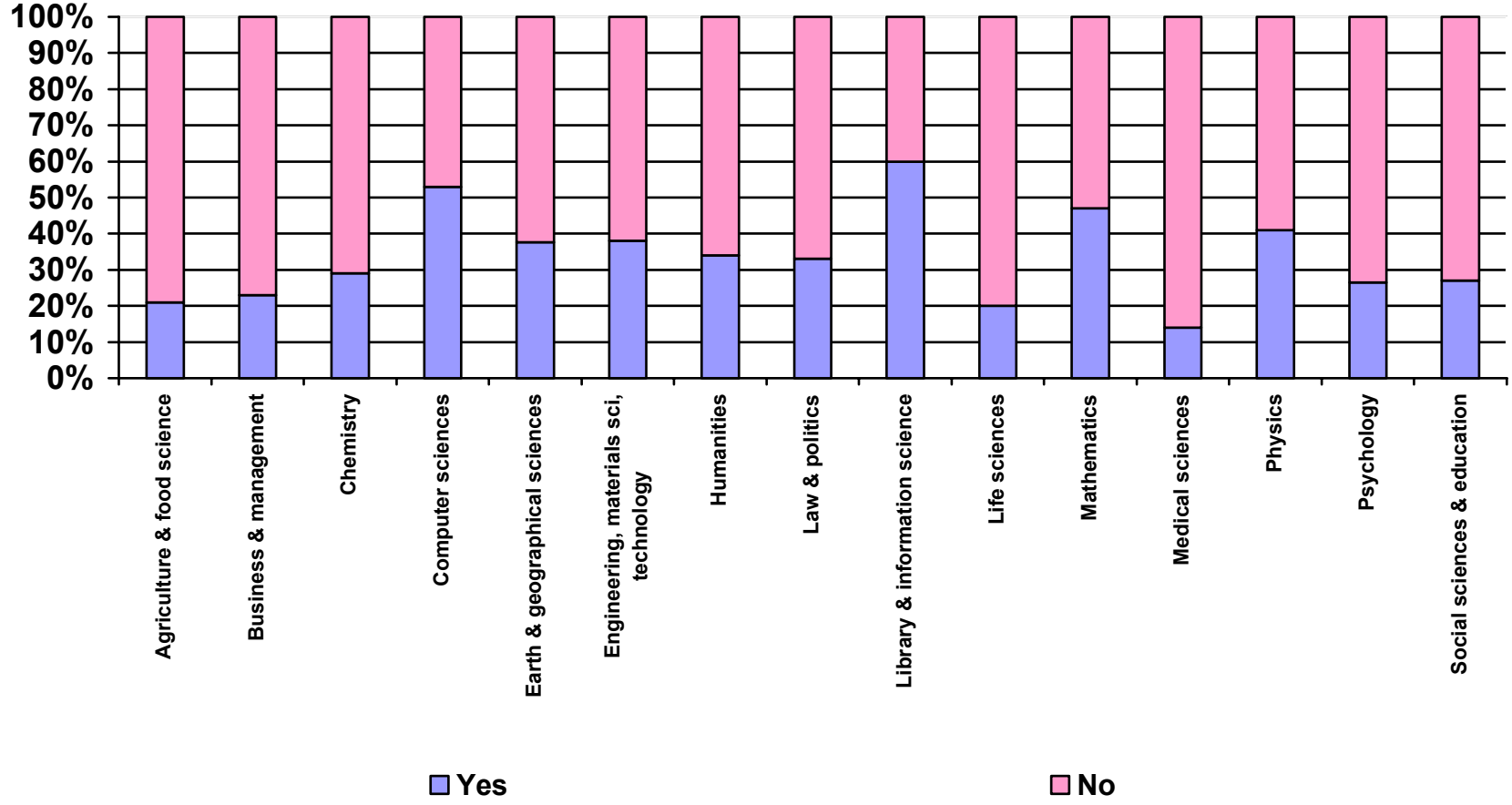

Figure 18: Awareness of self-archiving as a means to providing open access: results by subject area (stacked column chart)

A breakdown of the results for the awareness of self-archiving as a means to providing open access by the geographical origin of the respondents is shown in Table 22 and graphically in Figure 19. Figures are percentages of respondents and are rounded. Note that China and Central/South America score highly for awareness, but care should be taken when interpreting these figures since the numbers of respondents from these regions was small overall (see Section 2).

If you have not deposited drafts or refereed, published research articles in any of the ways listed, are you aware of the possibility of providing open access by self-archiving your work in open archives?		
Region	**Yes**	**No**
Australia/New Zealand	21	79
Asia (except China and Japan)	41	59
China	34	66
Japan	31	69
Canada	18	82
USA	27	73
Central/South America	39	61
European union (except UK)	31	69
Other European countries (except EU or UK)	31	69
UK	27	73
Middle East	26	74
Africa	27	73
All respondents	**29**	**71**

Table 22: Awareness of self-archiving as a means to providing open access; results by geographical origin of respondents

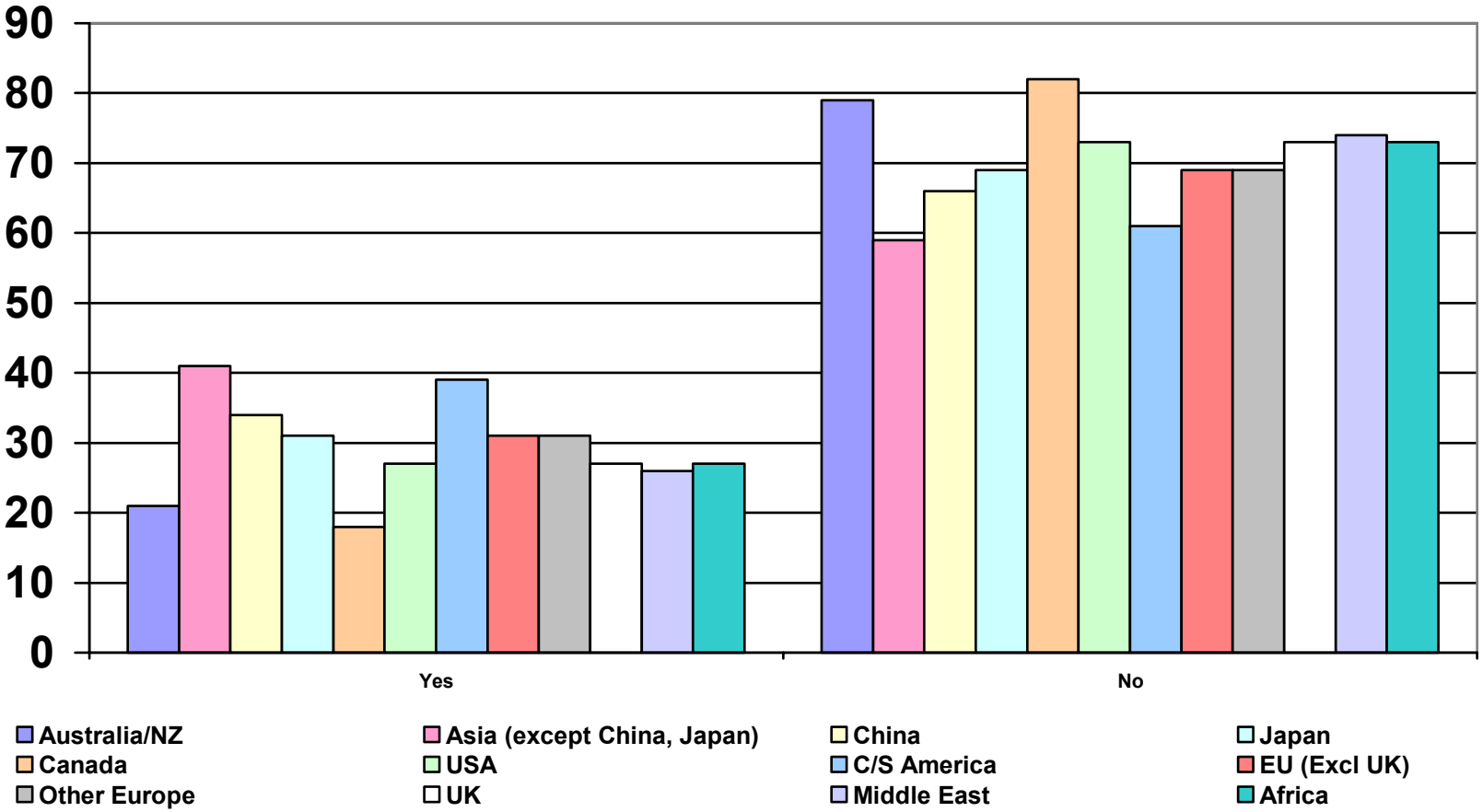

Figure 19: *Awareness of self-archiving as a means to providing open access; results by geographical origin of respondents*

There is a much greater lack of awareness about other issues surrounding self-archiving. Only 10% of self-archivers know about the SHERPA/RoMEO Directory of journal and publisher policies on self-archiving[23] and less than one quarter of them were familiar with the recent proposals on self-archiving by the House of Commons Select Committee on Science & Technology in the UK and by the National institutes of Health in the USA[24,25].

Those that *are* aware of the possibility of self-archiving learned about it from a range of sources. These are shown in Table 23 and Figure 20. Figures are percentages of people aware of self-archiving and are rounded.

How did you originally learn about self-archiving in an institutional or subject-based archive (as opposed to using your own web page) as a means to provide open access to your work? Please select any that apply	
How respondents originally learned about self-archiving	**Percentage of respondents**
Work in a field with established subject-based archives	14
Followed the debate on open Access	21
From information provided by institution or library	16
From information provided by department or school	9
From peers	22
From co-authors	9

Table 23: Original source of information on self-archiving

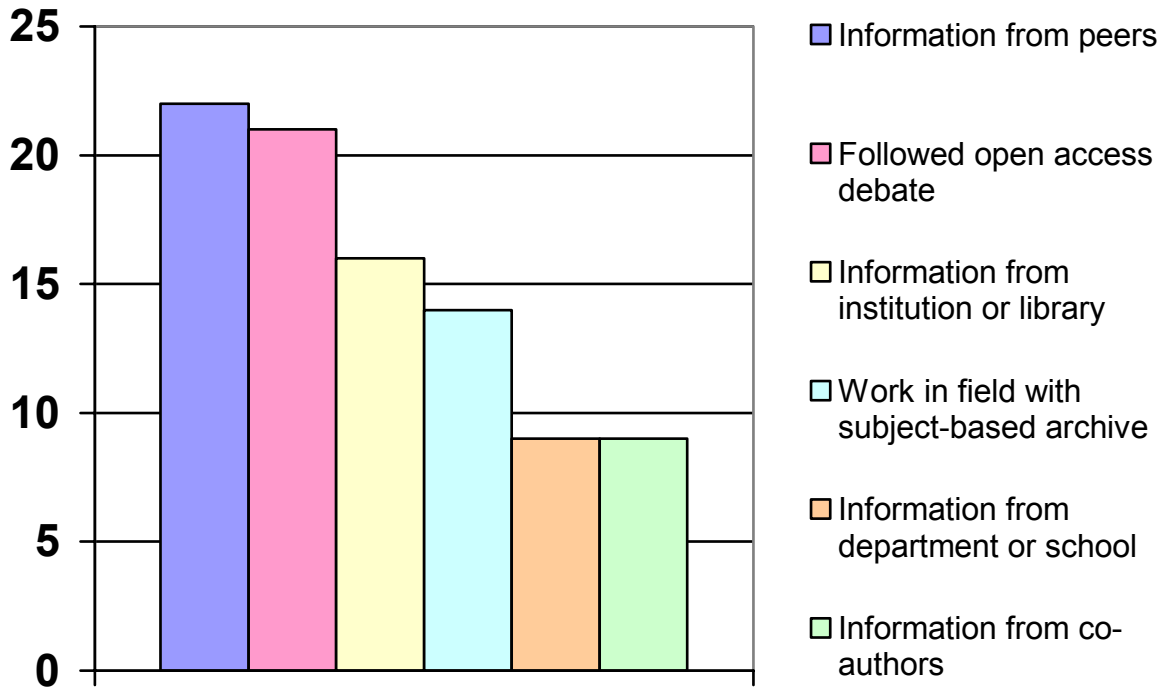

Figure 20: Original source of information on self-archiving

This corroborates anecdotal evidence we have learned from discussions with archive administrators who say that one of the most effective means of advocacy for an institutional archive is to identify a researcher who has self-archived his/her work, seen the benefit of doing so and is keen to promote this benefit to colleagues and peers; in other words, to act as a 'champion' within the organisation. An individual in this role has been found to be much more persuasive amongst their peer community than any of more formal advocacy methods such as seminars, posters, demonstrations and so forth, a finding borne out by the findings reported in the next section below.

5.3 Motivation issues

Self-archiving respondents were asked what the original motivator was for self-archiving their work. The results are shown in Table 24 and Figure 21. Figures are percentages of respondents and are rounded.

What was your original motivation for self-archiving your work? Please select any that apply	
Original motivator for self-archiving	**Percentage of respondents**
I was predominantly self-motivated	67
Encouragement from peers or co-authors	21
Open access articles are cited more often than articles accessible only in subscription journals	18
Encouragement from library or administrative personnel at respondent's institution	15
Encouragement from departmental advocates	9
Encouragement from research funders	7

Table 24: Original source of motivation for self-archiving

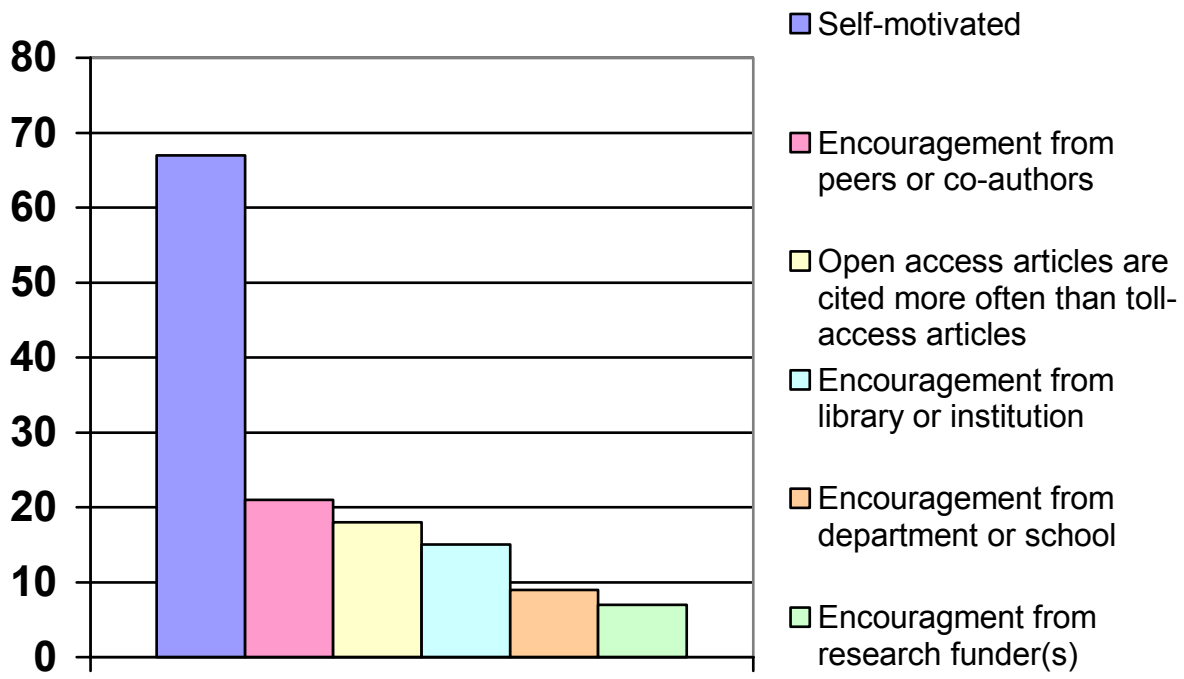

Figure 21: Original source of motivation for self-archiving

It is of interest to delve a little deeper into the data here, specifically to see what original sources of information about self-archiving were used by those who (a) were primarily self-motivated to self-archive and (b) declare that the citation-advantage of open access articles is the primary motivator for self-archiving.

The data are shown in Table 25. The figures show the percentage of people in these two groups (a) and (b) who originally learned about self-archiving in

different ways. The purpose of this analysis is to indicate the most effective advocacy tools so far and the extent of their influence.

How these self-archivers originally learned about the practice	Self-motivated self-archivers	Self-archivers for whom open access citation advantage was the primary motivation for self-archiving
Followed the debate on Open Access	38	45
From peers	36	35
Subject archive in their discipline	28	36
Information from institution or library	16	18
Information from co-authors	12	13
Information from department or school	12	10

Table 25: Sources of information on self-archiving used by people who were self-motivated to self-archive, or who were motivated by the citation advantage of self-archiving

5.4 The mechanics of self-archiving

5.4.1 Who has actually done the self-archiving?

There are a number of archives around the world where the majority of articles have been deposited by archive administrative staff in order to begin the archive-populating process. We were interested to know the extent of this by asking respondents who have self-archived how the actual archiving process had been effected. In the event, 80% of self-archivers have deposited their articles themselves; in 19% of cases the library staff archived articles for them and in 10% of cases this was carried out by students or assistants. Note that these answers were not mutually exclusive, so for the same person some articles may have been deposited by one means while others may have been deposited by another means.

5.4.2 How difficult is it to self-archive?

More than half of the self-archivers appear to have found it easy to carry out the procedure. Twenty two percent found it *very easy* to archive the first article they deposited and 32% found it *easy*. A further 20% found it *neither easy nor difficult*. Sixteen percent found it *somewhat difficult* and just 4% found it *difficult*. The procedure was carried out by a third party in 6% of cases. The data are shown in Figure 22 below.

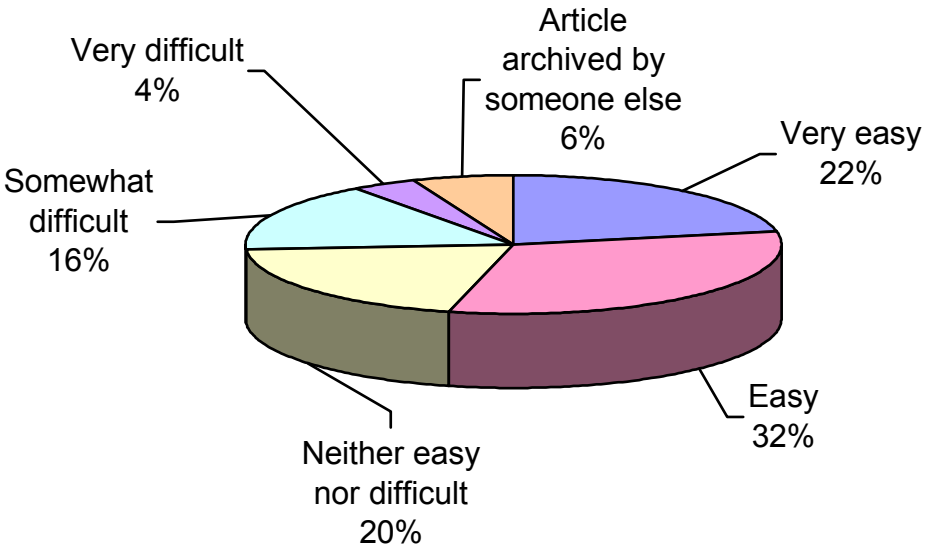

Figure 22: Ease of self-archiving an article for the first time

Subsequent article-deposition events are even easier. The results are shown in Figure 23 below.

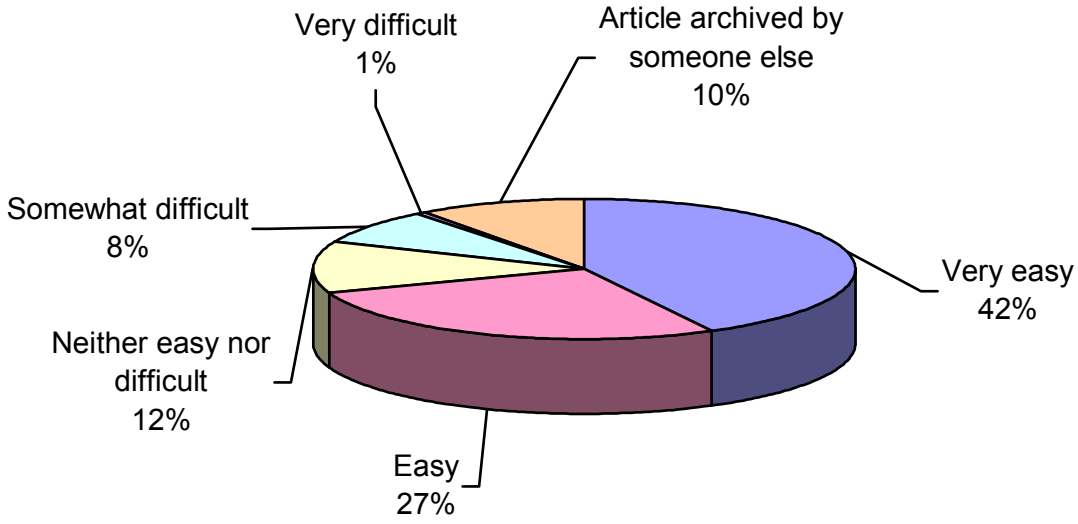

Figure 23: Ease of self-archiving subsequent articles

5.4.3 How long does it take to self-archive?

The first article that an author self-archives takes between a few minutes and an hour to deposit in most cases. Subsequent articles are deposited more quickly. The findings on this issue are shown in Figures 24 and 25 below. They tally with those recently determined from the archive of the School of Electronics & Computer Science at Southampton University[26]. An analysis of the computer records shows that it takes around 10 minutes for an author to deposit an article (that is, to enter the metadata and deposit the full-text file).

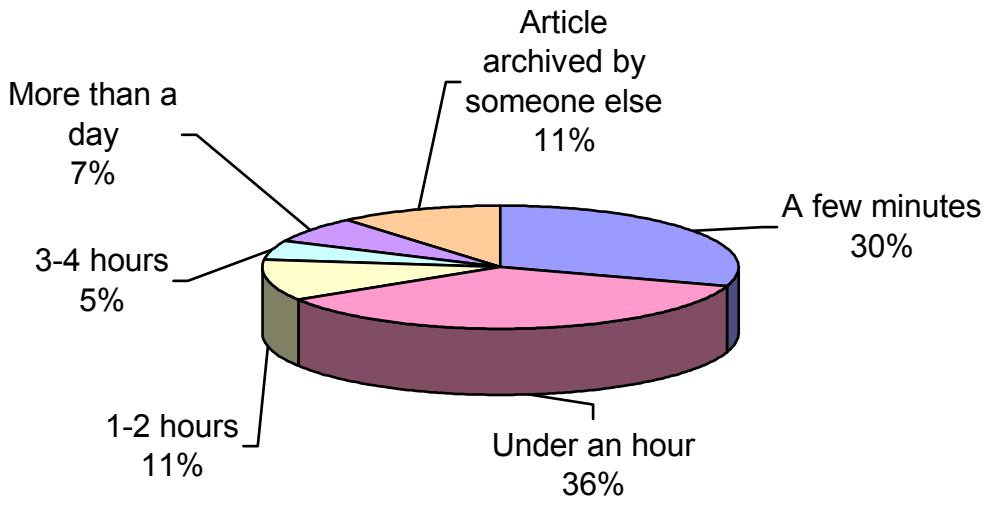

Figure 24: Time taken to self-archive the first article

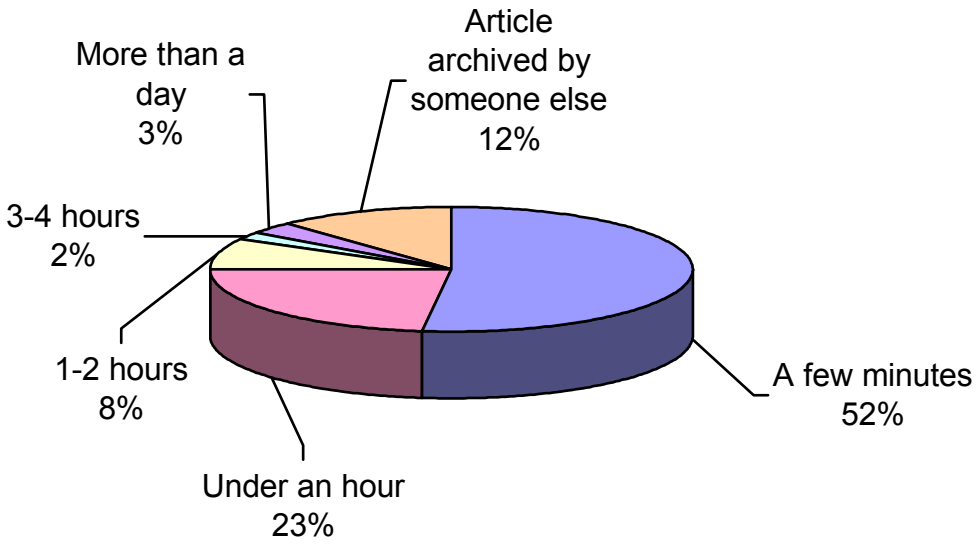

Figure 25: Time taken to self-archive subsequent articles

The data on the mechanics of self-archiving are also presented graphically in Figures 26 and 27.

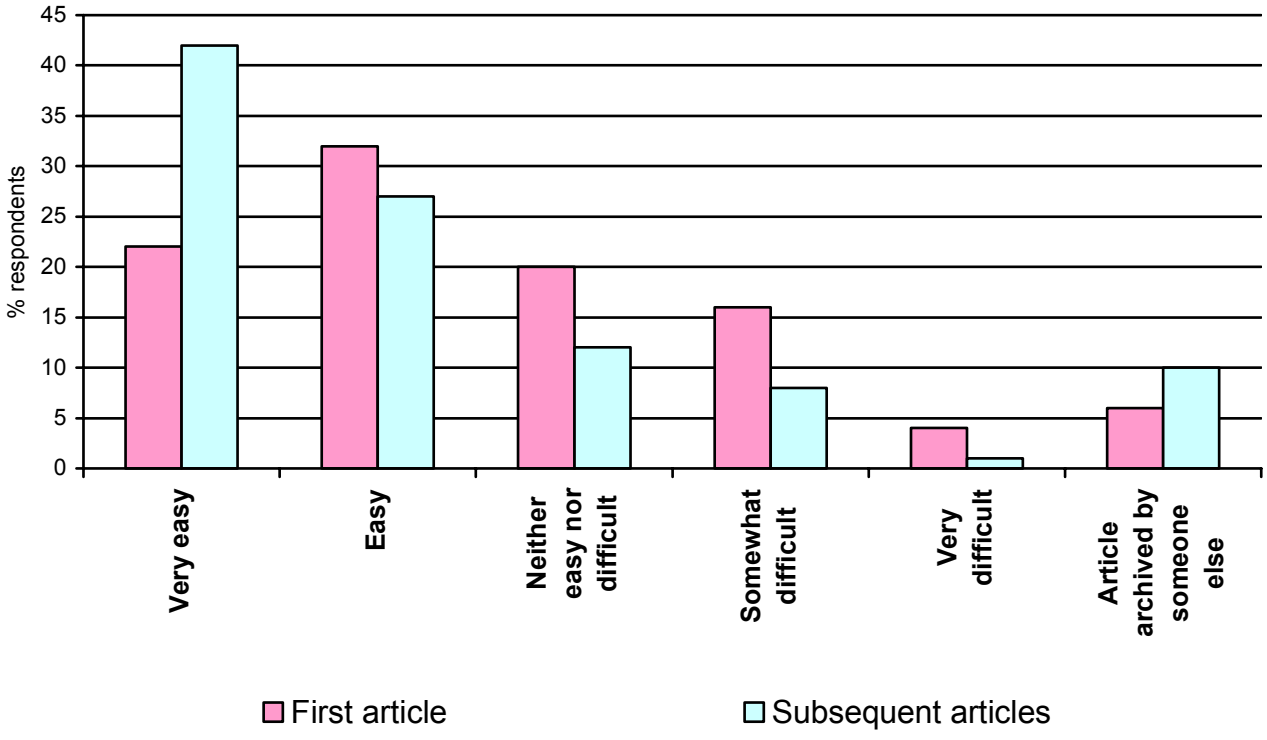

Figure 26: Ease of self-archiving articles

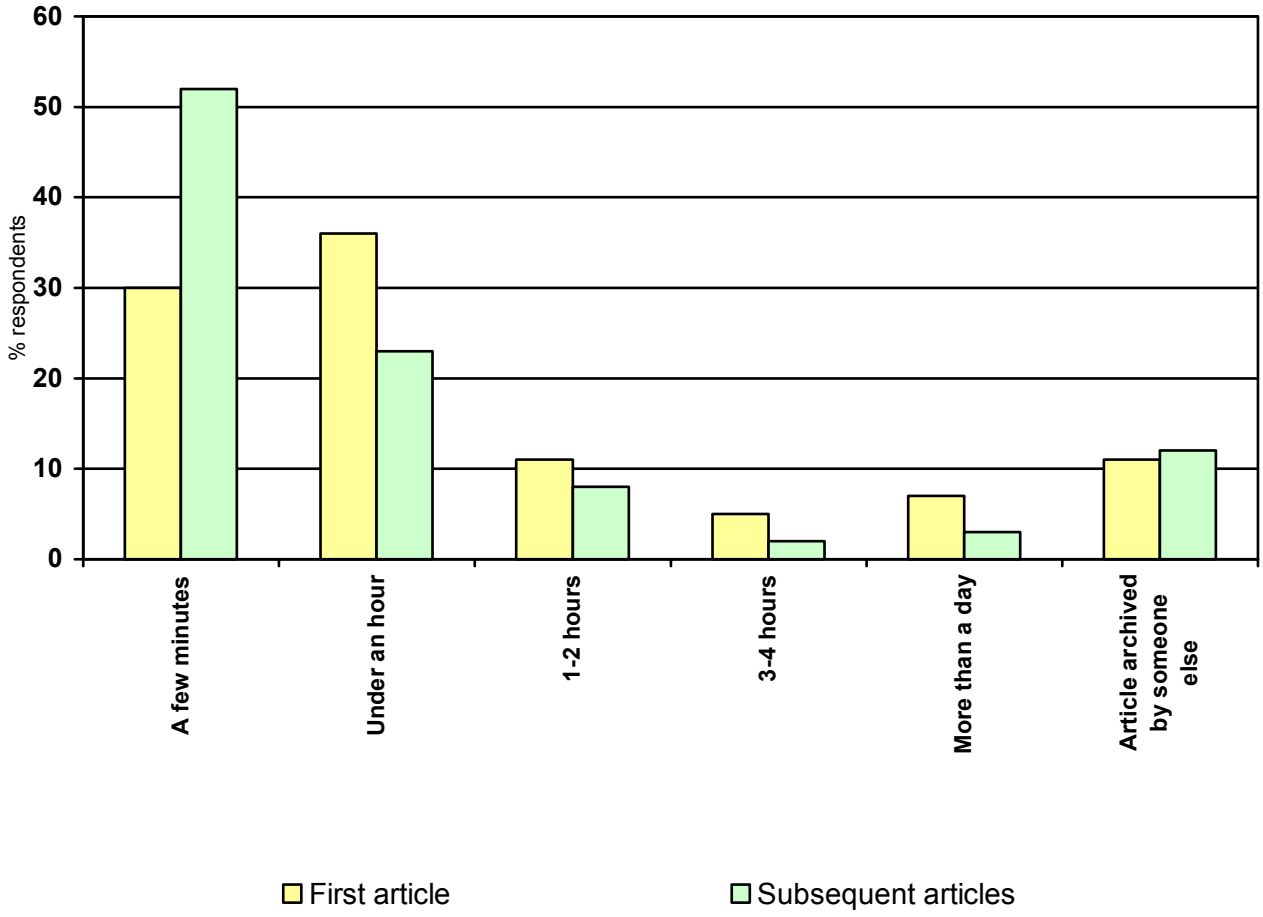

Figure 27: Time taken to self-archive articles

5.4.4 Preservation of archived articles

The issue of preservation of digital articles or research data is under much discussion so it was important to ascertain authors' expectations on this. Almost two thirds (61%) of respondents to this question said that they expected self-archived articles to be stored and made available for open access in perpetuity, technology permitting. Smaller percentages had more modest expectations: 13% expected articles to be stored for at least 10 years, 16% expected this to happen for at least 5 years and 11% expected the period of storage and access to be at least 2 years.

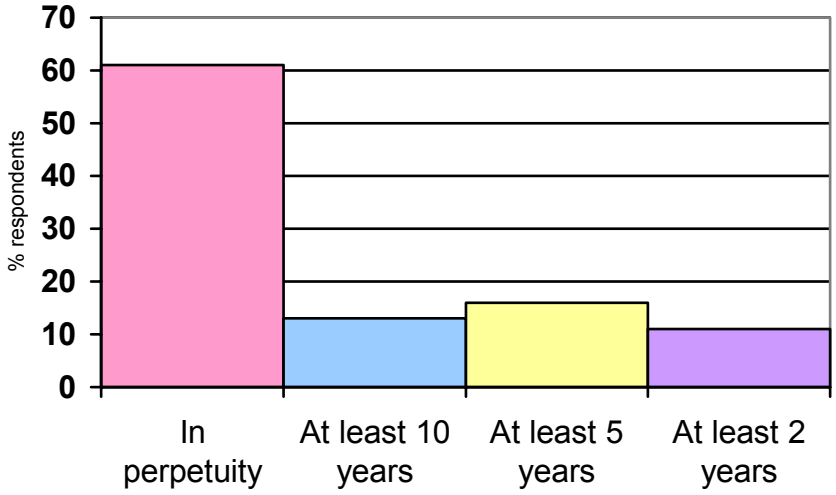

Figure 28: Authors' expectations on period of preservation of self-archived articles

5.4.5 Copyright

Respondents were asked several questions about copyright, which remains one of the reasons authors offer as a deterrent to self-archiving. We wished to explore authors' knowledge of and attitude towards copyright.

Authors were asked who retains the copyright to the last article they self-archived. Over a third (35%) said it was themselves; 37% said it remained with the publisher and 6% that it remained with another party (e.g. their employer). People working in industry or in non-commercial research institutions were most likely to say that copyright remains with their employer. Almost one quarter (22%) don't know who retains the copyright. This is comparable to the result we have obtained before when we have asked this question in previous surveys.

Nearly half of self-archivers (47%) said they were not required to ask permission to self-archive from their publisher. Thirty six percent don't know, but 17% said that permission was required. When asked if they *did* ask permission, 16% said they did but 84% did not. Almost all (95%) of those who said permission wasn't required went ahead and self-archived without asking for it, unsurprisingly. Interestingly, though, the people who said that permission was required did not all ask for it – in fact only just over two thirds of them did, the rest going ahead and self-archiving anyway. A few (7%) of those who didn't know if permission was required or not asked the publisher before self-archiving.

These results are shown in Table 26. All figures in the table are percentages of respondents.

For the last article you self-archived, who retains the copyright	
Myself	35
Journal publisher	37
Another party (e.g. employer)	6
Don't know	22
For the last article you self-archived, were you required by the publisher to ask permission to do so?	
Yes	17
No	47
Don't know	36
For the last article you self-archived, did you ask permission from the publisher?	
Yes	16
No	84

Table 26: Authors' experiences with respect to copyright on their articles

There have been instances of publishers refusing an article submitted by an author who has self-archived the preprint on the grounds of prior publication. Nine percent of the self-archivers here have experienced this. Of these, 16% work in the field of chemistry, and 14% in each of the fields of computer science and business & management. Other subject areas had experienced much lower levels of this.

5.4.6 Digital objects being deposited in open archives

Open access archives are used to store digital objects other than just preprints and postprints. We were interested to find out what other things people were depositing as well as these. The findings are presented in Table 27 and Figure 29.

What types of information have you deposited in the open archive(s) you use?	
Type of digital object deposited	**Percentage of respondents**
Postprint (peer-reviewed research article)	67
Conference paper	47
Preprint (research article before peer review)	41
Technical report	30
Working paper	21
Book chapter	20
Dissertation or thesis	19
Courseware (teaching materials)	19
Discussion paper	12
Software	11
Monograph	9
Manual	8
Video file	4
Audio file	3

Table 27: Digital objects deposited by self-archivers

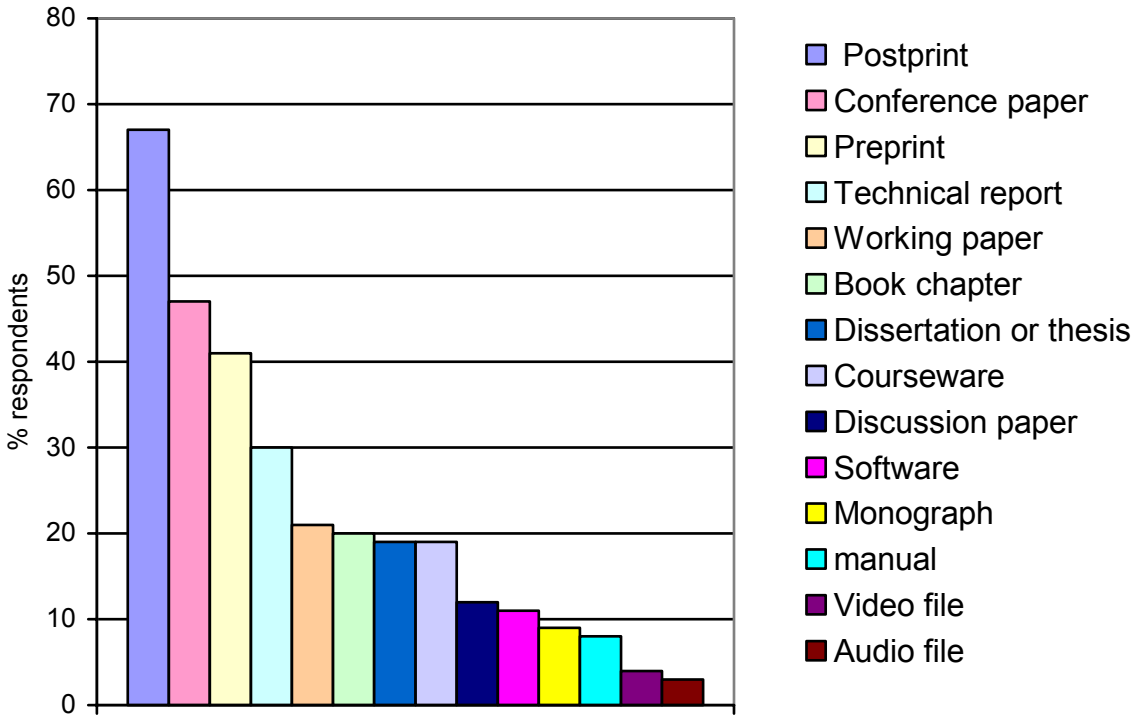

Figure 29: Digital objects deposited by self-archivers

For those interested in seeing how these deposition activities break down by subject area the full results are presented in Table 28. The largest group of people depositing conference papers, for example, is the computer scientists, as might be expected. Postprints are generally deposited more frequently than preprints, although in physics, a community with a long tradition of preprint archiving, these two forms of article are deposited equally frequently; and in computer sciences, another discipline with a long history of self-archiving, postprints are deposited by 42% of researchers, while 32% deposit preprints and 34% deposit technical reports (also a form of preprint in this community).

Subject area	Postprint	Preprint	Technical report	Monograph	Book chapter	Discussion paper	Working paper
Agriculture & food science	27	7	14	4	4	6	7
Business & management	11	17	13	2	4	4	15
Chemistry	23	10	5	1	5	1	2
Computer sciences	42	32	34	3	16	9	15
Earth & geographical sciences	39	9	15	12	3	3	6
Engineering, materials science, technology	32	14	14	4	6	3	7
Humanities	29	17	5	4	12	6	16
Law & politics	24	29	6	6	12	6	24
Library & information science	33	23	22	5	5	16	20
Life sciences	21	7	7	3	6	3	2
Mathematics	26	36	21	4	7	8	8
Medical sciences	15	5	4	2	2	1	1
Physics	32	31	9	1	7	5	5
Psychology	19	12	8	4	8	5	4
Social sciences & education	18	14	8	5	8	10	18

continued….

….continued from previous page

Subject area	Conference paper	Manual	Software	Dissertation or thesis	Audio file	Video file	Courseware
Agriculture & food science	13	4	3	9	1	1	6
Business & management	15	4	2	6	2	2	4
Chemistry	6	1	2	4	0	0	6
Computer sciences	46	7	13	21	1	3	15
Earth & geographical sciences	18	0	0	6	0	0	3
Engineering, materials science, technology	19	2	7	9	1	1	9
Humanities	22	1	4	14	1	0	11
Law & politics	24	6	12	0	0	0	18
Library & information science	27	1	1	7	3	1	8
Life sciences	7	4	4	6	1	1	7
Mathematics	26	7	11	11	0	0	11
Medical sciences	4	1	1	2	0	1	3
Physics	25	1	0	8	0	0	7
Psychology	12	4	5	4	1	3	5
Social sciences & education	17	4	3	3	1	0	6

Table 28: Types of item self-archived, by subject area

(all figures are percentages of respondents in that subject area)

5.4.7 Mandating self-archiving?

The question of whether employers or research funders should go so far as to mandate self-archiving has been the subject of much debate and is discussed fully by Pinfield[27,28]. Some employers, such as Queensland University of Technology in Australia[29], and some research funders (the Wellcome Foundation has announced a mandatory self-archiving policy for its own grant-holders[21]) see the benefit of providing open access by self-archiving to the research carried out under their auspices and have elected to mandate this activity. On the whole, though, employers and research funders have as yet not chosen to go down this path. Only 4% of the self-archivers in this present study say that they are *required* to make their work open access in this way, and 86% of these people are from Southampton University School of Electronics & Computer Science which has had a mandate in place since January 2003.

The reason most often given by employers for not imposing a mandate is that it would engender resentment within the academic community in their institution ("Researchers don't like being told what to do"). When, however, we asked in a previous survey[1,2] how authors would react to being *required* by their employer or research funder to deposit copies of their articles in an open archive the result was that the great majority (67%) would do so willingly, and that most of the rest would do so (albeit reluctantly). These statistics have been quoted often in support of the notion that employers and funders could impose a mandate on self-archiving without jeopardising the goodwill of their research staff. We repeated this exercise in this survey. The results are shown in Table 29 and in Figure 30. The data for three groups of respondents are shown: for the whole respondent population, for those who have self-archived their work already, and for those who have not done this.

If your employer or research funder REQUIRED you to deposit copies of your articles in an open archive, what would be your reaction?	
Total population	**% respondents**
I would comply willingly	81
I would comply reluctantly	13
I would not comply	5
Self-archivers	**% respondents**
I would comply willingly	82
I would comply reluctantly	16
I would not comply	6
Non-self-archivers	**% respondents**
I would comply willingly	78
I would comply reluctantly	15
I would not comply	7

Table 29: Compliance with an employer or funder mandate to self-archive

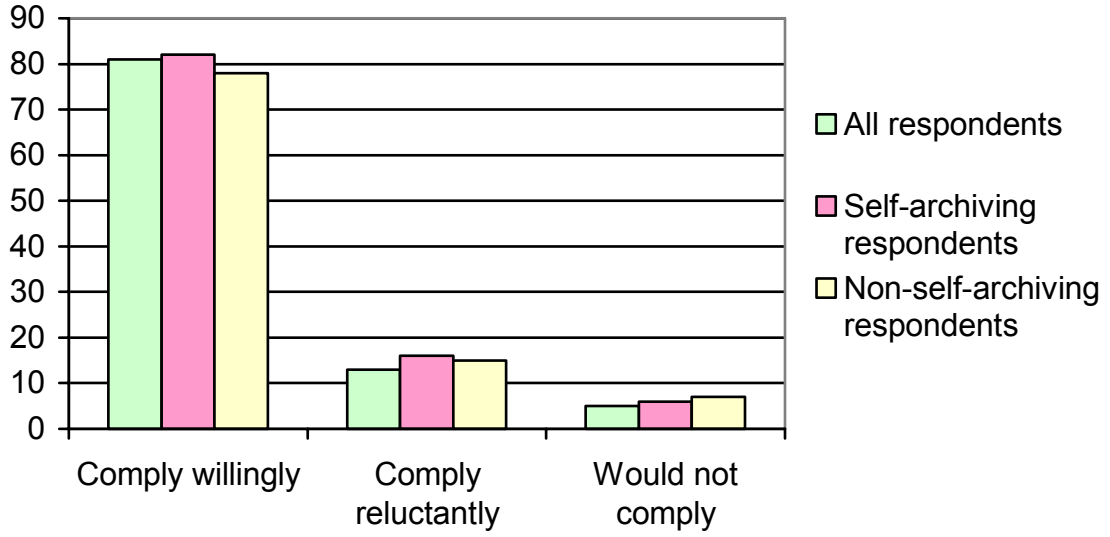

Figure 30: Compliance with an employer or funder mandate to self-archive

The breakdown of responses by subject area is shown in Table 30 and Figure 31.

If your employer or research funder REQUIRED you to deposit copies of your articles in an open archive, what would be your reaction?			
Subject area	I would comply willingly	I would comply reluctantly	I would not comply
Agriculture & food science	72	20	8
Business & management	65	30	5
Chemistry	63	19	19
Computer sciences	89	8	3
Earth & geographical sciences	79	11	11
Engineering, materials science & technology	74	18	8
Humanities	90	8	2
Law & politics	100	0	0
Library & information science	94	4	2
Life sciences	82	14	4
Mathematics	82	7	11
Medical sciences	85	8	8
Physics	84	12	4
Psychology	83	11	4
Social sciences & education	86	9	6
All respondents	**81**	**13**	**5**

Table 30: Compliance with an employer or funder mandate to self-archive by subject area

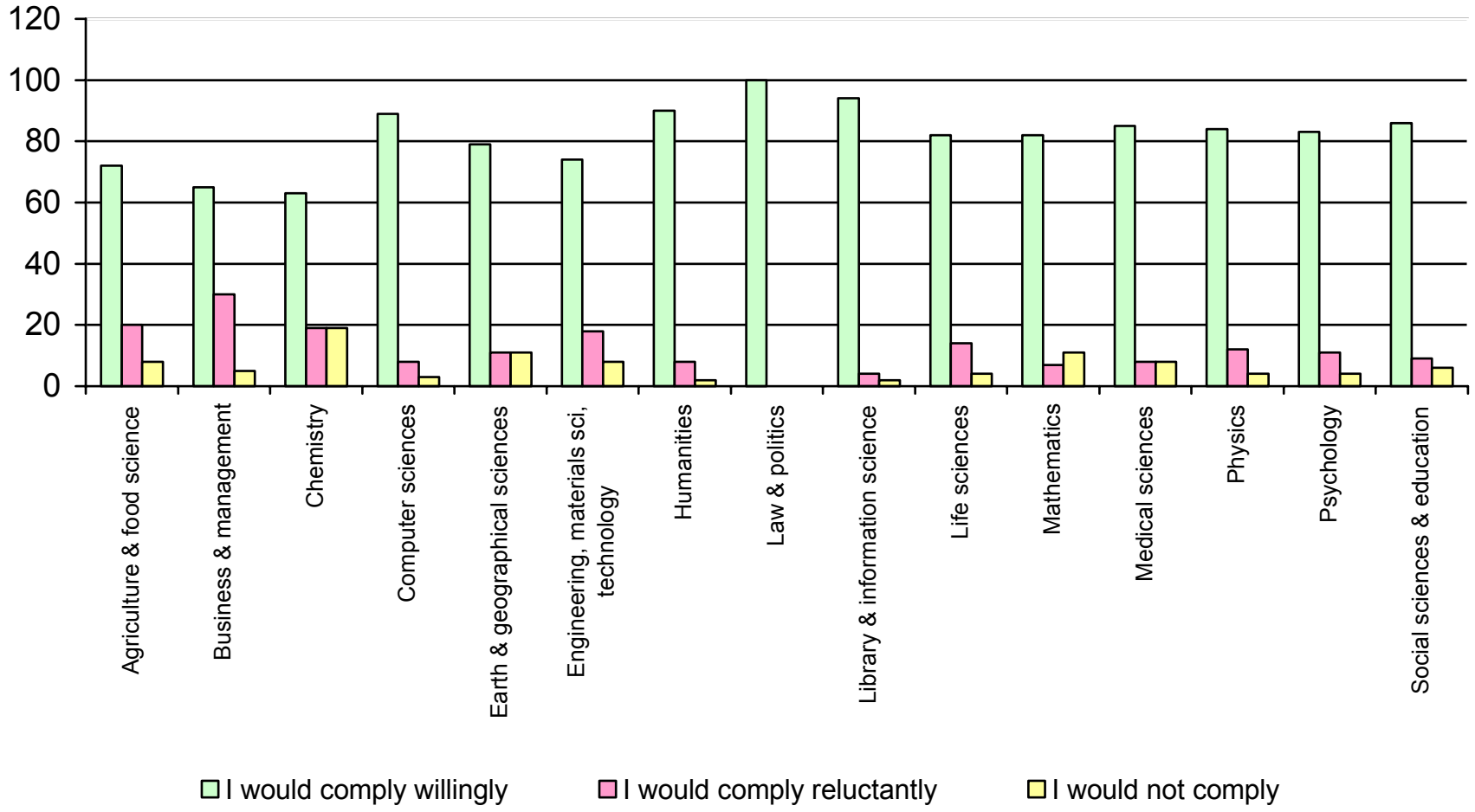

Figure 31: Compliance with an employer or funder mandate to self-archive by subject area

The results for this question are also broken down by geographical origin of respondents, shown in the table below and in the figure following. Figures are percentages of respondents and are rounded.

Geographical origin of respondents	Would comply willingly	Would comply reluctantly	Would not comply
Australia/ New Zealand	86	11	3
Asia (except China and Japan)	86	14	0
China	58	32	11
Japan	75	25	0
Canada	81	10	10
USA	88	11	2
Central or South America	80	10	10
European union (except UK)	82	12	7
Other European countries (except EU/UK)	78	11	11
UK	83	15	2
Middle East	78	11	11
Africa	75	17	8
All respondents	**81**	**13**	**5**

Table 31: Compliance with a self-archiving mandate by geographical origin of respondents

This set of results show that the great majority of authors — in the whole population, in each subject area and in each region of the world — have no objection to a mandate to self-archive imposed by their employer or research funder and would willingly comply with it. Most of the rest would comply reluctantly. Only a very small percentage would not comply at all.

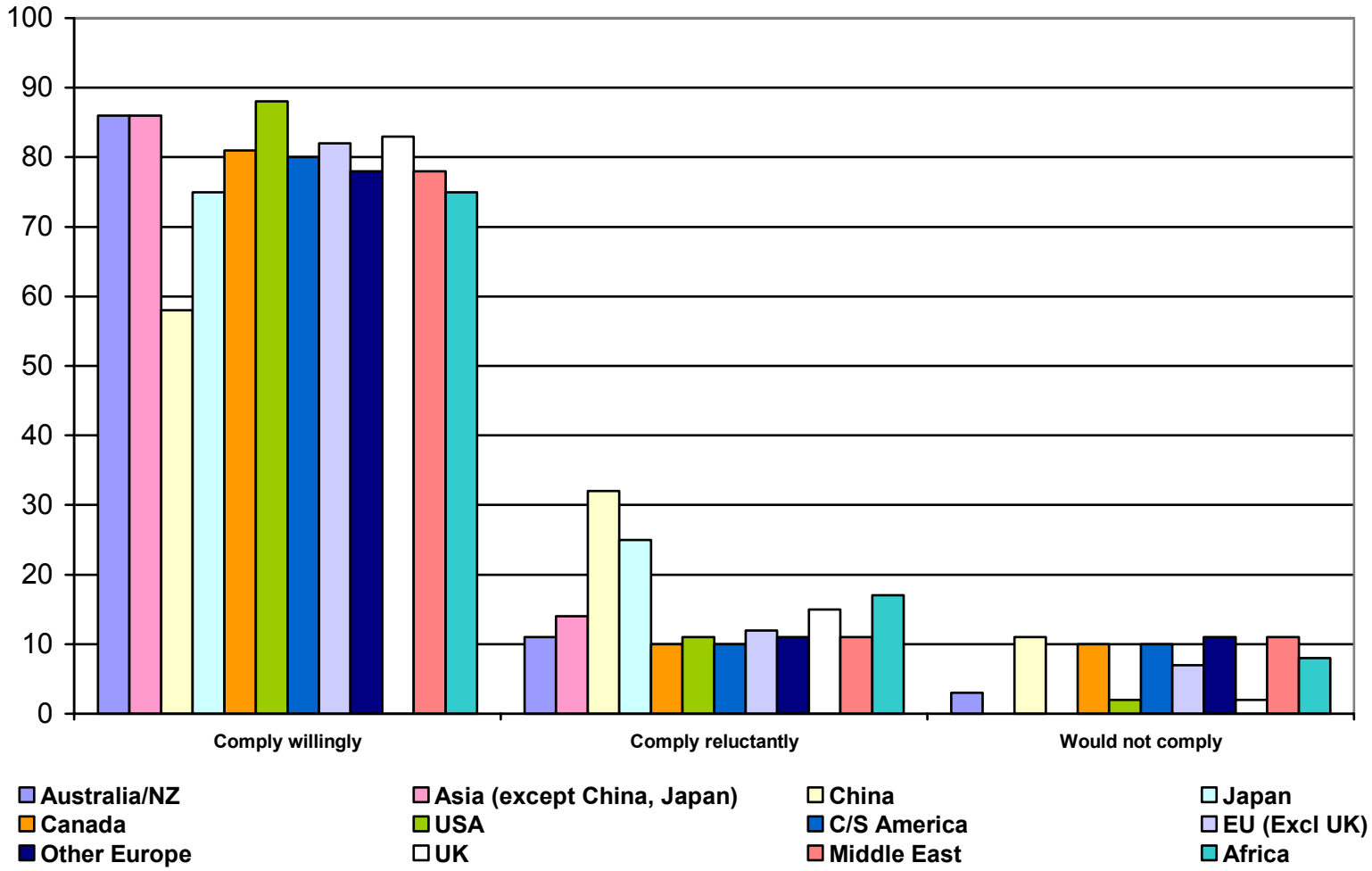

Figure 32: Compliance with employer or funder mandate on self-archiving by geographical origin of respondents (clustered column chart)

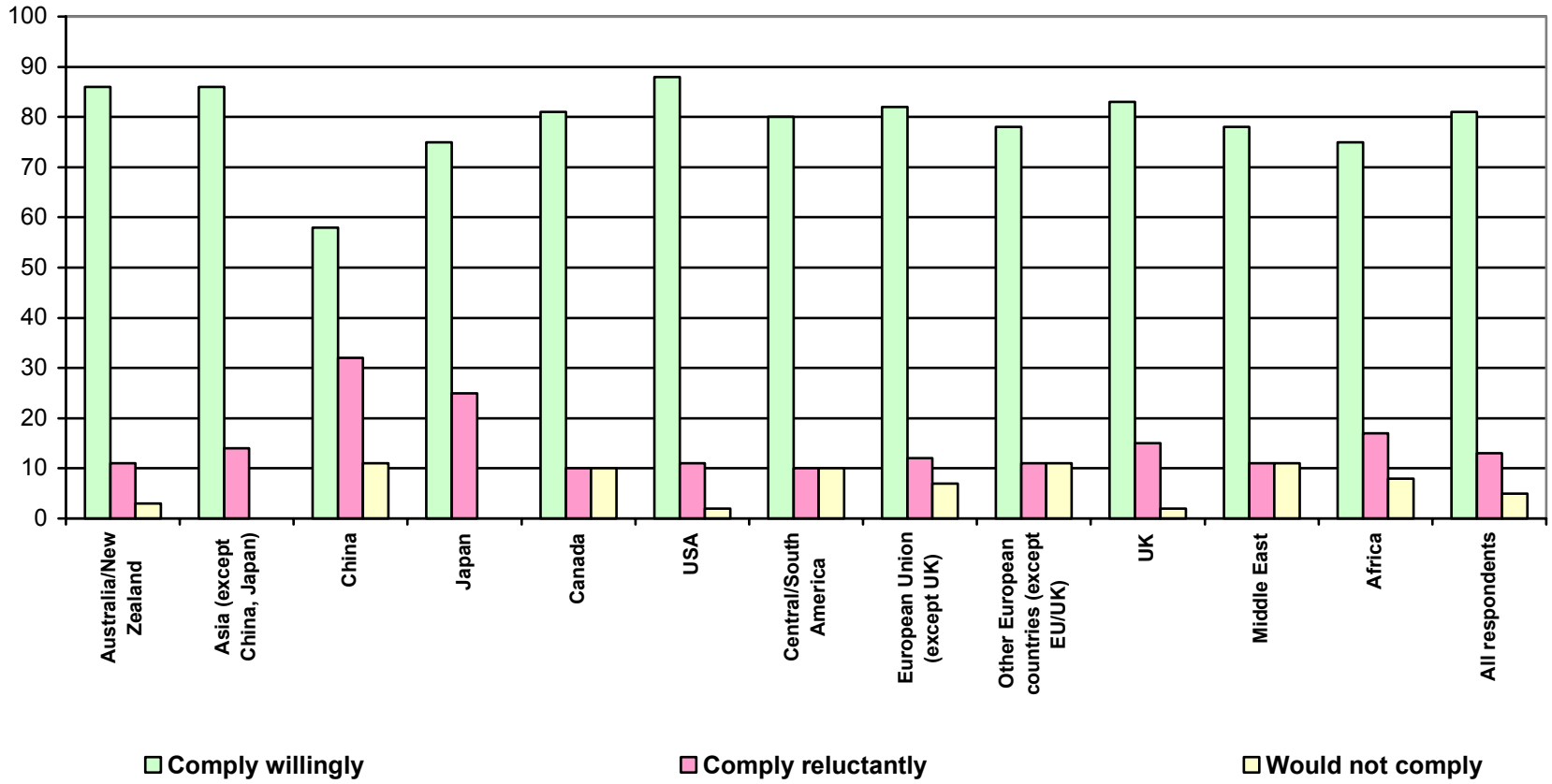

Figure 33: Compliance with employer or funder mandate on self-archiving by geographical origin of respondents (stacked column chart)

6. DISCUSSION

The findings reported here with respect to open access journals held no surprises because they matched those that arose in the previous survey we carried out on this topic[1]. The principle of free access is the strongest imperative for publishing in an open access journal, while authors also perceive the open access journals they have published in to have larger audiences than toll-access journals, to publish more rapidly and to be prestigious in their field. Those authors who have not published in an open access journal cite as the main reasons that they were unfamiliar with open access journals in their field and could not identify a suitable open access journal to publish in. Despite these misgivings, a sizeable proportion of all authors (open access authors or not) intend to publish in an open access journal within the next three years.

Publishing in an open access journal is not, however, the only means for an author to provide open access to his or her work. The alternative means is by self-archiving, whereby an author deposits a copy of each article s/he publishes in an open access archive or repository, or on his or her web page. What the previous survey did not explore in any detail was the experience of authors with respect to self-archiving. Only one question asked about this activity and the result suggested that it was something being carried out by only a small minority of authors. This current survey, reported here, set out to examine further the state of affairs with respect to self-archiving.

To set the self-archiving issues into context, authors' ways of working were explored. The aim was to get a general sense of how they are searching for, retrieving and manipulating research-related information, how they are using electronic information and what their own attitudes are to publishing their work. The main reason for publishing their work is to communicate their research findings to their peers so that they can be used and built upon in future research. In other words authors hope, by publishing the results of their work, to have some impact on how research in their field develops thereafter. It was important to establish this, for one of the planks of open access advocacy is that open access increases an article's citations and therefore impact[13,14,15,16]. If authors did not happen to care much about the impact of their research, this advocacy point would be ineffectual.

In fact, authors care very much about the impact of their own research, as the results here show. Not that they are particularly well-informed on how their research measures up in terms of personal citation scores. In some disciplines, scholars are better informed about this than in others but even so there is an overall poor level of awareness, with over half the population not knowing the

average citation rate for their own published articles. This can be expected to change as metrics used for measuring a scholar's impact shift from that of journal impact factor – never a good proxy measure of an individual's contribution to a field – to citation rates and other personalised measures for individual scholars. Open access will be important in this respect. It is well-documented that there is a direct relationship between article downloads and eventual citations[30,31], and thus the more an author can get his or her articles seen and downloaded the greater the number of citations that can be expected to result. Tools for tracking and measuring citations are increasing, and as authors become more aware of the existence of these, and how to ascertain their own citation impact, so the imperative to raise the visibility of their work will grow.

A lack of awareness is also seen with respect to open access-related issues generally, as has been shown in previous studies. For example, while almost every researcher (98%) uses bibliographic databases to search for information, and a very high proportion (86%) uses services like ScienceDirect or other publisher websites to search out full-text documents, only a very small number of people in comparison (30%) use the OAI search services to look for material in open access archives. This is hardly to be wondered at, because current open access archives contain so little content. Once the volume of research deposited in these archives increases, usage will rise. Also, since nearly three-quarters of researchers use Google to search the web for articles we can assume that the newly-launched Google Scholar service will redress this to some extent. Google Scholar crawls web pages and OAI-compliant open access archives specifically indexing academic material and so will return to searchers articles from open access archives that they may not have explored before out of ignorance of their existence.

There is one other issue to do with scholars' use of research information that has some bearing on open access and that is how far back the literature that they use commonly goes. There is substantial variation between disciplines here, as would be expected, with scholars in the fastest-moving ones mostly using material published in recent years; in contrast, in other disciplines, notably the humanities, material fifty or more years old is commonly consulted. Whilst the growing adoption of the self-archiving habit will account for the provision of current and recent literature on an open access basis, there will be much more of a problem freeing up old or ancient literature. Even though it is now commonplace for publishers to provide unrestricted electronic access to material in their backfiles (usually over 12 months old), it is still uncommon, though there are notable exceptions, especially in the sciences, to find a publisher who has digitised material more than ten years old. Moreover, there is at least one example of a research establishment which is attempting to digitise (scan) articles by its own authors — in this case articles that were published several decades

ago — so that it can deposit them in its own open access archive, and cannot get permission from one of the major publishers to do so.

Now we come to self-archiving activity amongst scholars. Fewer than half of authors have self-archived an article, even on their own website, the method which is the most popular way of making an article visible to all web users. This is, however, significantly higher than for the sample of authors used for our last survey in January 2004, when only 23% of authors had self-archived any work.

More people deposit postprints than preprints, except for the physics and computer science communities as already discussed (section 5.4.6). Some disciplines are particularly active in self-archiving, computer science and medicine showing the most pronounced activity, and the location of self-archived articles also varies with discipline, chemists using institutional repositories more than their websites, for example, while for other subjects this is reversed. Subject-based archives are used by the fewest people, but this is not surprising given the small number of subject-based archives in existence. It simply has not been an option for most scholars as yet. As distributed, institutional, archives accumulate content then subject-based, harvesting archives will increase and usage will grow.

The more prolific an author — that is, the more articles s/he publishes — the more likely they are to self-archive their work on websites or in institutional repositories. It is likely, therefore, that as greater numbers of the most productive authors become aware of self-archiving the number of articles in open access repositories will rise quite steeply.

The caveat here is that issue of awareness. Awareness of self-archiving amongst those who have not carried out this activity remains low, though scholars in the disciplines of library and information science, computer science, physics and mathematics are better informed than those in other subjects. But there are still many scholars who remain unaware of self-archiving and still others who, though aware, have not elected to undertake the activity, at least so far.

What clues are there in this study as to how open access advocates might most effectively go about changing these things, informing the ignorant and persuading the reluctant?

First, it has been shown here that the great majority of people who already self-archive have been self-motivated to do so: they have learned about the practice through their peers or by following the open access debate and have decided that it is a beneficial thing to do. Redoubling efforts to advocate self-archiving to the uninitiated would therefore be likely to bring more converts to the fold. What

messages can be delivered to them that would help? Presenting the case — and the evidence — on how open access increases downloads, citations and therefore the impact of an author's work is a most effective tactic. This is what has persuaded many of the people who have published in open access journals to do so; it is also one of the reasons cited by self-archivers for their adoption of the practice. All this is exactly as would be expected given that the main reason a researcher publishes his or her work is to inform others in their field of the results – to make some impact on future endeavour in that field.

Putting reasons for self-archiving before authors is one approach. It can be supported by other things that help to increase incentive. The practice of depositing an article in a repository itself, for example, may appear rather daunting to some. The data in this study show that the first article takes only a relatively short time to deposit, and that subsequent articles can be deposited in a few minutes. These are the experiences of the authors surveyed here. A study of the computer logs at the School of Electronics & Computer Science at Southampton University reveals the same thing, but with more precision: this real-life study shows that it averages around ten minutes for an article to be deposited in that School's institutional archive. In other cases, archive administrators have successfully set the ball rolling for a new archive by collecting and depositing articles from authors in that institution themselves, an alternative way to get an archive populated efficiently if the resources are available.

Finally, there is the question of insisting that authors self-archive their work. There has been a lot of debate on this issue but it may be informative to rehearse the essence of it. Some people say that authors cannot be *required* to do anything by their employer, especially those in academic institutions who are held to be such independent thinkers that they would rebel against any such requirement and institution-researcher relationships would deteriorate, perhaps irretrievably. Others hold the opposite view and argue that researchers should be under a mandate to deposit articles in their open access institutional repositories. They say that researchers are already implicitly required to publish in order to obtain tenure, advance their career, and hold onto their jobs. And certainly where funders are concerned there is always an explicit requirement for the recipient of a grant to produce a final report at the end of the funding period on the work that has been carried out. In most cases there is an expectation on the part of the funder that the grantholder also publishes the work in a recognised scholarly forum, usually a learned journal. Additionally, very often the journal's quality and average citation impact, as well as the article's own specific citation count, are weighed in evaluating and rewarding researchers' performance, rather than merely counting the number of journal articles published.

The evidence from this study (which corroborates our previous study that asked the same question) indicates that a mandate from an institutional employer or a research funder to self-archive would meet with very little resentment and even less resistance from researchers. Already, several institutions have decided this for themselves, institutions such as Queensland University of Technology in Australia, CERN in Switzerland and the University of Minho in Portugal[29]. So have funders, the Wellcome Foundation being one example already mentioned here. There are more, and there will be more to come. Probably, in the end, the usual recipe combining incentive and insistence will prevail as the most successful approach: if you are going to tell someone to do something it makes sense to start by explaining what benefits it will bring them at the same time as you finger the handle of your baton.

References

1. Swan, Alma and Brown, Sheridan (2004) Report of the JISC/OSI open access journal authors survey. pp 1-76.
http://www.jisc.ac.uk/uploaded_documents/JISCOAreport1.pdf

2. Swan, Alma and Brown, Sheridan (2004) Authors and open access publishing. *Learned Publishing*, **17 (3)**, 219-224.
http://lysander.ingentaselect.com/vl=15729124/cl=20/nw=1/rpsv/cgi-bin/linker?ini=alpsp&reqidx=/cw/alpsp/09531513/v17n3/s7/p219

3. www.citeseer.ist.psu.edu

4. www.arxiv.org

5. Harnad, Stevan (1995) A Subversive Proposal. In: Ann Okerson & James O'Donnell (Eds.) Scholarly Journals at the Crossroads; A Subversive Proposal for Electronic Publishing. Washington, DC., Association of Research Libraries, June 1995. http://www.arl.org/scomm/subversive/toc.html
(originally posted June 27 1994:
http://www.arl.org/scomm/subversive/sub01.html)

6. Harnad, S (1999) Free at last: the future of peer-reviewed journals. *D-Lib Magazine*, **5**, 12. http://www.dlib.org/dlib/december99/12harnad.html

7. Self archiving FAQ. http://www.eprints.org/self-faq/

8. American Scientist Open Access Forum.
http://www.cogsci.soton.ac.uk/~harnad/Hypermail/Amsci/index.html

9. SHERPA: Publisher copyright policies and self-archiving.
http://www.sherpa.ac.uk/romeo.php

10. Eprints.org: Journal self-archiving policies.
http://romeo.eprints.org/stats.php

11. Swan A, Needham P, Probets P, Muir A, O'Brien A, Oppenheim C, Hardy R and Rowland F (2004). Delivery, management and access model for E-prints and open access journals within further and higher education (Report of a JISC study). pp 1-121.
http://www.jisc.ac.uk/uploaded_documents/ACF1E88.pdf

12. Alma Swan, Paul Needham, Steve Probets, Adrienne Muir, Anne O'Brien, Charles Oppenheim, Rachel Hardy, Fytton Rowland and Sheridan Brown (2005). Developing a model for e-prints and open access journal content for UK higher and further education. *Learned Publishing*, **18 (1)**, 25-40. www.keyperspectives.co.uk/OpenAccessArchive/Eprints_LP_paper.pdf

13. Lawrence, S (2001) Online or invisible? *Nature* **411**, 6837, p521 http://www.neci.nec.com/~lawrence/papers/online-nature01/ or www.nature.com/nature/debates/e-access/Articles/lawrence.html

14. Kurtz, M (2004) Restrictive access policies cut readership of electronic research journal articles by a factor of two. http://opcit.eprints.org/feb19oa/kurtz.pdf

15. Harnad, S and Brody, T (2004) Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10 (6), (www.dlib.org/dlib/june04/harnad/06harnad.html).

16. Antelman, K (2005) Do open-access articles have a greater research impact? *College & Research Libraries*, **65 (1)**, 372-282.

17. Swan, A and Brown, S (2002) Authors and Electronic Publishing: The ALPSP research study on authors' and readers' views of electronic research communication. pp 1-76. ALPSP, Worthing.

18. Swan, A and Brown, S (2003) Authors and electronic publishing: what authors want from the new technology. *Learned Publishing*, **16 (1)**, 28-33. http://lysander.ingentaselect.com/vl=15729124/cl=20/nw=1/fm=docpdf/rpsv/cw/alpsp/09531513/v16n1/s6/p28

19. www.cogprints.soton.ac.uk

20. www.repec.org

21. Wellcome Trust and National Library of Medicine in talks for worldwide open access archive (press release). http://www.wellcome.ac.uk/doc_WTX022826.html

22. Rowlands, Ian, Nicholas, Dave and Huntingdon, Paul (2004). Scholarly communication in the digital environment: What do authors want? Findings of an international survey of author opinion: project report. Centre for Information Behaviour and the Evaluation of Research, City University, London, UK.

23. Publisher copyright policies and self-archiving. http://www.sherpa.ac.uk/romeo.php

24. UK House of Commons Science and Technology Select Committee: Tenth Report. Scientific publications: Free for all? http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39902.htm

25. US Government House Appropriations Bill HR 5006 recommendations: http://thomas.loc.gov/cgi-bin/cpquery/?&db_id=cp108&r_n=hr636.108&sel=TOC_338641&

26. Carr, L and Harnad, S (2005) Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving. http://eprints.ecs.soton.ac.uk/10688/

27. Pinfield, S (2005). A mandate to self archive? The role of open access institutional repositories. *Serials*, **18 (1)**, 30-34.

28. Pinfield, S (2004) Self-archiving publications. In: Gorman, G E and Rowland, F (eds). *International yearbook of Library and Information Management 2004-2005: Scholarly publishing in an electronic era*. London: Facet. Pp118-145. Available at http://eprints.nottingham.ac.uk/archive/00000142/

29. Registry of Institutional Self Archiving Policies. http://www.eprints.org/signup/fulllist.php

30. Perneger, T V (2004) Relation between online 'hit counts' and subsequent citations: prospective study of research papers in the BMJ. *BMJ* **329**, 546-7.

31. Brody, T and Harnad S (2005) Early web usage statistics as predictors of later citation impact. In press (Journal of the American Society for Information Science & Technology) http://eprints.ecs.soton.ac.uk/10712/

Appendix 1:

Reasons for publishing in open access journals: results broken down by subject area

Reason for publishing in an open access journal	Agriculture & food science	Business & management	Chemistry	Computer sciences	Earth & geographical sciences
The principle of free access for all readers	75	95	95	73	75
I perceive the readership to be larger than for a subscription-based journal	56	21	21	56	63
I perceive OA journals to have faster publication times than other types of journal	38	32	32	58	38
The OA journal(s) I have published in are prestigious in my field	31	37	37	48	38
I think my article will be more frequently cited	44	21	21	50	38
The OA journal(s) I have published in have a high impact in my field	44	32	32	35	13
I was attracted to the editor / editorial board	13	21	21	33	25
I am concerned about the cost to my institution of non OA journals	13	11	11	23	25
I object to publishing with a non-OA commercial publisher	6	5	5	19	38
My decision to publish in an OA journal was influenced by my co-publishing colleagues	13	21	21	13	13
The OA journal(s) I have published in are published from my own institution	13	5	5	2	13
My decision to publish in an OA journal was influenced by my institution	6	0	0	2	0
My decision to publish in an OA journal was influenced by my grant-awarding body	0	0	0	0	0

continued overleaf….

… continued from previous page

Reason for publishing in an open access journal	Engineering, materials science, technology	Humanities	Law & politics	Library & information science	Life sciences
The principle of free access for all readers	65	92	77	77	77
I perceive the readership to be larger than for a subscription-based journal	26	54	55	55	40
I perceive OA journals to have faster publication times than other types of journal	30	63	33	46	46
The OA journal(s) I have published in are prestigious in my field	44	33	50	55	49
I think my article will be more frequently cited	30	38	33	32	32
The OA journal(s) I have published in have a high impact in my field	17	13	33	36	35
I was attracted to the editor / editorial board	9	25	17	18	23
I am concerned about the cost to my institution of non OA journals	13	21	50	18	16
I object to publishing with a non-OA commercial publisher	9	17	17	27	14
My decision to publish in an OA journal was influenced by my co-publishing colleagues	9	4	0	9	16
The OA journal(s) I have published in are published from my own institution	4	17	17	9	2
My decision to publish in an OA journal was influenced by my institution	9	13	0	5	2
My decision to publish in an OA journal was influenced by my grant-awarding body	13	0	0	0	0

continued overleaf….

… continued from previous page

Reason for publishing in an open access journal	Mathematics	Medical sciences	Physics	Psychology	Social sciences & education
The principle of free access for all readers	70	80	76	76	77
I perceive the readership to be larger than for a subscription-based journal	50	51	33	44	59
I perceive OA journals to have faster publication times than other types of journal	55	37	38	60	54
The OA journal(s) I have published in are prestigious in my field	45	29	38	36	36
I think my article will be more frequently cited	40	35	29	32	41
The OA journal(s) I have published in have a high impact in my field	20	22	29	20	28
I was attracted to the editor / editorial board	40	18	14	24	28
I am concerned about the cost to my institution of non OA journals	20	10	5	16	13
I object to publishing with a non-OA commercial publisher	20	10	19	4	15
My decision to publish in an OA journal was influenced by my co-publishing colleagues	10	12	24	12	5
The OA journal(s) I have published in are published from my own institution	0	2	10	0	5
My decision to publish in an OA journal was influenced by my institution	0	0	5	4	0
My decision to publish in an OA journal was influenced by my grant-awarding body	0	0	0	4	0

Appendix 2:

Reasons for publishing in open access journals

Verbatim responses from respondents: responses are categorised under topic headings:

Article was invited

- The manuscript was invited
- Invited paper
- It was an article requested by the editor
- I was invited by the editor to submit a commentary.
- I was commissioned to write an article.
- I was invited o submit an article.
- I was invited to and the journal is an accredited one.
- The editor asked me to contribute.
- Invited by editor
- I was invited to write a review for an OA journal (the paper is still under review). The impact factor was good, so I was happy to oblige. The staff I worked with were very professional.

To support open access/the principle of free access

- To support OA journals and conferences.
- I wish to promote OA
- I am currently preparing a research paper that will be submitted to an open access journal. I am fully committed to the principle of free access to all recorded knowledge. I also wish to make my work readily available to all readers.
- I have been thanked by many researchers and educators for publishing my work in OA journals and therefore making it accessible to them and to their students.
- I wish to promote the concept of open access publishing, and also contribute to improving the impact of it by sending the best of my work to such journals.
- Electronic open access journals can be better for dissemination, particularly across disciplines, and for student access. One article I have published has over 10,000 hits on it. I consider it unlikely that a print journal would have anything approaching that number of cursory glances.
- I also wanted to try an Open Access Journal for the first time.
- I have published an e-book through Tufts University. I personally paid all repro costs for the illustrations, and the university paid a web master to put it up and maintain it. I did it because I want my research to be available world-wide free of charge.

Particular advantages

- The most important reason was: I perceive OA journals to have faster publication times than other types of journal
- Because I am an independent scholar outside academia now, I ran head on into difficulties in completing a literature review adequate to the peer reviewers of the collection I eventually was published in (I did not pay anything... this is an electronic publication of quite high stature, using a Creative Commons copyright). The mode of publication was my primary reason for submitting first in this route. And my work is being cited and linked to widely.

- Speed of peer review process and feedback
- They are easier to get papers accepted in.
- I wanted to be able to freely distribute software choose journals that had on-line archives tied to the articles that I publish.
- Rapid publication of lower impact results that nonetheless provide important information
- I have published articles with audiofiles and visual images in the article. By publishing on the Internet, I have the audiofiles accessible by readers of the article.
- They are easier to get papers accepted in.
- Speed of publication
- It was an online Journal which allows free colour and electronic material attachment
- I wanted to be able to freely distribute software choose journals that had on-line archives tied to the articles that I publish.

Copyright retention

- Retention of copyright, non-blind reviewers, prepublication record available
- Allows copyrights to be retained by authors
- My personal financial reward for publishing comes from citations and reputation. Restricted, copy-right nervous, publishing may serve the publisher's interest but it does not serve the scholar's interests.

The open access journal chosen was the most suitable vehicle

- I publish in journals because I am interested in something and write about it after researching. OA is important, inevitable, but at the time, not a deciding factor.
- I discovered it was OA after submitting to it. But the principal of free access is important to me.
- The particular journal I submitted to was a good fit for this paper.
- I published in this journal because its focus matched the focus of my paper. Open access was a bonus.
- The journal is respected in my field, and had a call for articles in my area of expertise.
- The topic of my article was best suited to the subject title of the selected OA journal
- The journal is more focussed on my area of research
- The OA journal happened to be an appropriate publication outlet for the material I wanted to publish. I thought my article would be more visible in there.
- Random submission, did not know it was an OA journal, and do not worry about the fact that it ended up in an OA journal.

Able to publish in an OA journal for less cost than normal

- No additional cost to include color prints, which were essential for the paper.
- Lower cost of publication

Object to commercial publishers/serials crisis

- I was formerly executive editor of a Kluwer journal, and the high price of library and individual subscriptions led myself and some of my colleagues to start a competing OA journal, which now has the highest impact factor in my subfield.
- I think most of researchers are financed by academic and research institutions, so the commercial publishers are abusing or taking advantage of these institutions, therefore OA journal is the solution against monopoly of publishers. I know that commercial publishers need money to improve their services, but not being exclusive in the investment of the scientific and academic research.

- I feel that OA is the way forward and that commercial publishers charge too much for journals. This means that libraries have to reduce the number of journals every year.

Appendix 3:

Reasons for not publishing in open access journals: results broken down by subject area

Reason for not publishing in an open access journal	Agriculture & food science	Business & management	Chemistry	Computer sciences	Earth & geographical sciences
I am not familiar enough with OA journals in my field to feel confident about submitting my work	36	30	34	29	49
I could not identify any OA journals to publish in	20	30	20	29	30
I perceive the OA journals in my field to have low prestige	11	26	21	13	21
I perceive the OA journals in my field to have low impact	10	15	21	10	21
I cannot find the funds to pay any publication fees required by OA journals	16	15	20	9	15
I object in principle to paying a publication fee to OA journals that charge one	11	19	11	11	21
I perceive the readership to be smaller than for a subscription-based journal	16	13	15	8	12
I always publish my work in the same journals and am satisfied with this way of working	16	9	13	5	12
I am concerned about the archiving of work published in OA journals	10	6	12	5	18
I perceive the OA journals in my field to have poor peer review procedures in place	4	11	9	2	12
My decision was influenced by my co-publishing colleagues	7	6	2	5	0
My decision was influenced by my institution	1	6	5	4	0
My decision was influenced by my grant-awarding body	0	0	7	3	0
I perceive the OA journals in my field to have slower publication times than traditional journals	1	4	1	0	0

continued overleaf….

… continued from previous page

Reason for publishing in an open access journal	Engineering, materials science, technology	Humanities	Law & politics	Library & information science	Life sciences
I am not familiar enough with OA journals in my field to feel confident about submitting my work	30	33	35	14	39
I could not identify any OA journals to publish in	27	23	18	18	22
I perceive the OA journals in my field to have low prestige	13	10	24	7	14
I perceive the OA journals in my field to have low impact	12	5	6	8	18
I cannot find the funds to pay any publication fees required by OA journals	19	10	18	17	15
I object in principle to paying a publication fee to OA journals that charge one	12	13	24	10	8
I perceive the readership to be smaller than for a subscription-based journal	9	4	0	0	13
I always publish my work in the same journals and am satisfied with this way of working	9	3	0	5	9
I am concerned about the archiving of work published in OA journals	6	10	6	5	7
I perceive the OA journals in my field to have poor peer review procedures in place	5	7	12	3	5
My decision was influenced by my co-publishing colleagues	4	1	0	3	7
My decision was influenced by my institution	4	1	0	5	3
My decision was influenced by my grant-awarding body	3	1	0	1	1
I perceive the OA journals in my field to have slower publication times than traditional journals	1	0	0	1	1

continued overleaf….

…continued from previous page

Reason for publishing in an open access journal	Mathematics	Medical sciences	Physics	Psychology	Social sciences & education
I am not familiar enough with OA journals in my field to feel confident about submitting my work	24	36	33	41	37
I could not identify any OA journals to publish in	22	15	27	23	22
I perceive the OA journals in my field to have low prestige	14	19	13	23	19
I perceive the OA journals in my field to have low impact	10	24	14	15	13
I cannot find the funds to pay any publication fees required by OA journals	14	14	21	12	12
I object in principle to paying a publication fee to OA journals that charge one	14	12	12	12	20
I perceive the readership to be smaller than for a subscription-based journal	7	15	11	14	8
I always publish my work in the same journals and am satisfied with this way of working	7	8	19	5	7
I am concerned about the archiving of work published in OA journals	7	6	4	8	8
I perceive the OA journals in my field to have poor peer review procedures in place	1	7	5	10	6
My decision was influenced by my co-publishing colleagues	7	5	5	5	5
My decision was influenced by my institution	3	6	4	5	1
My decision was influenced by my grant-awarding body	0	2	1	1	2
I perceive the OA journals in my field to have slower publication times than traditional journals	0	0	0	0	0

Reasons for publishing in an open access journal, by subject area

(all figures are percentages of respondents in that subject area)

Appendix 4:

Reasons for not publishing in open access journals

Verbatim comments by respondents, sorted and grouped under a number of main topic headings:

Unaware of open access

- I have never heard of Open Access Journals
- I would be pleased to get some hints were I could publish OA articles in the field of work and organizational psychology
- I know nothing about the OA project or your journals [sic]. This is the first I have heard of them.
- This is the first I've heard of this kind of journal, but I'm not favorably disposed toward it.
- Not aware of it
- Could you better define me the term "open access journal" ? Is it a Journal available in classical paper form plus paid reading it via internet? Thank you.
- I did not have informations about journal
- I do not even know if they exist in my field.
- I did not know that such journals existed. However, many of the above factors would probably have been taken into account if I were to consider such a journal.
- I simply don't know anything about OA journals.
- Knew nothing about them - but in any case I have no access to funds that would pay for any printing

Intend to publish in OA journals in future

- We are working on an article now to be published by you [sic].
- My next paper is not yet ready
- I am relatively new to the world of publishing, but have done a lot of reading about OA and plan to submit to OA journals in the future, perhaps exclusively.
- In the future I will publish in the OA journal. Actually I am end user for OA journals.
- I am a new academic and currently completing articles I plan to submit to OA journals...I have not done so yet, because they are not finished.
- I did not know about the Open Access initiative until 8 months ago. Now that I know about OA, I will try to publish in OA journals whenever I can. My 2 most recent submissions still were not to OA journals. One was because I submitted an article much related to a previous article, and I wanted them in the same journal. The other was because we found a non-OA journal the scope and readership of which matched what we were looking for (we could not find an OA journal that fit as well).
- Haven't got around to it yet, because no journal I know of has matched a paper I have written. But I intend to do so soon, because a paper I'm now working on will fit well with the subject matter of a particular online journal.

Promotion/tenure issues

- I have only just started publishing, so haven't been exposed to much to this kind of decision. The only thing that would prevent me from publishing in this kind of journal would be its lack of accreditation by my university. Only articles published in accredited journals receive government subsidy and "count" towards promotion. In principle, I support open access publication

- OA journals are not recognised by the South African government for research funding subsidies.
- OA journals are perceived poorly by the tenure review committee
- My university does not recommend to publish in OA-journals (Actually my supervisor forbade me). His view is that they are not "seriously taken" journals....if we start to publish in OA-journals commercial journals will refuse to accept our articles in future.
- I think job committees do not value publications in OA journals as highly as those in traditional print journals. I do not subscribe to this point of view myself, but I think this is still a common prejudice.
- My institution uses how many papers you have published in top tier journals as one of the criteria for getting tenure and do not rank the current open access journal as compared to Science, Nature or Cell.
- I am certain that paying to be published would not be well received in tenure and promotion decisions.
- I don't know how the promotion and tenure committee values publication in OA journals. My future decisions about where to submit a publication depend upon what the Departmental Appointment and Promotion committee judges on OA journal publications.

Impact, funding, RAE issues

- I am searching for a job, and the OA journals seem to have less "impact", or a lower profile, than non-OA journals. This especially seems to be true for the industry jobs I am applying for. Most companies (even law firms) have heard of Science, Nature, Cell, but they have not heard of PLoS, etc.
- If a journal is not ISI listed, there is no incentive at my University and the main funding body of South Africa to publish in such a journal.
- I am unsure if OA journals are recognised by the government body(or bodies) that fund/s the university. In Australia it is known as DEST (Department of Education, Science & Training). It is the body that rewards research quantum in this country.
- For me to publish in OA journals it would be essential that they have equivalent prestige and readership to the journals I currently publish in, otherwise this would have a serious impact on my career. Also much of my work is not grant-funded, so it would be necessary to persuade my institution to pay the publication fees.
- No credibility in RAE terms
- None currently can match the prestigious track record of the top print journals.
- ISI citations! Online paper citations are frequently not counted due to non-physical page numbering. It is imperative that OA journals attach page numbers to their articles and when they are cited, even if they are only published on the web.
- The Argentinian scientific grant institution (CONICET) and the Universities encourage to publish in journals with impact factor, by consider any other publication of lesser value. This is evaluated to get grants or job positions.
- They are not "real" journals. They do not have prestige. The internet is full of low grade essays such as students' term papers. The journals are too "open". Thy need to make it plain that they have normal review procedures.
- They key is the extent to which OA published journals will be effectively refereed and will manage the potential conflict between the author paying for publication and maintenance of quality. This will impact on their status for research assessment, personal career development etc.
- While my decisions have not been influenced by any outside body, I believe that my institution and grant-awarding bodies may consider OA sources non-acceptable when funding is considered. In particular, only journals in the ISI master journal list are regarded

as "true" journals by some bodies. I have not checked, however, whether any OA sources are in this list.
- RAE drives one towards 'prestige' journals. Some institutional type subscriptions (institution pays to put colleagues work on) is very expensive.
- Promotion and tenure committees do not rate these journals very highly - esp. as there is a fee charged to the author (which many people view as paying to have your work published).
- I work in an institution that takes the RAE very seriously. We are under pressure to publish in respected and prestigious journals - which currently do not include OA journals in my field. Perhaps in a few years time when well known authors publish and the impact fact ratings are seen to be good in these journals old and respected institutions like mine will change their policy. Another reason institutions like mine are not keen on OA journals at this juncture is the concern over quality - the issue
- They are not covered by indexing and abstracting services
- For the purposes of the Research Assessment Exercise in the UK I cannot afford to publish in any journal which is not old enough to have a citation rating.

Publication fees
- Funding agencies in the UK do not seem to have a policy on budgeting funds for OA journal publications. As a result, I don't have any funds available to pay the publication fees.
- In our country (India), generally we need not to pay to publish a paper. In some cases I presented the paper in conference and some I sent to journals to get it published. I am now putting those papers in Eprint archives.
- While I probably could find funds to publish in such a journal, that would be at the expense of other priorities in my lab. The bodies through which I am funded do not generally fund page charges or similar fees
- The average price is in the range of 500 US or 1500 Brasilian reals which represents about 15 percent of my monthly income. ie there is no hope as long as the price of publishing is born by the author. This system was of course the system used in the 19th century, perhaps it will have its day again, but those of us who have the patient material to base our studies may be stymied by cost rather than enthusiasm.
- Art historians already have to pay increasingly high fees to publish images of art works (even those well beyond copyright as far as the artist is concerned). An additional fee would be hard to bear by individuals. Most of already also pay our own research expenses. The culture of universities would have to change so that they bore the cost.
- I think that Open Access is potentially very dangerous. It has the potential to lead to only the richest people and institutions being able to publish their work, and to the stifling of originality and innovation. I am very skeptical about promises that the fees will be waived for poorer people, and it will add to the cost and lottery of the system to have to have people in place to make the decisions about whose fees are waived and whose aren't.
- It would depend on the size of the publication charge.
- Our Institute does not pay any fee for publishing in these open access journals
- I am perfectly happy to publish in OA journals which do not charge the author.
- I might have problems funding the fees. It of course depends on the size of them

Inertia; did not consider OA journals as a publishing vehicle
- It did not occur to me
- Inertia
- Inertia. I keep publishing in the same venues as I have in the past. If those journals adopt an open access policy, then I will be publishing in OA journals.
- Would be willing but am a creature of habit

No suitable OA journals; OA not a strong enough reason

- I am OAJ blind - I would certainly publish in one if it were held in high esteem by the workers in my field. As it is, none of the places I look to publish happen to be OA.
- There are not a lot of OA journals in my field to publish in (that are any good). And OA or *not* OA is not a determining factor in where to publish.
- When I was actively publishing in the past, open access journals did not exist.
- There are relatively few OA journals in linguistics and none of them seemed to be a good match for the papers that I was submitting.
- I'm not aware of OA journals in my field.
- Have not had the opportunity, or been made aware of such journals in my field
- I don't believe there are open-access journals in my field yet, so there's been no opportunity to publish in one. A traditional journal I already publish in has page charges. These are optional up to 12 pages and mandatory for more. I have paid these charges when I've had grant support.
- I don't know any and I try to publish in the best journals possible
- As far as I know, there are no OA journals in the fields of earth sciences in which I am interested. I prefer to publish in society journals which do keep their costs down. I am also basically still mostly paper-oriented.
- My decision was influenced by the journal I choose to send my manuscripts to. These are not always the same. The ones I have chosen do not happen to have OA publishing.
- I had not given the matter any thought, to be quite frank, since I publish in journal related to my areas of expertise, and to the best of my knowledge, none of these areas has an open access journal at the moment. However, I do object to paying a publication fee. The majority of my research is non-funded and I would be paying for publication out of my own pocket.
- Law reviews and similar publications often provide access, but don't follow any formal OA model. I don't believe there is an OA legal periodical at this time.

Not publishing at present; not in a position to choose journals to publish in

- No first authored publication in the last 3 years except for Hungarian language publications
- Just beginning an academic career..have not yet published much as yet (or at all, really): one article in review in an non-OA journal. The journal was chosen for its more British/European audience in my field (ornithology). I was unable to locate a similar OA journal...though admittedly, at the time, didn't look very hard (other authors in question rather in a hurry).
- Confidentiality conditions imposed by my employer prevent me from publishing in ANY journal.
- New to field, so have not published much yet.
- I'm still a PhD, so have only one paper published in a journal, so "last 3 years" statistics are not good :)
- I would publish exclusively in OA journals but am not high enough on the food chain to force the issue (postdoc).
- I currently am not writing to publish in journals.
- My work (translations, in any case) is published by my organization.
- Haven't completed work yet.
- I work for a small biotech company and for most cases company don't publish the work.
- I haven't had any new data ready yet
- Not ready to submit yet
- Just haven't had time to write more papers!
- I have been writing mainly book chapters and reviews, rather than primary research papers
- Haven't published in any journals recently.
- I feel they are more highly cited and they have higher impact, but I am a junior faculty member trying to get tenure. My senior faculty members put weight on the print-based peer-

reviewed journals. Since I only have a limited number of articles I will be publishing between now and my tenure decision, it would be foolish to "throw an article away" to a publication venue that would not have value in the tenure decision. Personally, I disagree with this, and plan to act differently when I have tenure

Already self-archiving so choice of journal does not matter

- I find ArXiv "publication" to be sufficient.
- I self-archive on the web in my own websites, using searchable html.
- 1. Like many computer scientists, I tend to publish in archival-quality conferences more than in journals. 2. In my field (computer science and more specifically computational linguistics), most content is already freely available online. Journals, as well as the archival-quality conferences where most work is published, are already sufficiently enlightened that they allow authors to post papers or at least preprints on their home pages.
- For my last peer-reviewed publication, the editor of the journal, through a colleague, expressed an interest in my article. This saved me the time of shopping it around. The journal allows for submission to an institutional archive, other forms of self-archiving, and other uses (e.g. presentation rights) so I was OK with having the article published in this journal.
- I prefer to start up my own sites and submit these to relevant communities and linking pages
- The top journals in my field are not currently OA. Journal ranking trumps whether the journal itself is OA in my mind. I would support any initiative to make the journals in which I publish OA. Were they OA now, I would continue to submit to them. (I self archive preprints and where possible post-prints of all my research in central archives. I know perfectly well that OA improves impact)
- I choose the journals to publish in based on the quality and the likely type of audience not any other characteristics. All our work is also published as working papers so is publicly accessible.
- There is a danger that it is perceived as vanity publishing, and is not of as high calibre. At our institution, we retain copyright so that we make our pubs available free on the web at our own site, just in a different format to the published version.

Effect of OA on learned societies

- OA may harm the scientific societies who publish academic journals
- As a matter of principal I try only publish my research in journals published by scholarly societies. I am strongly opposed to commercial academic journals and I have worked hard on various library committees to get the libraries to subscribe only to journals published by academic societies. I also actively lobby my colleagues to do the same. If the societies I belong to decide to publish OA then I will support them.

Peer review issues

- Traditional peer reviewing procedure helps keep the quality of publication better than anything else, OA journals may tend to lead to uncritical mass publication of mediocre or low-quality research among which it will be harder to find important research papers
- There will be a conflict of interest between high standards and accepting articles, if the journals get paid by authors rather than readers. I object to this as a matter of principle. Although, it is not clear why anyone would pay a journal for putting stuff on the web---after all the contribution of an electronic journal is not distribution (any webpage will do), but the review process---and that is usually unpaid.

Other/miscellaneous

- In comp sci people go to conferences, write conference proceedings, and prepublish in the web. The one OA journal (JMLR) is like Nature, publishing there requires more than will. :)
- The reference style for BMC Ecology is strange, and you are supposed to put the "methods" section at the end of the paper which is very uncommon and not convenient for the reader. Concerning the reference style, I suggest to define an universal format for all scientific publications, so we get rid of all complication concerning all the various formats. The author name and date should appear in the text, it is more convenient for the reader and the writer.
- I find OA journals difficult to browse
- I usually publish in conferences which put the proceedings in OA
- In general in computer science and related fields papers are published in journals for reasons of prestige and no other. We can get hold of nearly all papers from the author's own web sites. I believe the cost component that normal journals claim is excessive because we do all the formatting and editing, we provide ready to publish latex or pdf. I notice that in the humanities e.g. the journal of literary and linguistic computer they are still sending texts to typesetters, rather than requesting pdfs
- The bandwidth to access journals is a major issue in countries such as South Africa. We find that scholars in other countries do not have the patience to wait out our access delays. On our part we often find that articles are presented in a way that is very unfriendly to low bandwidth access, articles are too big and servers time out too soon for a download to complete.
- 1. Regardless of personal opinion, electronic publications are academically still not as widely accepted as print publications. 2. URLs can change over time, which may make it difficult to retrieve an electronic publication, i.e. there is a problem with stability of electronic resources that does not exist for more traditional forms of publication. 3. Data standards, soft- and hard ware change, which means that there is no guarantee that electronic data is still accessible in 10, 20, 50, 100 .
- Having authors pay the costs of publication will lead to both good and bad things. The good will be: 1. authors will write smaller papers in the style of Nature or J. Geophysical Letters and Geology because of page charges. This change will tighten the literature, and supporting material will be put into electronic archives. 2. There will be fewer published "dogs" out there because authors will not want to pay for page charges for crap. The bad outweighs the good
- I wish to support scientific societies which publish in my discipline. If they switched entirely to OA journal format, I would probably use them more. I have tried to access some OA journals in Current contents and other reviews and could not find them cited.
- I wish you could define OA better. Does this include online journals that require subscriptions? How about journals that claimed to be forever free (a few years ago) but now charge exorbitant fees?
- I object in principle to a system where well-funded persons may be able to publish more easily than someone that may not afford the fees. I believe that it will be open to corruption, whereby to fill an edition, papers may be passed that are not necessarily up to par, thereby reducing the quality, and eventually prestige of the journal. It is the possibility of corruption that I object to. However, I do have a problem with some publishing companies charging exorbitant fees to libraries for subscriptions
- I question whether the open-access cost model is sustainable in the long run. Also, I do NOT agree that scholarly information should be available without cost to everyone. My work has economic value, and I prefer to work with publishers who recognize that fact.
- I would be willing to publish in OA journals if offered a fee for producing work. Subscription sites charge for access to articles which authors receive no payment for. Authors spend a

great deal of time and effort on producing papers but receive no reward for this except standing in academic circles.

- It needs the structure change of the whole society to give the money currently used for library subscription for the scientists to pay for publishing. The library will not need to subscribe anything (do we need a library any more?)

Appendix 5:

Ease of access to research articles needed for work: results broken down by subject area

Subject area	I have easy access to all of the articles I need to read	I have easy access to most of the articles I need to read	I have easy access to some of the articles I need to read	I have easy access to very few of the articles I need to read
Agriculture & food science	9	50	27	14
Business & management	13	47	32	9
Chemistry	11	54	24	11
Computer sciences	6	68	20	5
Earth & geographical sciences	15	58	21	6
Engineering, materials science, technology	15	53	20	11
Humanities	7	44	32	15
Law & politics	18	41	29	12
Library & information science	7	53	29	12
Life sciences	9	59	26	6
Mathematics	4	64	21	11
Medical sciences	10	53	27	10
Physics	9	55	22	13
Psychology	10	49	36	5
Social sciences & education	10	55	23	12

Ease of access to research articles by subject area

(all figures are percentages of respondents in that subject area)

Appendix 6:

Reasons for publishing research results

Verbatim comments by respondents: comments categorised under various headings:

Advancement of scholarship/society/mankind

- I publish so as not to perish.
- I publish to put an end to a project so that I can move on to the next one. Publishing is the only way to do this without wasting years of work
- Contribute (in a small way) to the progress of science and the advancement of mankind. Help mentor students in writing papers.
- I think it's more important to bring the results of research to the general public.
- I publish to advance science and to communicate my results to the scientific community and public
- It´s fundamental to science.
- Intellectual curiosity and to contribute to the field
- Hopefully, my research will contribute to advancing the state of knowledge in my field.
- I publish because I enjoy feeling like I am part of the building knowledge base.
- I publish to communicate not only with peers but others such as teachers and parents.
- To further scientific knowledge and carry on the tradition of scientific progress
- The question of impact is only partially covered in your questions. If I want my basic research to have some influence on others doing applicable research I publish the results
- I publish for the sake of advancing progress in understanding infectious diseases, and for getting my country away from neglect and margination
- I publish to help create a more socially just society
- I publish out of the altruistic desire to share my findings with whoever finds them useful. This of course has an impact on the kind of topics
- I publish to deliver newly created knowledge to society
- I publish to advance knowledge in my field, in particular logic/philosophy/foundations of science.
- I publish to fix/specify the attained advancement in knowledge/understanding as such
- Dissemination of good practice and sharing it with peers
- I publish to let to let the scientific community learn about new findings they may be interested in.
- When I feel that a result I've obtained is not likely to be obtained by others, I consider publishing as my basic duty.
- Despite being retired I still wish to communicate my ideas to others and still seek to increase my standing in my field.
- I publish because I wish to discourse with others about important ideas
- Many of my articles are used in other people's courses.
- To assist in the education of students, peers and professionals in tourism and hospitality and related fields.
- I publish to advance science and knowledge in my field
- To (indirectly) communicate results to decision makers to contribute to (human) development
- I also publish to promote open inquiry among academic and professional people.

- I publish because I am employed by co-authors that wish to publish
- I publish also to share my results with other people working in my field.

Personal career progress/assessment

- You didn't list tenure, but I want to expressly state that my research is very HAPPILY outside the coerced research of the tenure system.
- I publish (with my students) in order to advance their careers.
- With students to help them accomplish all the above
- In the UK, publications by staff in a department directly affect the funding made to that department (via the "Research Assessment Exercise")

Stamp claim on work, document results, posterity

- I publish to archive information for posterity
- I publish to document my research
- I publish to claim an idea as my own

Requirement of job

- I publish because writing papers is in itself a research activity
- Technology transfer is a requirement in my job description.
- As a scientist who has gained much knowledge from other people's work, I feel it is my basic duty to communicate my own research findings to others
- I publish because it's my job (as translator for WIPO).
- Publication is a prerequisite of accepting public funds for research.
- It is a requirement of my job.
- Requirement of my post
- I see carrying out research and seeing it published as part of my job as a university professor.
- Because my boss will want to know why if I don't
- I publish because my job requires it.
- I publish because it is a professional responsibility, and demanded by my employment contract.
- I am a full-time researcher with publication targets as part of my employment contract

Feedback from peers and scholarly community

- To get peer feedback on my work! This is very important to me.

Personal satisfaction

- I publish out of interest.
- For fun
- I publish for the inherent satisfaction of writing and seeing my work in print
- I am 58 years of age, I am not looking for recognition a new job or money, I do it because it is fun and intellectually stimulating
- It is emotionally and aesthetically appealing to see your freshly published paper in print
- It's a challenge to set out clearly and competently what one has learned and feels personally rewarding to have accomplished that.
- For self esteem and satisfaction

Enhancement of the reputation of their institution

- To enhance prestige of my institution
- I publish to increase the prestige of my institution.

- I publish to gain prestige for my company

Other

- I publish a lot by 'accident'. I do joint research and often don't know until after the fact that my collaborators have put my name on a publication

Appendix 7:

Use of closed archives: results broken down by subject area

Subject area	ScienceDirect	CrossRef Search	Other cross-subject full-text services	Subject-specific full-text services	Individual publishers' own websites
Agriculture & food science	74	7	14	20	33
Business & management	51	9	28	32	38
Chemistry	79	17	24	12	32
Computer sciences	55	8	23	21	48
Earth & geographical sciences	61	3	6	18	24
Engineering, materials science, technology	68	10	20	12	36
Humanities	33	6	35	42	44
Law & politics	24	0	53	47	65
Library & information science	43	12	38	52	42
Life sciences	62	5	15	16	43
Mathematics	60	13	19	24	43
Medical sciences	52	6	22	26	34
Physics	62	12	14	22	49
Psychology	60	5	28	37	36
Social sciences & education	41	4	34	35	40

Use of closed archives by subject area

(all figures are percentages of respondents in that subject area)