Loughborough University

This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

PEER behavioural research: authors and users vis-à-vis journals and repositories. D4.2 final report.

PLEASE CITE THE PUBLISHED VERSION

http://www.peerproject.eu/reports/

PUBLISHER

LISU and Loughborough University

VERSION

VoR (Version of Record)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Fry, Jenny, Steve Probets, Claire Creaser, Helen Greenwood, Valerie C.L. Spezi, and Sonya White. 2019. “PEER Behavioural Research: Authors and Users Vis-à-vis Journals and Repositories. D4.2 Final Report.”. figshare. https://hdl.handle.net/2134/12967.

Loughborough University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

creative commons

COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:



Attribution. You must attribute the work in the manner specified by the author or licensor.



Noncommercial. You may not use this work for commercial purposes.



No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license).

Disclaimer

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

PEER Behavioural Research:

Authors and Users vis-à-vis Journals and Repositories

Final Report

Jenny Fry, Steve Probets
Department of Information Science

Claire Creaser, Helen Greenwood, Valérie Spezi, Sonya White
LISU

August 2011

LISU Research & consultancy for performance management
Information, cultural & academic services

Loughborough University

Contents

Acknowledgments ii

List of abbreviations iii

Executive summary iv

1 Introduction 1
- 1.1 Context 1
- 1.2 Research objectives 1
- 1.3 Notes on methodology 2

2 Awareness of Open Access and Open Access Repositories 7
- 2.1 Different types of repositories 10
- 2.2 Disciplinary similarities and differences 12

3 Open Access Repositories in the research process 13
- 3.1 The use of Open Access Repositories by readers 13
- 3.2 Authors' open access behaviour and self-archiving practice 29

4 Researchers' perceptions of the best way(s) of achieving Open Access 43
- 4.1 Motivations to self-archive and perceived benefits of OA 43
- 4.2 Perceived barriers of Open Access Repositories 51
- 4.3 Open Access Publishing 54
- 4.4 Perceived influence of OAR on scholarly publishing 55
- 4.5 Disciplinary similarities and differences 59

5 Discussion 62

6 Key highlights and conclusions 74

Bibliography 77

Appendices 79

Appendix 1: Methodology 79

Appendix 2: Representativeness of the survey samples 92

Appendix 3: Summary of Phase 2 survey responses 94

Acknowledgments

The authors would like to thank all those who have assisted with this research and the report:

- Michael Mabe and members of the PEER Executive
- The Research Oversight Group and Special Advisor
- Chris Armbruster and Julia Wallace
- The publishers and repositories participating in the PEER Observatory
- Those researchers who completed the surveys and participated in the focus groups and participatory workshop
- Kerstin Michaels and colleagues in the Department of Information Studies at UCL
- Sharon Fletcher and Mary Ashworth (LISU)

List of abbreviations

arXiv	An e-print service popular in physics, mathematics, computer science and other related fields
CTA	Copyright Transfer Agreement
DOI	Digital Object Identifier
HAL	Hyper Article en Ligne
HE	Higher Education
HEP	High Energy Physics
OA	Open Access
OAR	Open Access Repository / Repositories
OAJ	Open Access Journal(s)
PEER	Publishing and the Ecology of European Research
PLoS	Public Library of Science
PMC	PubMed Central
RePEc	Research Papers in Economics
ROAR	Registry of Open Access Repositories
SSHA	Social sciences, humanities & arts
SSRN	Social Science Research Network
STM	Science Technology and Medicine

Executive summary

The Behavioural research project is one of three independent research projects commissioned and managed by PEER as part of the PEER Observatory. The aim of the Behavioural research project was to address the role of stage-two manuscript repositories in the scholarly and scientific communication system by exploring perceptions, motivations and behaviours of authors and readers. The research was carried out between April 2009 and August 2011 by the Department of Information Science and LISU at Loughborough University, UK.

Key conclusions

Over the period of Phases 1 and 2 of the Behavioural research project the increase in the number of researchers who reported placing a version of their journal article(s) into an Open Access Repository was negligible.

Researchers who associated Open Access with 'self-archiving' were in the minority.

Open Access is more likely to be associated with 'self-archiving' (Green Road) by researchers in the Physical sciences & mathematics and the Social sciences, humanities & arts, than those in the Life sciences and the Medical sciences who are more likely to associate Open Access with Open Access Journals (Gold Road).

There is anecdotal evidence that some researchers consider making journal articles accessible via Open Access to be beyond their remit.

Authors tend to be favourable to Open Access and receptive to the benefits of self-archiving in terms of greater readership and wider dissemination of their research, with the caveat that self-archiving does not compromise the pivotal role of the published journal article.

Readers have concerns about the authority of article content and the extent to which it can be cited when the version they have accessed is not the published final version. These concerns are more prevalent where the purpose of reading is to produce a published journal article, and are perceived as less of an issue for other types of reading purpose.

Academic researchers have a conservative set of attitudes, perceptions and behaviours towards the scholarly communication system and do not desire fundamental changes in the way research is currently disseminated and published.

Open Access Repositories are perceived by researchers as complementary to, rather than replacing, current forums for disseminating and publishing research.

Approach to the project

The research consisted of two phases using a mixed methods approach:

- Phase 1 consisted of an electronic survey of European journal article authors for which 3,139 valid responses were received. A series of four broad disciplinary-based focus groups were conducted in parallel to the survey and were held in London (Medical sciences), Berlin (Social sciences, humanities & arts), Rome (Physical sciences & mathematics) and Budapest (Life sciences). Collectively there were 21 participants for the focus groups.

- Phase 2 consisted of a follow-up electronic survey of European journal authors, which drilled down into some of the phase 1 results in more depth. A total of 1,427 valid responses were received. A small scale repository exit survey was also implemented at six of the participating PEER repositories. The number of valid responses received was very low, which is disappointing, and had the number of respondents been higher the exit survey may have yielded some very interesting findings about readers who are not necessarily authors. A one-day participatory workshop was held in London with 22 participants recruited from across Europe. The aim was to drill down from the broad disciplinary groupings used for the surveys and focus groups to individual disciplines, and the four disciplines included in the workshop were: Earth, marine & environmental sciences; Chemistry; Engineering; Applied social sciences.

Key findings

Authors

- The central role of the peer-reviewed journal article in academic careers is not diminishing despite the emergence of digital scholarship and novel modes of dissemination.
- Important factors influencing researchers' decisions to disseminate via peer-reviewed journal articles are wide-spread visibility of their research and dissemination to specific target audience(s), as well as career advancement. Researchers have knowledge of the readership for each journal title and this influences choice of journal for publication. Researchers who were in the early stages of their careers were more likely to rank career advancement as being the most important influence in choosing to publish in a peer-reviewed journal.
- The process of evaluating where to publish and what to cite places a high value on peer review as being the primary mechanism for quality control in a discipline. Put another way, researchers linked good quality peer review with highly esteemed journal titles.
- Anecdotal evidence from the qualitative research suggests that in some European countries journal impact factors are becoming increasingly important in terms of evaluating the quality of researchers' outputs.
- Open Access was perceived by focus group and workshop participants as an effective solution to the common tension between institutional-based motivations, e.g. reputation and career advancement, to publish in peer-reviewed journals and audience-based motivations to make research outcomes as widely available as possible. There appears to be confusion amongst researchers, however, about the distinction between Open Access journals, which may have a journal impact factor, and Open Access Repositories (OAR), which do not.
- There were some reservations amongst authors about having their peer-reviewed published journal articles held in an OAR with other content of variable quality, which they felt in some way might influence the perceived quality of their own articles. Researchers felt that as long as the peer-reviewed journal article was the major mechanism for quality control in their discipline, then OAR and peer-reviewed journals would need to co-exist in some mutually beneficial way.
- Subject-based and institutional repositories have emerged in a different context and with a different set of goals. In some disciplines there are de facto centralised repositories, such as arXiv (an e-print service popular in physics, mathematics, computer science and other related fields) in physics, RePEc (Research Papers in Economics) in economics and PMC (PubMed Central) in the medical sciences, while in other disciplines the repository landscape

is less well-established. In terms of authors' preference for type of repository in which to place a version of their article(s), there was a slight shift between phase 1 and phase 2 towards institutional repositories.

- The nature of the motivation to self-archive seems to influence in which type of OAR authors are likely to place (or give permission to have placed) a version of their article(s). If the motivation is voluntary, requested by a co-author, or an invitation by a publisher then authors are more likely to choose a subject-based repository. If, on the other hand, authors are invited by repository/library staff, or were mandated by their employer, they are more likely to choose an institutional repository. Respondents to the phase 2 survey did, however, consider funder and institutional mandates to be relatively unimportant as motivators for repository deposit. The participatory workshop threw some light on this, with researchers reporting that where funder or institutional mandate policies did exist they tended to be insufficiently enforced.
- Researchers' perceptions of how they would respond to the enforcement of multiple mandates were explored in the phase 2 participatory workshop, and the anecdotal evidence seems to suggest that if institutions were to enforce mandates then researchers would feel compelled to prioritise their employer's mandates over other types of mandate.
- Across the phase 1 and 2 surveys approximately half of the authors reported having placed (or had placed on their behalf) a version of their journal article(s) in an OAR. Authors from the Physical sciences & mathematics and the Life sciences were more likely to place a version of their article(s) in an OAR themselves, whilst authors in the Medical sciences or the Social sciences, humanities & arts were more likely to have a version of their article(s) placed in an OAR by somebody else.
- The published final version (publishers' PDF file) is the version most likely to be placed in an OAR. Authors from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to place (or have placed on their behalf) a pre-print or author's final peer-reviewed accepted version. Authors from the Medical sciences were most likely to be uncertain about which version of their article had been placed in an OAR, which correlates with the high proportion of Medical science researchers who reported that their article(s) had been placed in an OAR by someone else.
- Almost one-third of phase 2 survey respondents indicated that someone else had made their work available in an OAR on their behalf.
- In terms of authors' experiences of depositing a copy of their article(s) in an OAR themselves the survey results suggest that authors do not generally experience much difficulty, although anecdotal evidence from the phase 2 participatory workshop indicates that researchers do at times find this process 'tedious' and 'time-consuming'. Nonetheless, findings from the phase 2 survey indicate that checking publishers' Open Access policies is the main barrier to authors managing the self-archiving process themselves.
- In terms of weighing the perceived benefits of OA against the effort of placing a copy of their journal article(s) in an OAR, the majority of the authors surveyed in phase 2 felt that it was either 'definitely' or 'probably' worth the additional work involved. Authors within the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to feel that placing article(s) in an OAR was 'definitely worthwhile', compared to authors in the Life sciences and the Medical sciences.

- There appears to be a lack of awareness of publishers' open access embargo periods, with just over half of authors surveyed in phase 2 stating that they did not know or could not remember what embargo period, if any, was enforced by the publisher when they placed their article in an OAR.
- Respondents to the phase 2 survey who specified that shorter publisher embargo periods would make them more likely to make their article(s) available via open access, with longer periods making them less likely to do so, were almost equal in number to those authors who perceived that the length of publisher embargo periods would not influence their behaviours.
- Widening the accessibility of published journal articles to researchers beyond Higher Education, i.e. those in the public, private and third sectors, or researchers in countries with emerging and developing economies, was perceived by researchers as an area of concern that OAR could effectively address, but that responsibility for widening accessibility should be taken up by other stakeholders in the scholarly communication system.

Readers

- Researchers typically select a narrow range of information resources that they use on a regular basis to locate research-based sources, and for most disciplines the search strategies deployed within the scope of these resources appear to be highly individualised.
- Traces of disciplinary differences in the information resources used were found, with bibliographic literature databases more likely to be used by researchers from the Life sciences, and subject-based portals/repositories more likely to used by researchers from the Medical sciences, whilst researchers from the other broad-based disciplinary groups are more likely to start their search strategy from a publisher‘s journal platform or Google/Google Scholar. Career length also seems to play an influential role, with less experienced researchers (i.e. fewer than five years) being more likely to initiate a search using Google Scholar, and the most experienced (i.e. 25 years or more) being more likely to browse print-based journals.
- In most disciplines researchers are unlikely to go directly to a repository to search for journal articles, with Google and Google Scholar being the most likely route by which researchers locate material within OAR. The exceptions to this were researchers who reported using well-established subject-based repositories, such as PMC, arXiv, RePEc or SSRN (Social Science Research Network).
- Researchers' satisfaction with a version other than the published final version of a journal article is closely related to the purpose of their reading. Researchers seeking articles in order to cite them in their own articles are most likely try to locate the published final version. Accessing the published final version of an article, however, appears to be less critical for writing research outputs other than peer-reviewed journal articles.
- Researchers from the Medical sciences and the Social sciences, humanities & arts were most likely to consider the article version important and least likely to 'trust' versions of articles held in a repository unless it was very clear to them that they had accessed the published final version. On the other hand, there was a greater acceptance of pre-prints by researchers from the Physical sciences & mathematics than researchers from the other broad-disciplinary groups.
- The clarity of repository metadata and readers' ability to distinguish between the different versions of an article seems critical to how the 'quality' (e.g. whether it is authoritative) of repository content is perceived. Researchers reported that whilst the difference between a

pre-print and a published final version was reasonably clear, it was much more difficult to distinguish between a submitted stage-one article and an accepted stage-two article.

1 Introduction

1.1 Context

The *Behavioural research: Authors and users vis-à-vis journals and repositories* project was commissioned by PEER in April 2009 as part of a broader initiative to investigate the effects of the large-scale, systematic deposit of authors' final peer-reviewed manuscripts (also called stage-two research outputs) on reader access, author visibility, and journal viability, as well as on the broader ecology of European research. The specific aim of the behavioural research was to understand the extent to which authors and users are aware of Open Access (OA), the different ways of achieving it, and the (de)motivating factors that influence its uptake.

This report is the final deliverable of this project. The research was carried out in two phases between 2009 and 2011. In Phase 1, an extensive survey of European researchers was supported by a series of focus groups that ran concurrently and explored specific issues in greater depth. In Phase 2, a second survey of European researchers was used to gather information in more detail on specific issues raised by the phase 1 findings. A final workshop for researchers from across Europe, covering a range of disciplines, served to consolidate the key findings.

A baseline report covering the first phase of the study (Fry *et al*., 2010), carried out in 2009, is available from the PEER website[1]. This final report covers the second phase of the research, carried out in 2010-11, and provides a synthesis of the results of both phases.

1.2 Research objectives

The overall goal of the behavioural research was to develop an understanding of the perceptions, motivations and behaviours of authors and readers with respect to the use of authors' final draft manuscript in Open Access Repositories (OAR).

The specific research objectives were:

- To identify the choices authors and readers make in locating and selecting sources in the context of publication and dissemination, and information seeking behaviours, and the major influences on those choices.
- To examine ways in which author and reader choices influence the role played by repositories in the scholarly communication landscape.
- To identify the common perceptions authors and readers have in relation to OAR, and the ways in which such perceptions influence publication, dissemination, and information seeking behaviours.
- To explore researchers' green OA experience both as authors and readers.
- To investigate authors' perceptions of the values/benefits of OAR in relation to the effort involved in making their work available via OAR.

[1] http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf [accessed 29.09.11]

- To explore researchers' perceptions of publishers' OA embargoes and their impact, if any, on OAR behaviours.
- To investigate the behaviours and characteristics of readers and researchers using OAR.
- To explore the influence of context on readers' behaviours in relation to OAR, for example, reading for current awareness, in the preparation of research grant proposals, or for article writing.
- To investigate whether there are identifiable coarse-grained characteristics of authors and readers that influence their OAR behaviour (e.g. institutional type, region, discipline, career status etc.).

This report summarises the findings of both phases of the research, and relates these to relevant evidence in the literature.

1.3 Notes on methodology

Broad disciplinary groupings, based on the categorisation of journals included in the PEER Observatory, were used throughout the research, both in the analysis of the quantitative data (phase 1 and phase 2 surveys) and in the selection of the participants in the two qualitative data gathering methods (phase 1 focus groups and phase 2 participatory workshop). Details about the broad disciplinary groupings and their composition are provided in Table 1.1.

Table 1.1 Broad disciplinary groupings

Broad disciplinary grouping	Disciplines included:
Medical sciences	Clinical medicine
	Clinical dentistry
	Anatomy & physiology
	Nursing & paramedical studies
	Health & community studies
	Pharmacy & pharmacology
Life sciences	Biosciences
	Psychology & behavioural sciences
	Earth, marine & environmental sciences
	Veterinary science
	Agriculture & forestry
Physical sciences & mathematics	Chemistry
	Physics
	Mathematics
	General engineering
	Chemical engineering
	Mineral, metallurgy & materials engineering
	Civil engineering
	Electrical, electronic & computer engineering
	Mechanical, aero & production engineering
	Information technology & systems sciences & computer software engineering

Broad disciplinary grouping	Disciplines included:
Social sciences, humanities & arts	Architecture, built environment & planning
	Catering & hospitality management
	Business & management studies
	Economics
	Geography
	Social studies
	Media studies
	Humanities & language based studies
	History
	Archaeology
	Modern languages
	Design & creative arts
	Education & Sports

In the surveys, for both Phase 1 and Phase 2 of the study, respondents were asked to indicate in which discipline(s) they carried out research, and were allocated to a broad disciplinary group accordingly. Respondents were allocated to an Interdisciplinary group if they ticked disciplines from two or more of the broad areas included; the majority of these had indicated disciplines in both the Medical sciences and the Life sciences. Respondents were also asked to select their institution type from a list of options. Details are provided in Table 1.2 and Table 1.3.

Table 1.2 Distribution of survey respondents by discipline

Disciplinary groupings	Phase 1 survey (No. of respondents)	Phase 2 survey (No. of respondents)
Life sciences	416	311
Physical sciences & mathematics	1,773	454
Social sciences, humanities & arts	259	167
Medical sciences	248	194
Interdisciplinary	440	300
Total[2]	3,136	1,426

[2] This is the total number of respondents who have selected at least one discipline, not the total number of valid responses to the survey(s).

Table 1.3 Distribution of survey respondents by type of institution

Institution type	Phase 1 survey (No. of respondents)	Phase 2 survey (No. of respondents)
University or College	2,211	922
Hospital or medical school	161	145
Research Institute	619	265
Industry or commercial	47	29
Government	63	36
Other	38	30
Total	3,139	1,427

The analysis of the phase 1 focus groups was based on these broad disciplinary groupings, with some minor references to participants' specific disciplines where appropriate, whilst for the phase 2 participatory workshop a more fine-grained approach was desired, and participants were selected as far as possible, from the individual disciplines within each broad disciplinary grouping *(Table 1.4).*

Table 1.4 Demographics of the focus groups and participatory workshop

Phase 1	Broad disciplinary group	Discipline & country	Institution types	No. accepting the invitation / No. of participants
Focus groups	Life sciences	Hungary-based researchers from a broad range of disciplines within the Life sciences group; focus group held in Budapest	University or Research centre	6 / 6
	Physical sciences & mathematics	Italy-based researchers from Physical sciences, Structural Engineering, Chemistry; focus group held in Rome	University or Research centre	9 / 7
	Social sciences, humanities & arts	Germany-based researchers from a broad range of disciplines, including Law, Sociology, Gender studies; focus group held in Berlin	University or Research centre	7 / 5
	Medical sciences	UK-based researchers from Immunology, Anatomy, Ethno-botany & pharmacology; focus group held in London	University or Research centre	6 / 3
	Total			21
Phase 2	Broad disciplinary group	Discipline	Institution types	No. of participants
Participatory workshop	Life sciences	Earth, marine & environmental sciences	University or College x5 Charity x1	6
	Physical sciences & mathematics	Chemistry	University or College x2 Research institute x3	5
		Engineering	University or College x4 Research institute x2	6
	Social sciences, humanities & arts	Mixed	University or College x3 Research institute x1 Private research institute x1	5
	Total			22

The questionnaire distributed in Phase 1 of the project aimed to ascertain levels of awareness of Open Access, and in particular Open Access Repositories (OAR), as well as researchers' attitudes and perceptions of OAR. A total of 3,139 responses were included in the analysis. Four focus groups were conducted in Phase 1, held in London, Rome, Budapest and Berlin, which aimed to explore further some of the early findings of the phase 1 survey. They included

a total of 21 participants. Full details are provided in the baseline report (Fry *et al.,* 2010), available from the PEER website.[3]

Although this research was initially planned as a longitudinal study, it became evident that little would have changed between the first survey and the second, so the longitudinal approach was dropped in favour of a greater exploration of authors' and readers' behaviours in relation to OAR in Phase 2. The second survey was therefore geared at digging down into the findings of Phase 1, and examining the details of researchers' behaviours both as authors and as readers. A total of 1,427 valid responses were received and analysed. The workshop, conducted after the survey had taken place, aimed to enhance the survey findings by giving them some context and texture. A total of 22 researchers from across Europe participated in this workshop. Full details of the methodology employed in Phase 2, together with demographic information on phase 2 survey respondents, are given in Appendix 1.

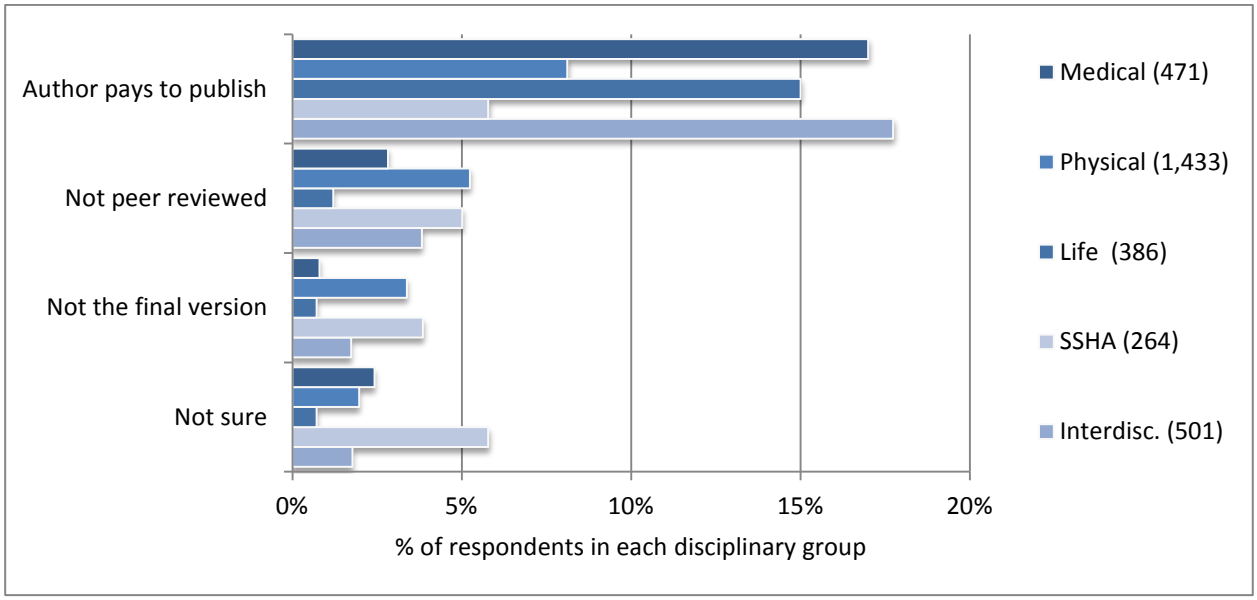
Following the phase 2 survey, data were made available concerning the broad disciplinary basis on which the invitations to complete both surveys had been distributed. This enabled the research team to investigate the representativeness of the responses in terms of that distribution. Details are given in Appendix 2; in broad terms medical scientists were under-represented in the samples, and life scientists over-represented, while responses from physical scientists and those in the Social sciences, humanities & arts were broadly in line with the initial distribution. As a result, the analyses presented in this report have been weighted to account for the different response rates by discipline, and phase 1 results presented here may differ in detail from those in the baseline report.

[3] http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf [accessed 29.09.11]

2 Awareness of Open Access and Open Access Repositories

The phase 1 survey data indicated that there is a general awareness of OA, with more than two-thirds of respondents understanding OA to mean free electronic access to full-text journal articles. Less than five percent of the phase 1 survey respondents considered OA to mean low quality, not peer-reviewed, not the final version of an article, or vanity publishing. However during the project there was some uncertainty identified over the precise meaning of the term 'Open Access', with focus group participants expressing uncertainty over what 'Open Access' really entails. This is supported by the survey findings, where only six percent of respondents equated OA with self-archiving, and 11% with author-side payment. However, the phase 1 findings indicate that general awareness of OA is growing compared to results from earlier seminal studies conducted by Rowlands *et al*. (2004) and Swan and Brown (2004, 2005).

Figure 2.1 Perception of the term 'Open Access', by subject



(Weighted total numbers of responses given in brackets)

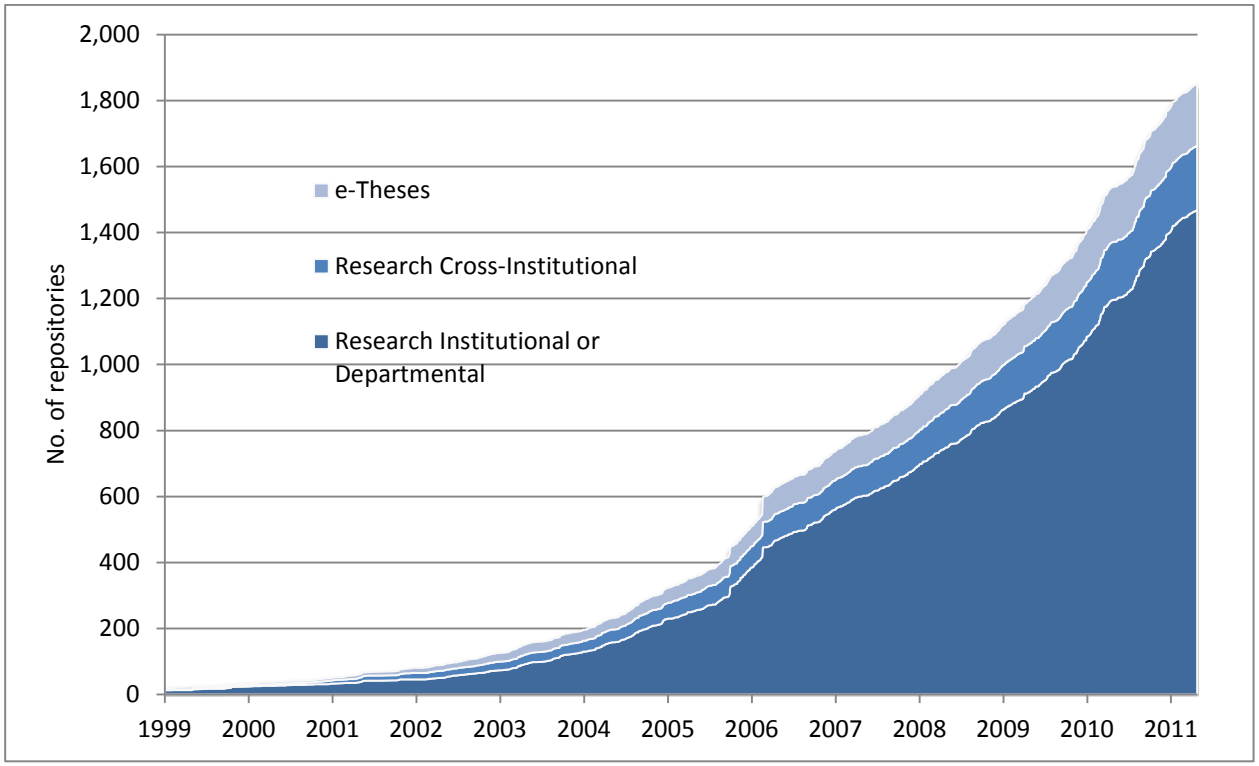
Amongst phase 1 survey respondents, the general level of awareness of OA did not differ greatly across disciplines, although Figure 2.1 shows that researchers in the Life sciences and the Medical sciences had slightly differing views on what OA entails to researchers from other disciplines. Outside of the Life sciences and the Medical sciences, some researchers expressed concern that OA equated to 'not peer-reviewed', although the percentage of respondents who held this opinion was low across all disciplines (4.1% overall). Within the Life sciences and the Medical sciences, and amongst Interdisciplinary researchers, OA was more closely associated with the 'author pays' model of Open Access Journals (OAJ), rather than self-archiving. This echoes findings from recent studies which indicated that the 'author pays' model (or Gold OA) is very widespread and developed in the Life sciences and the Medical sciences. Björk *et al.* (2010) found that Gold OA is prominent in the Life sciences, whereas Green OA is well-developed in disciplines such as Earth sciences, Physics and Astronomy. In their study, Dallmeier-Tiessen *et al* (2010) found that STM is represented by 66% of OAJ (pure OAJ and hybrid journals) and contributes to 77% of articles, of which Biology and life sciences represent 19% of OAJ and 21% of articles, while Medicine and health sciences represent 28% of OAJ and 28% of articles.

The phase 1 survey and focus groups enabled the research team to further explore levels of awareness of OA. Whereas the survey findings indicated a reasonably good understanding of OA and OAR, the focus groups provided a more mixed picture. Focus group participants expressed uncertainty about what OA entails and the different ways of achieving it. The analysis of the free text responses received in the phase 1 survey also revealed a discrepancy between what was reported in the multiple choice questions and what researchers really understand OAR to be, showing confusion during discussions with regard to what defines an OAR. For example, phase 1 survey respondents were asked to name the OAR that they know or use in order to gain an insight into the repository landscape of the surveyed population. A considerable number of responses listed OAJ and/or OA publishers (e.g. Biomed central, Public Library of Science) and subscription-based journal platforms; the latter indicating the difficulty that respondents had in distinguishing between OA resources and subscription resources where institutional libraries provide seamless access to subscribed resources. The archive most often mentioned was arXiv (by 52% of respondents to that question, mostly from the Physical sciences & mathematics). Other major archives mentioned included PubMed Central (4% of respondents, from the Medical sciences, the Life sciences and Interdisciplinary areas); Citeseer (4% of respondents, from the Medical sciences, the Physical sciences & mathematics and Interdisciplinary areas); HAL (4% of respondents, from the Physical sciences & mathematics); and RePEc (2% of respondents, from the Social sciences, humanities & arts).

Although longitudinal analysis of the findings was not the primary focus of the second survey, there is anecdotal evidence that researchers' levels of awareness of OAR have changed little, while the number of institutional repositories has continued to increase *(Figure 2.2).* The Registry of Open Access Repositories (ROAR)[4] reported 425 new repositories being added to their lists in 2010, with a further 169 to date in 2011. However, phase 2 workshop participants indicated that they felt levels of awareness of OAR remained low amongst researchers. Supporting a view expressed by Harnad *et al.* (2009) that OAR were not really on scholars' radar, they indicated that OAR did not have sufficient visibility. OAR do not seem to be perceived as bringing something essential to researchers, nor something that is missing from current scholarly communication. Examples of low levels of awareness of institutional repositories included workshop participants learning of the existence of a repository at their institution through discussion with the research team at the workshop. Others were not always sure whether their institutional repository was OA or available for access only by members of the institution.

[4] http://roar.eprints.org/ [accessed 29.09.11]

Figure 2.2 Growth of institutional repositories internationally



Source: data supplied by ROAR (July 2011)

In addition, some workshop participants showed confusion about OAR, notably in the Chemistry group, where participants were, to a certain extent, aware of arXiv[5] and the physicists' pre-prints culture, and thus tended to conflate OAR with repositories of pre-prints. Such misunderstandings from the Chemistry group were surprising, given that some of them regularly use another subject-based repository, PubMed Central (PMC), which holds publishers' PDF versions of biomedical and life sciences articles. One point arising from the focus groups and participatory workshops was the unspoken perception, at least in the Chemistry group, that PMC is not an OAR but rather a commercial database or aggregator of published resources that one can access easily, without thinking further why it is possible for them to access this database. This reinforces the finding from the phase 1 survey noted earlier, that only 4% of respondents listing subject-based repositories mentioned PMC.

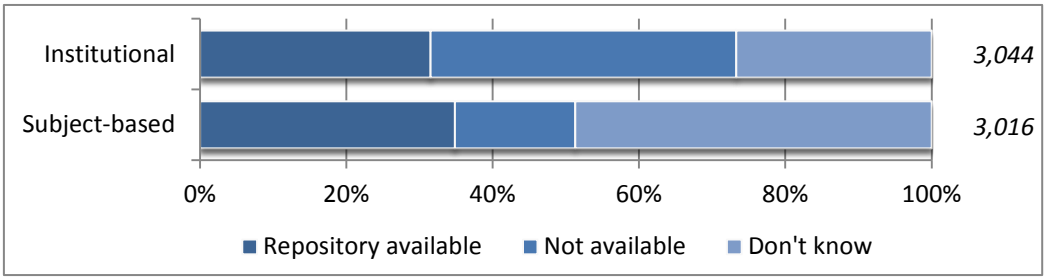
Although the workshop sessions explicitly focused on OAR, at no point did it occur to the Chemistry group to consider their OA practice in relation to PMC. There may be two reasons for this: firstly, participants tended to conflate OA material with pre-prints, and PMC primarily contains publishers' PDFs. This may mislead users into thinking that they are accessing a commercial platform to which their institution subscribes. Secondly, the deposit process into PMC is mainly managed by publishers, so authors may not be fully aware of how and when articles are added. This makes it difficult for them to recognise PMC as a subject-based repository, since it operates differently to self-archiving OAR. The immediate implication of this is that authors would probably trust more OAR if publishers were involved, in the same way that they are involved in PMC.

[5] ArXiv is the repository of e-prints in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance and statistics

2.1 Different types of repositories

Amongst the focus group participants, the term 'repository' was not universally understood to mean the same thing and some did not have a clear idea of what a publicly available OA repository was. A lack of awareness of the existence of OAR was also indicated by phase 1 survey respondents, with over 26% of respondents being unsure whether or not their institution had a repository *(Figure 2.3),* and 47% being unsure whether a subject-based repository was available to them.

Figure 2.3 Availability of repositories

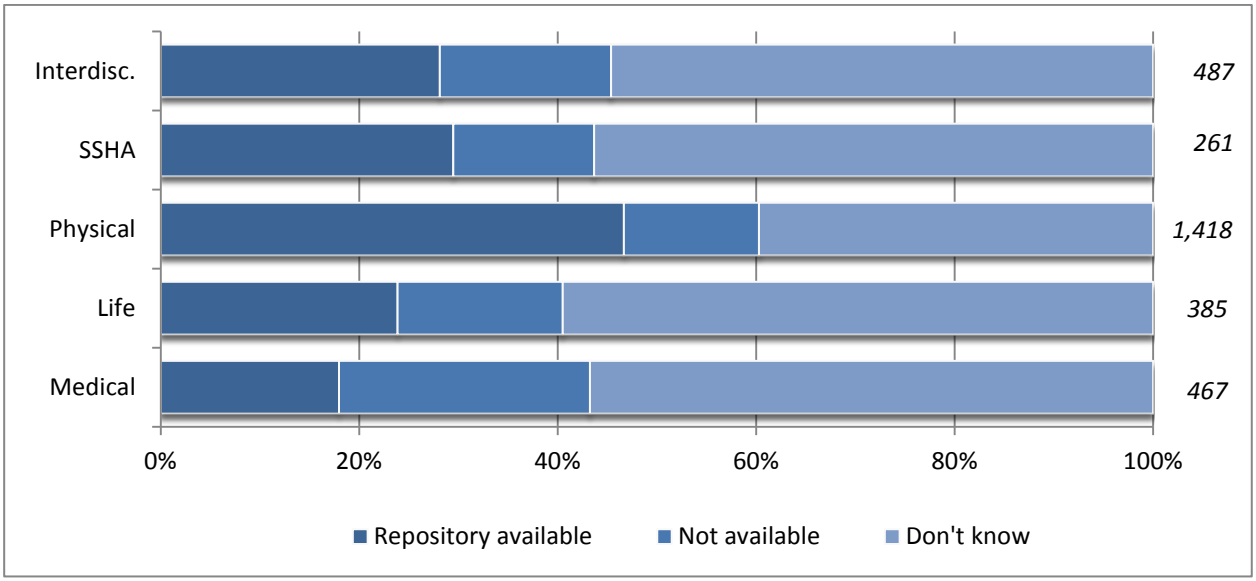


(Weighted total numbers of responses given in italics)

Physical sciences & mathematics phase 1 survey respondents were significantly more aware of the subject-based repository options available to them *(Figure 2.4).* This increased awareness was also reflected amongst participants in the Physical sciences & mathematics focus group, where, for a number of participants, the use of repositories, such as arXiv, was described as part of the daily or weekly workflow, both in terms of depositing and locating papers. Community-based information services, such as SPIRES[6] or arXiv, play a major role in the information landscape of the High Energy Physics (HEP) community, with 87.9% of HEP scientists reporting using these two information sources in most of their information searches (Gentil-Beccot *et al.,* 2009). It has been recognised that the uptake and use of non-market based information resources, such as institutional and subject-based repositories, has been problematic - with long lead times in terms of content population (Duranceau, 2008). Yet amongst the HEP community almost the opposite is true. This is likely to be explained by the fundamental cultural characteristics of the HEP community. Our findings do, however, emphasise the importance of connective structures in digital information environments (Palmer, 2005) and the 'location' of OAR in the broader context of competing resources, with search preferences likely to play an influential role in resource discovery and use.

[6] SPIRES is a metadata-only search database for the field of high energy physics. It collects metadata from repositories and journal literature and proposes additional services such as citation analysis, keywords, authors' affiliation, matching pre-prints with publications, etc. Search facilities in arXiv are extremely limited and, as pointed out by Gentil-Beccot *et al.,* (2009), SPIRES and arXiv may be seen as the two ends of a single information system, with arXiv focusing on data storage and SPIRES offering users search facilities and other services.

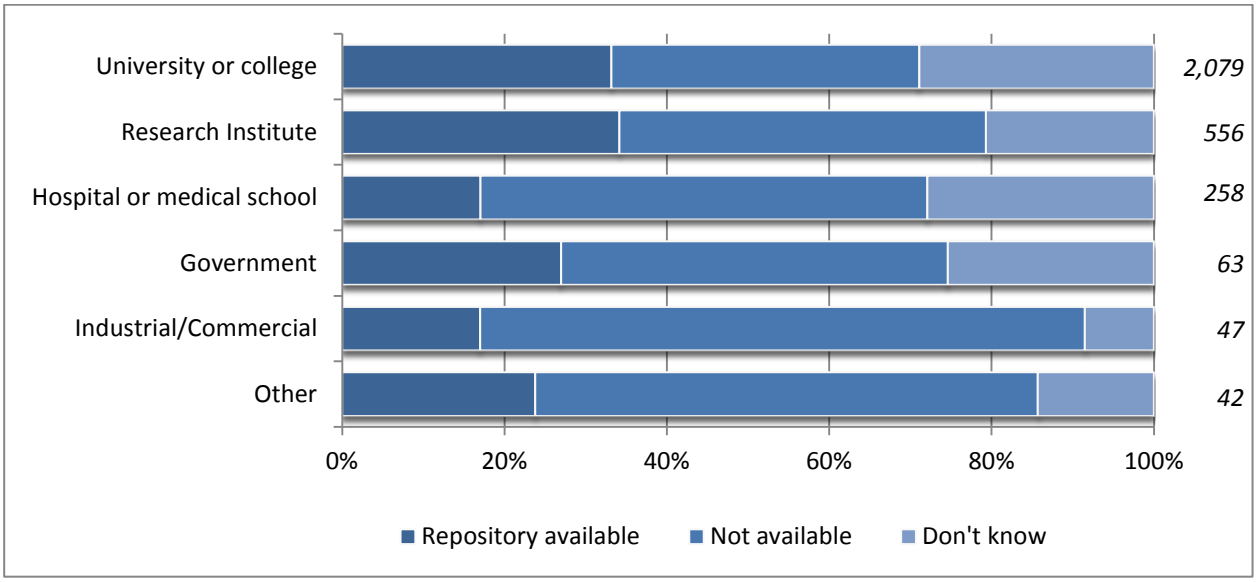
Figure 2.4 Availability of subject repositories, by subject



(Weighted total numbers of responses given in italics)

In the phase 1 survey, the type of institution in which researchers were based was found to influence their awareness of OAR. Figure 2.5 shows that survey respondents from the industrial and commercial sectors were most likely to know whether their institution had its own publicly available repository, and least likely to report that their institution did have one. The percentage of researchers from universities, colleges, medical schools and research institutes who did not know whether or not their institution had an OA repository implies that awareness of the existence of institutional OAR could be improved. One implication is that even if the European Universities Association recommendation that institutions create their own repositories to hold the institution's research is adopted, ways to increase levels of awareness amongst researchers and academics will still need to be considered (European Universities Association 2008).

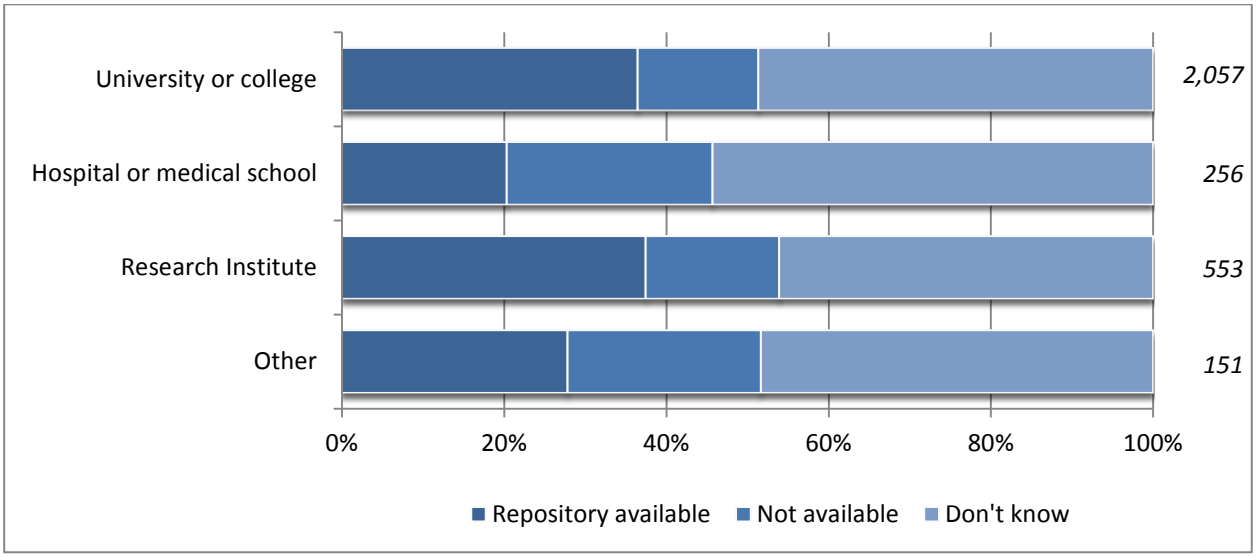
Figure 2.5 Availability of institutional repositories, by institution type



(Weighted total numbers of responses given in italics)

In considering awareness of subject-based repositories by institutional type, researchers from hospitals and medical schools are more likely to be unsure as to whether or not an appropriate subject-based repository exists for their research, than researchers from universities, colleges, research institutes, and other types of institution *(Figure 2.6).* Researchers from universities, colleges and research institutes were more likely to report the existence of a suitable subject-based repository than those from hospitals, medical schools and other types of institution.

Figure 2.6 Awareness of subject repositories, by institution type



(Weighted total numbers of responses given in italics. Owing to the low number of respondents from Government and Industrial/commercial institutions these are included in the 'other' category here)

2.2 Disciplinary similarities and differences

Awareness of OA and OAR is influenced by disciplines and existing OA practices within these disciplines. Medical scientists (together with Life scientists and Interdisciplinary researchers) are more likely to equate OA with an 'author pays' model; this tends to support the findings of Björk *et al.,* (2010) and may be owing to the prominence and visibility of well-known OA journals (such as those from PLoS journals) within this discipline. This may also account for why OA is not associated with the term 'not peer-reviewed' within the Life sciences and the Medical sciences *(Figure 2.1).* It is of note that both Medical and Life scientists tended to be less aware of subject-based repositories available to them – even though one of the most well-known subject-based repositories (PMC) is focused on this domain. Reasons for this are speculative; however the role of journals/publishers in facilitating deposit procedures for this subject-based repository may be contributory.

Perhaps unsurprisingly, given the existence of arXiv, Physical scientists & mathematicians were more aware of subject-based repositories than other disciplinary groups. They were more likely to consider that articles in OAR were not peer-reviewed, and not the final version, than other disciplinary groups. They were also less likely (along with researchers from the Social sciences, humanities & arts) to associate OA with the 'author pays' model.

3 Open Access Repositories in the research process

Studying the role and use of Open Access Repositories in the scholarly research process is a complex matter as it involves looking at researchers' workflow both from a reader perspective, i.e. as a consumer of journal articles, and an author perspective, i.e. as a producer of journal articles. Findings from the PEER Behavioural study suggest that readers and authors show different perceptions, attitudes and use of OAR according to their role at the time and the tasks being undertaken in the research process.

Although open access was not an unfamiliar concept to most participants and respondents in Phase 1 of the PEER Behavioural project, influences and perceptions of the use of publicly available OAR varied greatly. In some cases these differences can be attributed to disciplinary practices, and in others they are seemingly due to individual idiosyncrasies. Important influences highlighted in Phase 1 appear to be perceptions of quality, peer review, confidence, trust and visibility. The phase 2 survey and workshop brought in-depth analysis of the issue in order to untangle the relationship between influences, perceptions and use.

3.1 The use of Open Access Repositories by readers

3.1.1 Readers' information searching behaviour

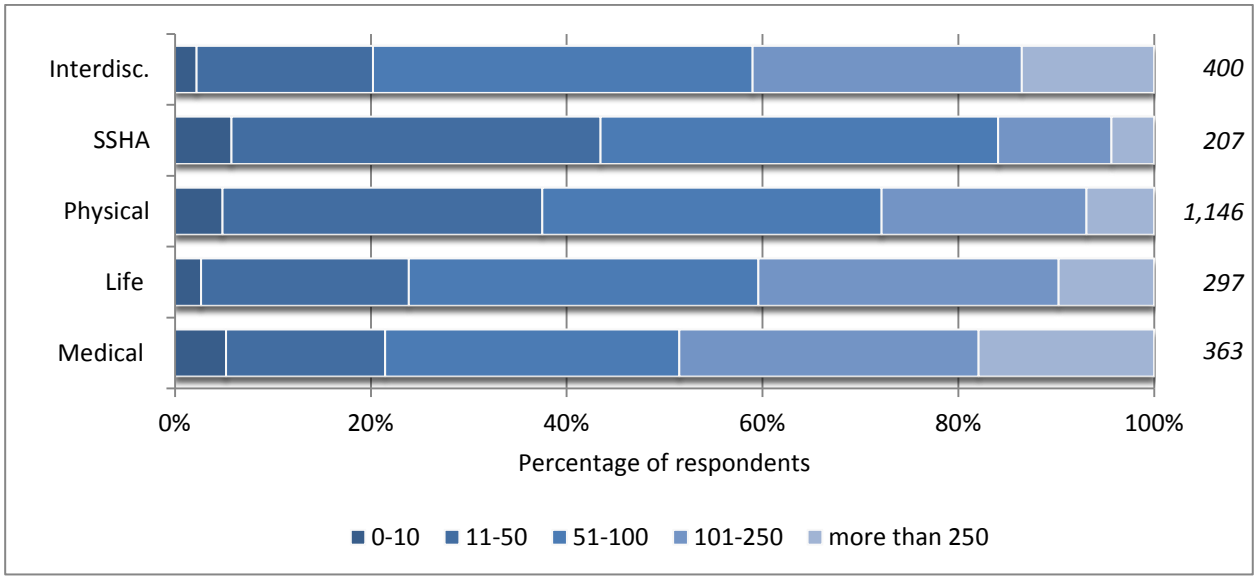
In the phase 1 survey, over 90% of respondents rated peer-reviewed journals as 'very important' to their research, although this result may be biased owing to the method used to contact authors. These findings regarding the importance of scholarly articles in scholars' information landscape corroborate results recently published by Tenopir *et al.,* (2009) and King.*et al.,* (2009) which both found that scholarly articles accounted for over 90% of scholars' information sources.

The importance of journal articles as reported in our phase 1 survey varied by subject, from 95% in the Life sciences to 86% in the Social sciences, humanities & arts rating these as 'very important'. Phase 1 respondents from the Physical sciences & mathematics and from the Social sciences, humanities & arts also highly valued the use of monographs in their information search process. Conferences were more likely to be used by respondents from the Physical sciences & mathematics. Further analysis regarding the relative importance of other types of research outputs by disciplines is provided in the Baseline report *(Section 2.3)* [7].

The importance of journal articles in readers' research practice highlighted in the survey was well supported by the amount of journal articles read by scholars, with over 95% of phase 1 survey respondents reporting that they read, on average, more than 10 peer-reviewed articles per year. Frequency of reading was very high for phase 1 respondents from the Medical sciences and the Life sciences, with almost half of Medical scientists (49%) and 40% of Life scientists reading more than 100 peer-reviewed articles per year, compared to just over one quarter (28%) of Physical scientists & mathematicians and 16% of researchers from the Social sciences, humanities & arts. Researchers from the Social sciences, humanities & arts were more likely to read between 51 and 100 articles per year *(Figure 3.1).*

[7] http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf [accessed 29.09.11]

Figure 3.1 Number of articles read per year



(Weighted total numbers of responses given in italics)

Readers' information behaviours in relation to journal articles were explored further in the phase 2 workshop. Findings suggest that researchers, across disciplines, use only a few sources to search for research information. Readers develop their own information search strategy over time, possibly constructed on the basis of previous trial and error methods. Researchers tend to use the handful of information sources they have used in the past, and expand their search from there. When asked about their 'preferred' sources, the two most frequent research resources used as a starting point by workshop participants, and spontaneously mentioned (i.e. without any prompts), were journal platforms and Google / Google Scholar. This was the case across all workshop disciplines, especially in the Chemistry, Earth, marine & environmental sciences and the Social sciences, humanities & arts groups. These, qualitative, findings corroborate phase 1 survey findings, which indicated that researchers tend to prefer search engines, including Google and Google Scholar, over libraries' gateways (library portal or catalogue) when it comes to identifying relevant journal literature.

Differences between broad discipline groups were reported in the phase 1 survey, with Medical scientists more likely to use subject portals/repositories, Life scientists more likely to use bibliographic databases, and respondents in the Social sciences, humanities & arts more likely to use library resources and search engines, than respondents from the other broad disciplinary groups. Some interesting differences by length of career also emerged from the phase 1 survey; in particular the less experienced researchers (fewer than five years) are more likely to use Google Scholar, while the most experienced (25 or more years) are more likely to browse print journals, than the other groups.

There is some evidence from the workshop that researchers, across disciplines, generally tend to go first to journal platforms. Researchers seem to have their own hierarchy of journals and they usually go directly to the journals they trust, if they have a good knowledge of the field they are searching – researchers indicated that the process could be different if they were new to the field. Earth, marine & environmental sciences participants elaborated further on this, saying that trusted journals, where one goes more often, would be where renowned authors in the field have published and/or where they have had a good experience themselves either as an author or a reviewer. However they emphasised that trusted journals did not necessarily equate to

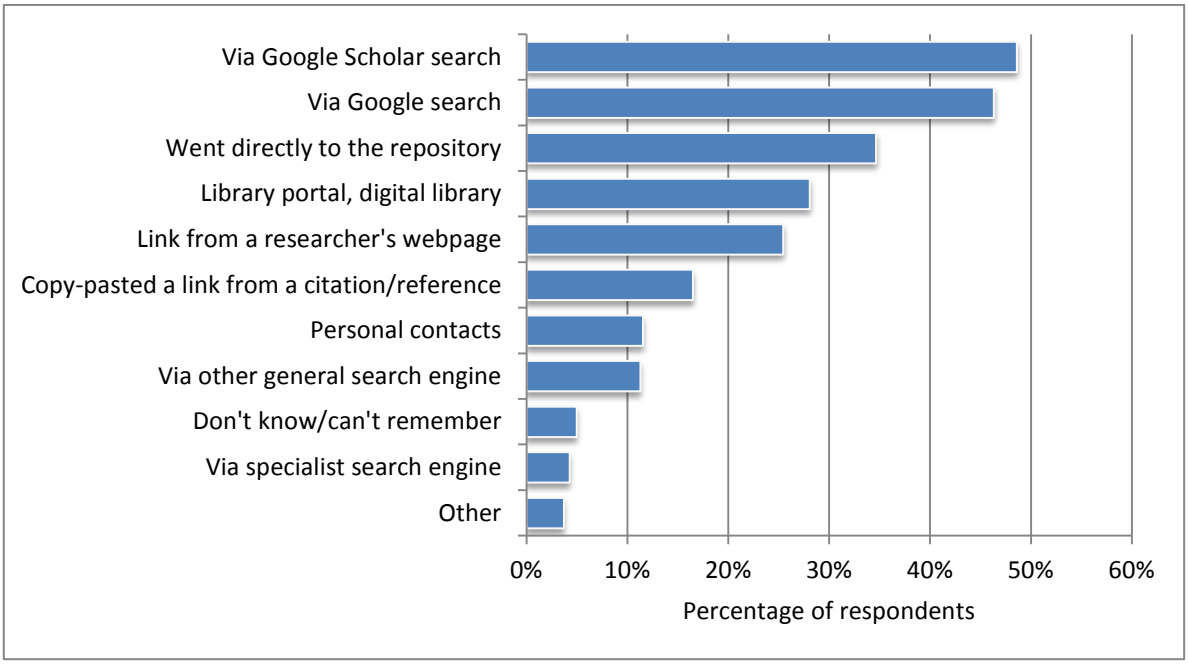
quality, although this was generally a good filter for quality. Most workshop participants, across the four disciplines, were generally ready to accept journal reputation as a guarantee of quality. Impact factor, reputation and quality seemed to be closely intertwined.

When exploring a new research field, workshop participants indicated they would generally try to identify key authors or key journals in the field – though it was suggested that it is difficult to infer quality from a specific journal when one is not familiar with the field. Conversely participants felt that in their own field they know which journals they can trust.

Overall, 72% of phase 2 survey respondents indicated that on the most recent occasion that they had accessed an OAR, they were looking for articles on a specific topic (keyword search), with 69% reporting that they were looking for a specific article, whilst 40% were looking for work by a particular author/research group (these percentages total more than 100 as respondents could tick more than one response to this question). One difference was noted between the subject areas, with respondents from the Physical sciences & mathematics, and the Social sciences, humanities & arts more likely to be looking for work by a particular author/research group than respondents from the Medical sciences.

As outlined in Figure 3.2, almost half of researchers reached the repository via Google or Google Scholar, whilst only 35% reported that they went directly to the repository. Less popular methods of accessing repositories included through specialist search engines, with just 4% of researchers using these. Irrespective of how users arrived at repositories, just over two-thirds of researchers had accessed journal articles held in an OAR in the last year, with 16% stating that they had not accessed such articles, and a further 16% claiming that they were unsure if they had or not.

Figure 3.2 Methods of accessing repositories



Total respondents = 1,012

Searching specifically for OA resources did not really figure in researchers' information searching landscape. The fact that Google and Google Scholar were the most common ways to

arrive at an OAR does not mean that this was the primary destination. Although many workshop participants tended to think of OAR as a place where one goes to find resources, most were aware that OA resources could be retrieved via search engines. The few participants who had this mechanism clear in their minds, thought that what really matters is how well the material is indexed so that it can be retrieved by Google. However, it was said that searches through search engines do not always deliver good results; for example, one often needs to know some words from the title, which makes the process more difficult when one is exploring a topic without any specific articles in mind. It was also indicated that relying on Google to find articles in OAR may not be a satisfactory discovery tool, as one may end up with too much information, especially when the search query is rather vague (e.g. no known authors or title).

3.1.2 Factors influencing readers' use of OAR versions of journal articles

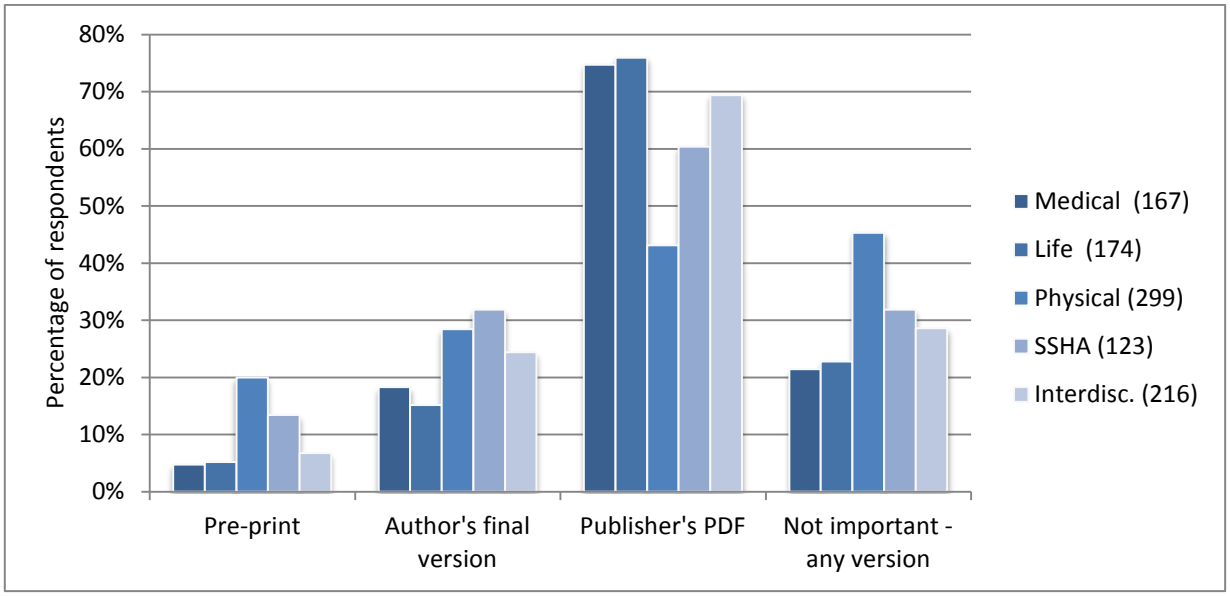
As reported in the previous section, scholars generally rely on scholarly articles to inform their research work, with 90% of phase 1 survey respondents rating scholarly journals as 'very important' for their research. The phase 1 survey data showed that researchers from the Medical sciences and the Social sciences, humanities & arts were most likely to consider the article version important, and least likely to trust documents (other than the published version) they found in repositories.

3.1.2.1 Satisfaction with the quality of the material accessed in an OAR

Sixty-two percent of phase 2 respondents indicated they were looking for the published final version on their last visit to an OAR, whilst 24% stated that they were looking for the author's final version (stage-two accepted manuscript). Interestingly, almost one-third of respondents indicated that it was not important which version of the article they were able to access. Differences were noted by subject area *(Figure 3.3):*

- Researchers from the Physical sciences & mathematics were more likely to be looking for a pre-print than those from the Medical sciences or the Life sciences and Interdisciplinary researchers; this may reflect their long-standing pre-print culture.
- Researchers from the Physical sciences & mathematics, and the Social sciences, humanities & arts were more likely to be looking for the author's final version than respondents from the Medical sciences or the Life sciences. In contrast to this, researchers from the Medical sciences or the Life sciences were more likely to be looking for the published final version (publisher's PDF file) than respondents from the Physical sciences & mathematics, and the Social sciences, humanities & arts.
- Respondents from the Physical sciences & mathematics were most likely to indicate that it was not important which version they found when visiting a repository.

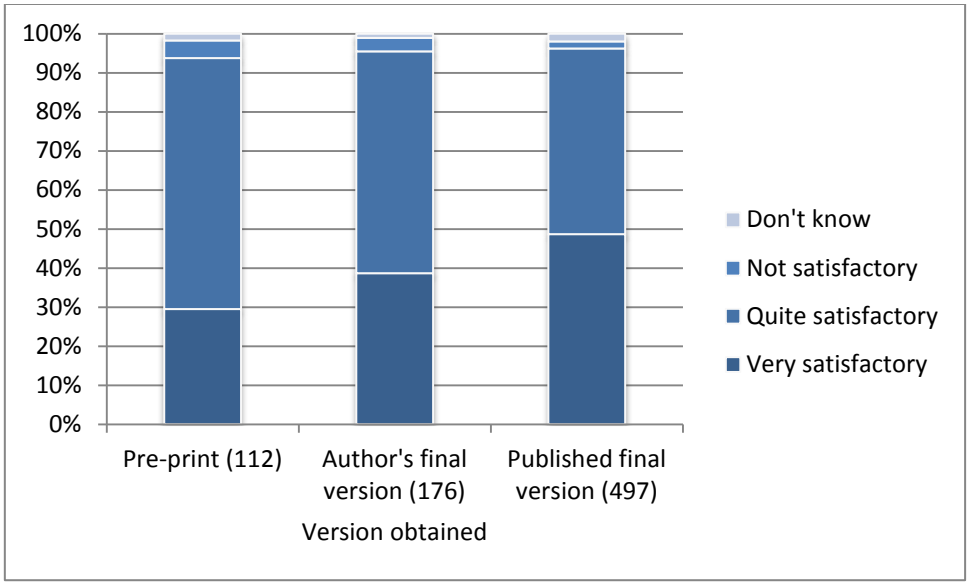
Figure 3.3 Version of the article hoping to find



(Weighted total numbers of responses given in brackets)

In terms of whether what the researchers found in the OAR met their requirements, just over half of the researchers surveyed in Phase 2 indicated that the article they accessed was 'quite' satisfactory for the task they were undertaking, with a further 40% indicating that it was 'very' satisfactory. There is also a clear association between the version found and the level of satisfaction with it *(Figure 3.4).*

Figure 3.4 Satisfaction with the version obtained



(Weighted total numbers of responses given in brackets)

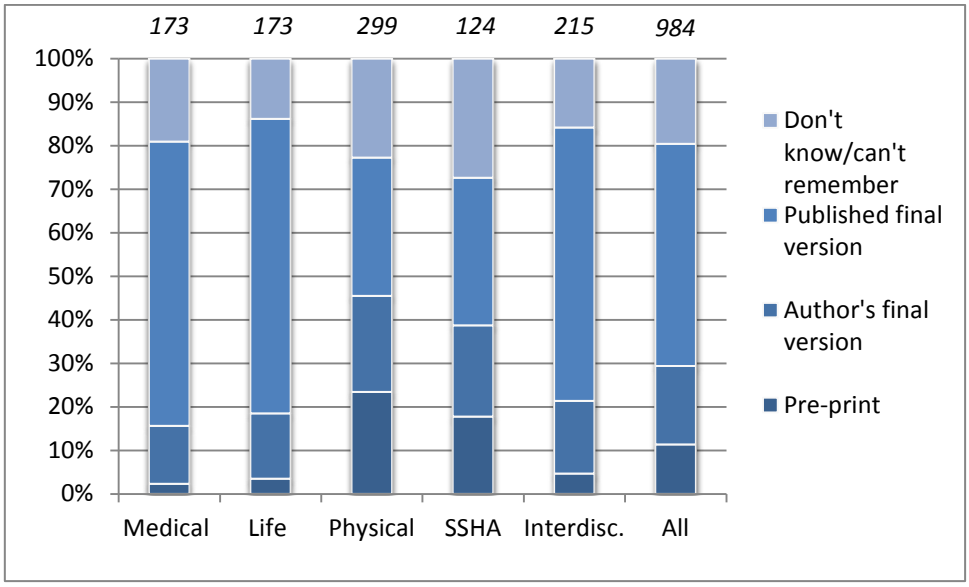
These findings resonate with the idea aforementioned whereby academics take different roles, and so the purpose of reading varies accordingly. This may impact on their acceptance of versions other than the published final version of a paper. This also resonates with findings from studies looking at the main differences observed before and after publisher's copy-editing, i.e. between stage-two accepted manuscripts and their respective published versions. Those

studies revealed that most corrections are made for spelling, grammar or style mistakes, though a certain number of corrections deal with mistakes spotted in the references or missing references (Thatcher, 2011; Wates and Campbell, 2007; Goodman *et al*., 2007). They conclude that the quality of repository content is probably good enough for certain types of academic use, which could be categorised more as public consumption than scholarship, though the question of the perpetration of erroneous references may cause problems further down the line. This resonates with the phase 2 workshop findings, whereby it was suggested that, for readers who are not authors themselves, OAR versions could be satisfactory, in that their reading involves learning about a topic and using this knowledge without the need to formally cite material in scholarly works. However, it is important to reiterate the importance that workshop participants placed on accessing, reading and citing the published version, which remains the version of choice for the purpose of writing a journal article *(also see Section 3.1.2.4)*. Participants across disciplines indicated on many occasions that they highly value publishers' versions as they were thought to be easier to read than pre-print or authors' final accepted manuscripts, especially when it comes to graphs, tables and pictures.

Evaluating the suitability of an article for the task they were undertaking did not seem to be problematic, with 64% of researchers finding the article they accessed either 'easy' or 'very easy' to evaluate. Workshop participants indicated that they often found it difficult to ascertain which version of the paper they had accessed from OAR platforms. This is in line with phase 1 survey results, indicating that 16% of respondents could 'rarely' or 'never' identify the article version from the repository information provided, while 31% could do so only 'sometimes', and 20% did not know.

Half of all researchers reported finding the published final version on their most recent visit to a repository, although this varied by discipline *(Figure 3.5).* It is unclear whether such a high rate of published versions reportedly found in repositories is representative of the version content held in OAR worldwide, although the disciplinary differences shown in Figure 3.5 are in line with known disciplinary OA practices. This raises the question of whether researchers conflate Open Access Repositories with Open Access Journals, or whether the clarity of the metadata for the version accessed in OAR is not sufficient to enable readers to distinguish between an author-produced PDF of a stage-two accepted manuscript and the publisher-produced PDF of the published final version of the article. The main visible distinction between these two versions generally relies on the publisher's branding, which may, on some occasions, be overlooked by readers. It is also possible that readers never see the repository metadata page - there is some evidence (reported by repository download statistics) that they may be directed from the search engines directly to the full text document, despite repositories' efforts to bring their readers to the metadata page first (Organ 2006).

Figure 3.5 Version found, by subject



(Weighted total numbers of responses given in italics)

The clarity of repository metadata, as well as readers' ability to distinguish between the different versions of an article (pre-print; author's final accepted version; published final version), was also discussed in the phase 2 workshop, and findings show that readers' perceptions seemed to differ greatly at the individual level. Some disciplinary themes emerged – although it is difficult to know how representative they are of the wider discipline – with participants in the Social sciences, humanities & arts group being of the opinion that it was not always clear which version they were accessing from a repository. It was also indicated that the process of identifying and ascertaining whether the paper accessed was a pre-print or an author's final accepted manuscript could sometimes take readers quite some time – though it was recognised that this may differ from one repository to another, and from one individual to another. As for distinguishing between an author's accepted manuscript and a published final version accessed online, it was felt, especially in the Chemistry group, that the distinction was generally fairly clear by the look and feel of the article (publisher branding on the published article) or by looking at the URL. To illustrate what they meant by 'look and feel' one commented "*on the publisher's website it's always nice, but on the repository it's not…*". The issue of clarity of the bibliographic records held in OAR seemed to be centred on the distinction between pre-refereed papers and post-refereed papers. The quality of OAR was also raised at the workshop, mostly by Social sciences, humanities & arts participants, in that the quality of OAR was also thought to depend on the clarity of the metadata, thus enabling readers to know exactly which version of a paper they are about to access. The issue of repositories' reputation briefly came up in the Earth, marine & environmental sciences group, where a participant thought that it was important that OAR are authoritative, and that authority starts with quality metadata and research outputs. Indeed, the lack of clarity about whether or not the paper accessed has been through peer review seems to be the biggest obstacle to using OAR sources. It emerged from the discussion that a standardised way of presenting different versions of an article would greatly help readers in ascertaining whether the version accessed is satisfactory for the purpose of the task at hand and harnessing trust and confidence in the use of repository content.

3.1.2.2 Confidence issues

The issue of readers' confidence in using repository versions of journal articles was raised in phase 1 focus groups, which showed some minor disciplinary differences in relation to the acceptance of the different versions of a published paper. There was a greater acceptance of pre-prints in the Physical sciences & mathematics than in the other disciplines. In the Medical sciences, knowledge of the author was key to acceptance of, and trust in, pre-prints, although the reputation of the journal was equally important when considering peer-reviewed articles in OAR. Social sciences, humanities & arts and Life sciences focus group participants placed great emphasis on peer review and journal reputation to assess the quality of research articles. In the Physical sciences & mathematics focus group, there was a dichotomy between those who used repositories and those who did not, and emphasis was placed on the users' own critical judgement rather than a peer-review 'stamp'.

This issue of confidence in versions other than the published version was also captured in the phase 2 survey, in which readers indicated they were most likely to be looking for the published final version of a journal article whilst using the OAR. Almost one-third, however, indicated that it was not important to them which version of the article they were able to access *(Figure 3.3, page 16)*. Within these overall results there were broad disciplinary differences. Unsurprisingly, these seem to reflect the repository infrastructure available in each discipline, with researchers from the Physical sciences & mathematics again more likely to be looking for a pre-print than those in the Medical sciences or the Life sciences, or those categorised as being Interdisciplinary researchers, whereas researchers in the Medical sciences or the Life sciences were more likely to be looking for the definitive published version than respondents from the Physical sciences & mathematics or the Social sciences, humanities & arts.

3.1.2.3 Peer review and the purpose of reading

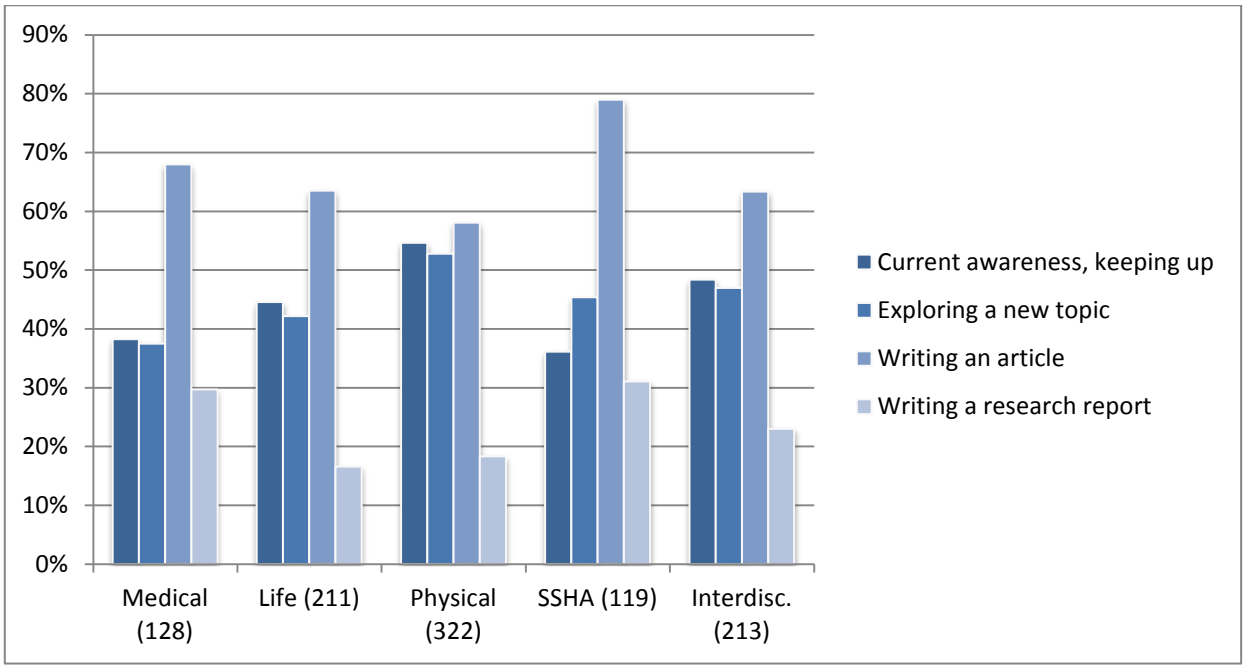
Phase 2 of the project brought some interesting results as to which factors are contributing to the use of repository versions of journal articles. The issue of validation of the results through the process of peer review was a recurrent argument in the workshop discussion about readers' use of OA articles in their research process. A great majority of participants recognised that they would not trust a repository version of an article as much as the published article itself. Only one workshop participant, from the Chemistry group and a regular user of arXiv, thought that it did not really matter where one gets a paper from, whether from an OAR or a subscription-based journal, as judgement is often made on the reputation of the authors, and, ultimately, when reading the content. Another chemist, with a biology background, thought it was not acceptable to 'look up' OA versions as one cannot be sure whether they have been through the peer review process or not. There was also mention of the reluctance to trust the content of pre-prints, as there can sometimes be considerable changes between the pre-print and the published final version.

There was some evidence in the workshop findings that, for researchers exploring a new field, peer review was an important indicator of quality and would guide them in their information search. It was thought that because they were unfamiliar with a field, researchers should only seek peer-reviewed papers as they felt they did not have the requisite knowledge to make a sound judgement about the validity of the results. The Social sciences, humanities & arts group felt that peer review provided confirmation that the research published is 'good'. Although not all participants shared this view, it was noted that the feeling that repository material is generally not peer-reviewed is an enduring one amongst researchers, which is yet another hurdle to

readers' use of OA sources, be it OAR or OAJ. Others, in the Social sciences, humanities & arts group, thought OAR versions, for example working papers in economics, were a good source to learn about the issues on which peers are currently working, what is likely to be published in the near future and generally to get an idea about how the field is evolving. The tension between accessing validated results and quick access to scholarly literature was captured in the workshop discussion when participants were asked to consider whether there would be occasions when they may consider looking up OAR versions of a paper. Although it was suggested, in the Chemistry group, that a researcher already familiar with a topic may want to look up OAR versions when working on a 'hot topic' (i.e. a fast moving field) or when browsing for current awareness, most participants were happy to wait for the published final version. Some were adamant they would not read pre-prints at all but would consider reading authors' final accepted papers – though they insisted they would be happy to wait for the published final version.

The phase 2 survey found that over 60% of researchers reported that, on their last visit to an OAR, they were in the process of writing an article, nearly 50% were exploring a new topic or browsing literature for current awareness, and nearly 30% were in the process of writing a research proposal. Unsurprisingly, the fact that writing an article was the most common task on researchers' last visit to an OAR echoes studies of journal article reading showing that reading is positively correlated with success (King *et al*., 2009; Tenopir *et al*., 2009; RIN, 2009), using an academic's publication rate as a proxy for success. In terms of the purpose for which readers were visiting a repository, the phase 2 survey highlighted some broad disciplinary differences *(Figure 3.6).* Readers from the Social sciences, humanities & arts were most likely to have arrived at the OAR for the purpose of writing an article, whereas readers from the Physical sciences & mathematics were more likely to visit for the purpose of current awareness, than readers in the Medical sciences or the Social sciences, humanities & arts.

Figure 3.6 Reasons for visiting an OAR, by subject



(Weighted total numbers of responses given in brackets)

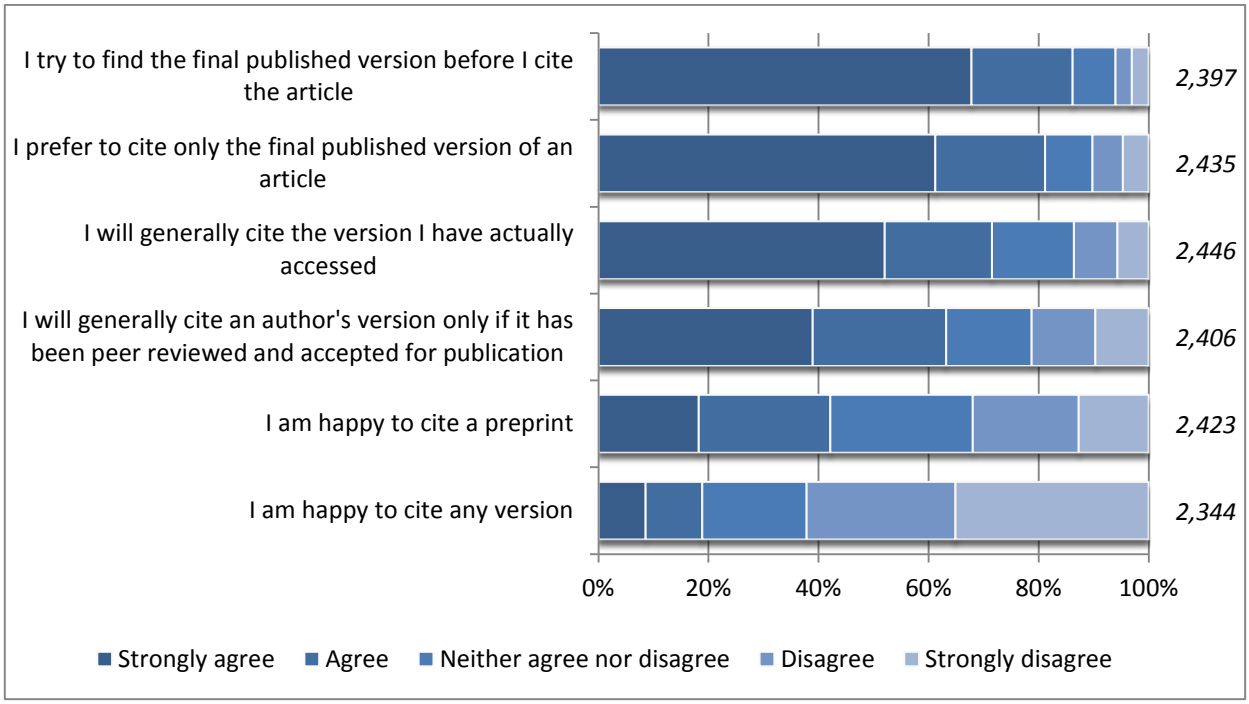
The importance of purpose when accessing materials held in OAR was well-supported by workshop findings, which showed that scholars seem to hold different attitudes towards repository versions of published articles according to the purpose of their reading. Workshop participants tended to describe themselves as wearing different hats in their research process (commercial as opposed to academic research; writing an article as opposed to current awareness).

Opinions diverged greatly between individuals with regard to exploring and searching a new disciplinary field, although there were no clear disciplinary patterns emerging from the workshop findings; attitudes depended very much on how confident one felt in exploring new fields of research. It was indicated, notably by engineers, that one could happily use author's final accepted manuscripts for pieces of work conducted for the commercial sector, but they would definitely need to access the published article for any academic research work. In the Engineering group it was also suggested that repository versions could be used at early stages of the research when one wants 'to find out' about a new field, but again at later stages (when writing up the research) they would require access to the published final version. This latter practice was not universally shared by participants.

3.1.2.4 Use of repository material as it relates to citation practice

Another point of importance in relation to the version researchers access and use was the perceived difficulty to cite a version other than that published in their own research, i.e. when they write up research with a view to publishing it in a scholarly journal. This was highlighted in both phases of the study. Indeed, whilst reputable repositories were recognized as important sources of material, it was clear that when citing articles, researchers in all disciplines preferred to cite the published final version rather than any other version. From the phase 1 survey, over 80% of researchers indicated that they prefer to cite only the published final version of an article and over 85% indicated they would try to find the published version before they cite the article. Interestingly, 60% of researchers also indicated that they would cite an author's version as long as it had been refereed *(Figure 3.7).*

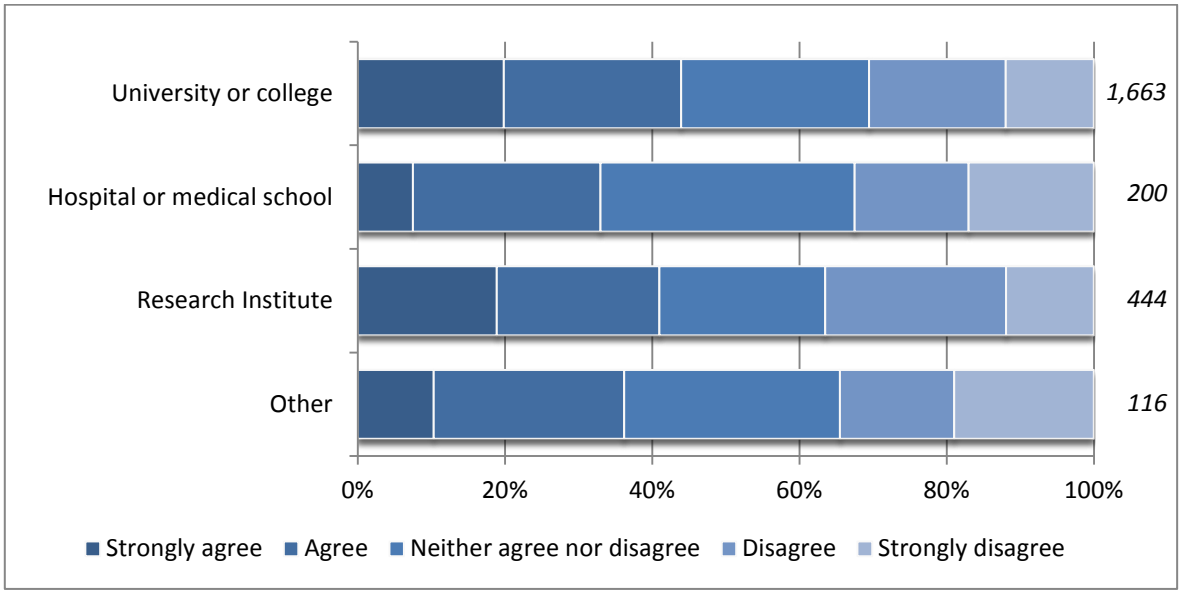
Figure 3.7 Citation preferences



(Weighted total numbers of responses given in italics)

Citation preferences amongst phase 1 survey respondents varied according to institution type, with some differences in levels of willingness to cite unpublished versions of journal articles. Researchers from hospitals, medical schools, and other types of non-HE institutions were less-likely to 'strongly agree' that they would be happy to cite a pre-print, than respondents from universities, colleges or research institutes *(Figure 3.8).*

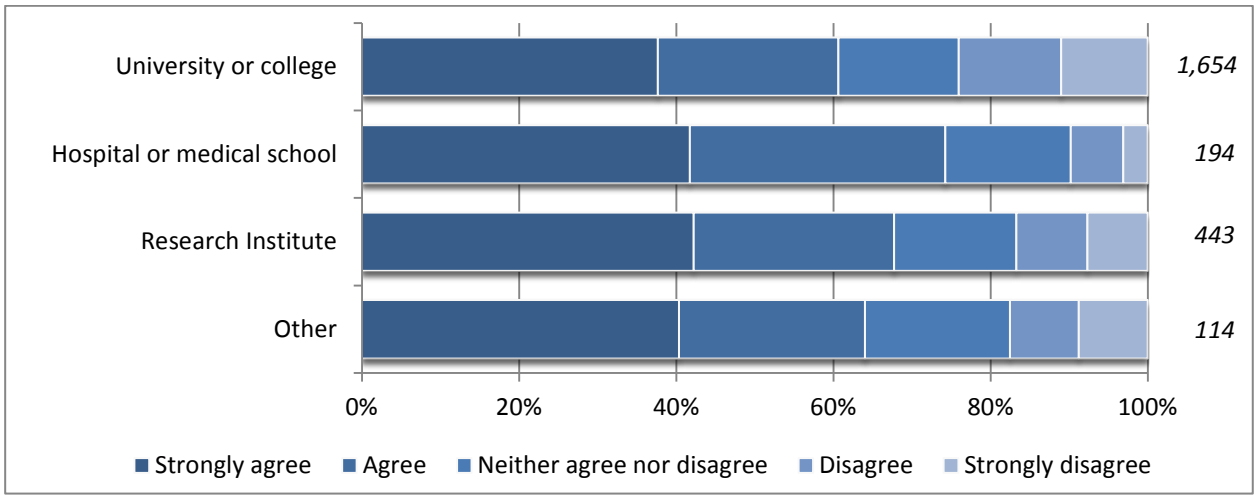
Figure 3.8 Willingness to cite a pre-print, by institution type



(Weighted total numbers of responses given in italics. Owing to the low number of respondents from Government and Industrial/commercial institutions, these are included in the 'other' category here)

Respondents from hospitals or medical schools were more likely to indicate that they 'agreed' or 'strongly agreed' that they would only cite an author's version if it had been through the peer review process than respondents from universities or colleges. However, around 40% of all researchers, regardless of institution type, indicated that they strongly agreed with this statement *(Figure 3.9).*

Figure 3.9 Willingness to cite an author's version only if it has been peer reviewed and accepted for publication, by institution type



(Weighted total numbers of responses given in italics. Owing to the low number of respondents from Government and Industrial/commercial institutions, these are included in the 'other' category here)

The fact that respondents showed a strong preference for citing the published final version was well supported by the workshop findings, which again showed that researchers generally find it extremely difficult to cite anything other than the published final version when writing a journal article. Various reasons were put forward to explain authors' preference for citing the published final version, but there did not seem to be any differences between disciplines, which suggests that this is more related to the general conduct of academic research than disciplinary norms. The notion of trust in the validity of the results was again put forward to explain the need to cite the published final version, even in the hypothetical case that one has consulted a repository version at an earlier stage of the research process.

In addition, some concerns were voiced, across the disciplines, about the technical difficulties pertaining to citing an OA version. The issue of the absence of page numbering for repository versions seemed to be a major problem to citing these, although some workshop participants recognised that there were ways to find the exact page numbering without accessing the published final version. The phase 1 focus groups indicated that the citation policy of the journal publishing the article may also have an influence on authors' citation behaviour and choice of version. Participants generally agreed that citation practices depended on the purpose of the piece of work they were working on.

Workshop participants generally indicated that they would be happy to cite versions other than the published articles if the purpose was not to write an academic journal article. As an illustration, an Engineering workshop participant with a commercial background indicated that using and citing an author's final version would be acceptable when preparing a report for a commercial client, but that if a scholarly article was to result from the work, then access to (and citation of) the published final version would be necessary as it was felt that the published

version has a ‘purity’ about it that is not always strictly necessary when preparing commercial reports. This sentiment seemed to be shared across disciplines. A Social sciences, humanities & arts participant indicated that it may also happen that one uses non-peer-reviewed material (such as pre-prints) when working in a new field for which not many papers have yet been published. There was also a consensus in the Earth, marine & environmental sciences group, that if an article has not been published it cannot be cited. Others thought that one can always put the OA version when writing the paper as an interim measure – “*by the time you write your own paper, the pre-print will be published in a journal*” (Chemistry participant).

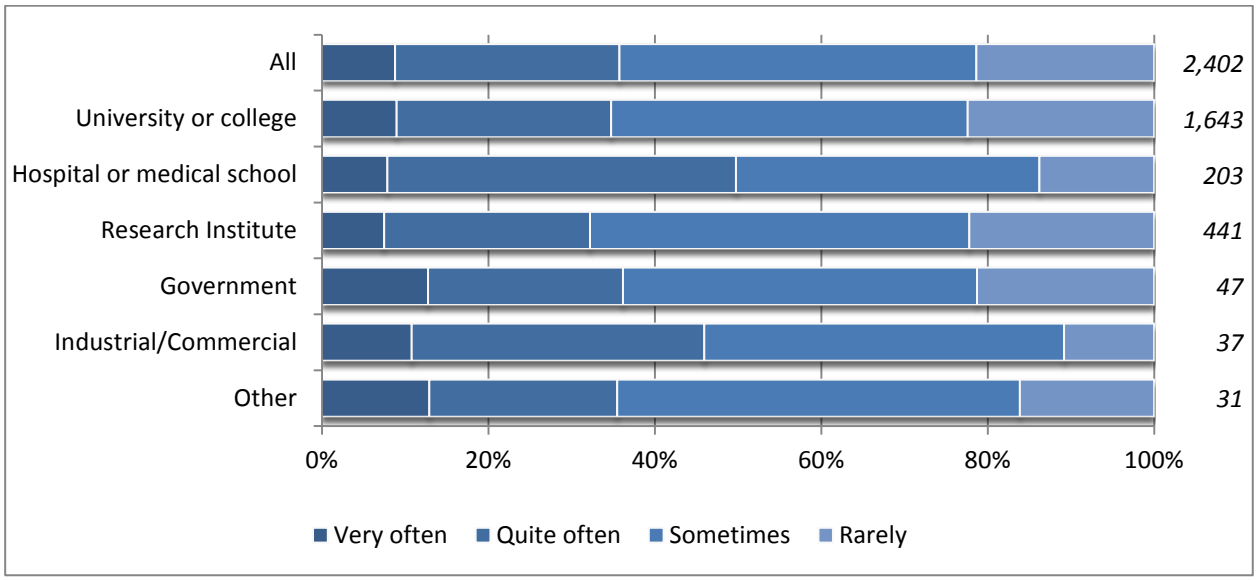
Similar comments were made by a participant in the Earth, marine & environmental sciences group, who indicated that the use of OA author’s final drafts would depend on the publication lag of the journal targeted for submission. For instance, for slow publication journals this participant would access and cite repository versions as, by the time the paper gets published, the unpublished paper would have been published too. Conversely, for fast moving fields, it was felt that only published articles could be used in the process of writing up the research.

Workshop participants generally thought it was not good practice, nor common practice, to cite OA material for several reasons (in no specific order): firstly, it is often required by the publisher to provide the citations’ DOIs when submitting an article; secondly, OA material can be seen as ephemeral, whereas published final versions are seen as permanent. Finally, there was a feeling that repository versions tend not to be cited in important indexes such as the Science Citation Index [Web of Science]. The Social sciences, humanities & arts group elaborated on this specific point and expressed the view that authors may worry about how their readers may perceive citations of versions other than the published article. It was felt in the Chemistry group that it was probably better to put the citation with the ‘in press’ mention rather than an OAR version, even if it is the OAR version they have actually accessed and read.

3.1.3 Access to information sources

From the perspective of readers, an inability to access journal articles was identified in the phase 1 survey, with only 22% of researchers stating that access to peer-reviewed journal articles was ‘rarely’ a problem. Figure 3.10 shows that more than three quarters of all researchers reported being unable to obtain quick and easy access at least ‘sometimes’ to a peer-reviewed journal article which they had identified, with 9% reporting that this happened ‘very often’, and 26% ‘quite often’. Note that these findings are based on researchers' perceptions and should be considered in relation to the amount of reading that they do, although there was no statistically significant association between the amount of reading and difficulty of access found in the phase 1 survey.

Figure 3.10 Frequency of lack of quick and easy access to peer-reviewed articles, by institution type



(Weighted total numbers of responses given in italics)

The workshop revealed that the extent of the 'access problem' highlighted in Phase 1 seemed to vary from one individual to another, and more generally seemed to differ at institution level. No significant differences between the four disciplines making up the workshop audience were noted. Although all workshop participants felt that they were, to some extent, experiencing problems in accessing all the journal literature they need for their research, they found it particularly difficult to give more detail regarding this issue. No clear and identifiable pattern emerged from the discussion; however, in an attempt to quantify the perceived access problems identified in the phase 1 survey, the Chemistry workshop participants mentioned that difficulties occurred 'almost daily' or 'daily' in some cases, whilst the Engineering group indicated that approximately between 20% and 30% of the material they retrieved could not be accessed through their institutions' journal subscriptions.

Other points of interest in relation to the issue of access to scholarly journal articles included the following:

- Without any prompt from the research team, the problem of accessing older articles was mentioned independently in three workshop groups out of four, which suggest that this is a common problem across disciplines.

- Workshop participants generally felt that they had more limited access to resources than they used to; this could infer library budget cuts as they thought their access to subscription-based resources was shrinking. Despite the fact that workshop participants indicated a strong preference for the published version, both in terms of reading, trust and citation, some of them thought that there will come a day when there will be a switch in favour of OA versions (pre-prints or authors' final drafts) and researchers will use these more as they become more accessible than the published versions.

- Researchers from the commercial/industrial sector present at the workshop seemed to experience great difficulties in accessing research articles, supporting the findings of the phase 1 survey *(Figure 3.10, page 24)*, which show that the percentage of respondents

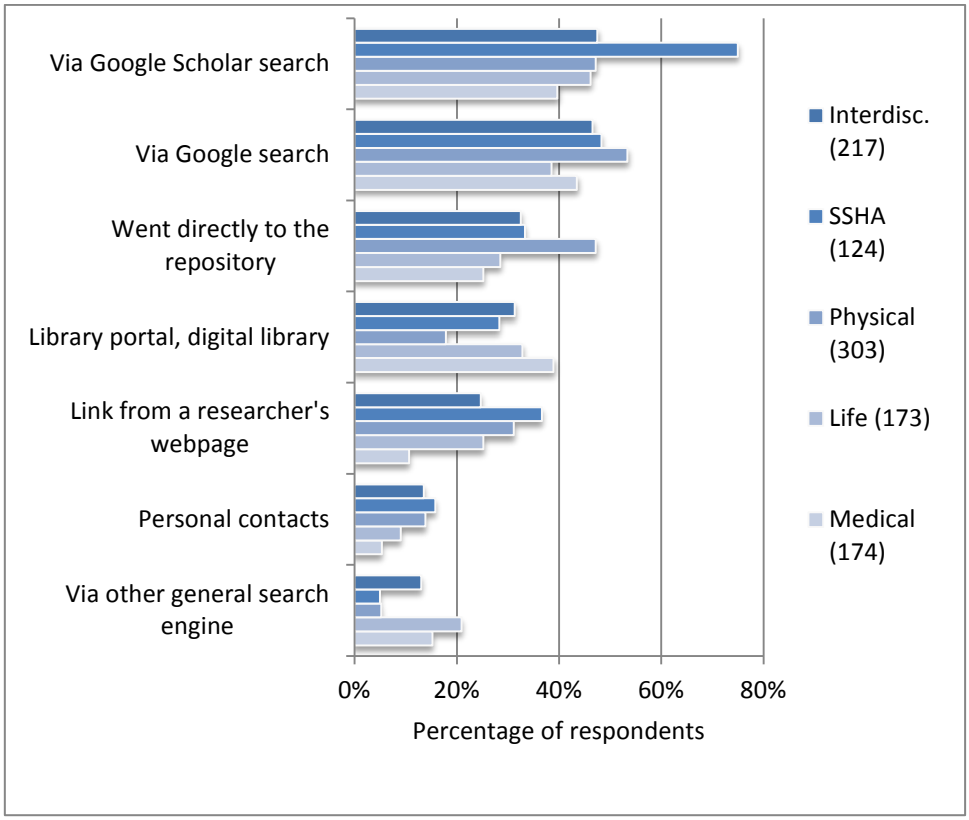
from industrial/commercial institutions reporting that they ‘very often’ or ‘quite often’ experience a lack of quick and easy access to peer-reviewed articles is above the average for all researchers. The industrial/commercial sector researchers at the workshop suggested that these difficulties were owing to publishers’ perceived high prices. In order to overcome this problem, they tended to look for parallel Higher Education affiliation or to identify OA sources via Google. An issue was raised by a commercial researcher about licence restrictions and commercial research; insofar as licences often forbid the use of a resource for commercial research, while at the same time universities conduct commercial research as well as academic research.

3.1.3.1 Alternative ways of accessing material

Phase 1 findings indicated that researchers from all disciplines were likely to seek an OA version of an article when they cannot access a journal article they have identified during their information search, though this behaviour was more noticeable amongst researchers in the Physical sciences & mathematics. Interestingly, physical scientists seemed slightly more prepared to forego using a resource if it is not easily accessible to them – this may be a consequence of having a well-developed culture of OAR deposit. The use of OAR to access journal articles in the past year was reported by just over two-thirds of the researchers surveyed in Phase 2. Figure 3.2, page 14, shows that Google and Google Scholar search engines were the most popular means of accessing OAR, which echoes previous studies (Organ, 2006), whereas 35% reported that they went directly to the repository. Except for the case of well-established subject-based repositories such as PMC, arXiv, RePEc or SSRN where researchers tend to go directly to the repository platform to find articles, results highlighting entry via search engines reflect the current repository infrastructure.

There was disciplinary variation in the ways in which OAR materials were accessed, illustrated in Figure 3.11. In particular, researchers from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to go directly to the repository than researchers in the Medical sciences and the Life Sciences. Researchers from the Medical sciences and the Life sciences are more likely to have accessed the repositories they used through a library portal or digital library than researchers in the Physical sciences & mathematics or the Social sciences, humanities & arts. Researchers from the Social sciences, humanities & arts were most likely to use Google Scholar.

Figure 3.11 Methods of accessing repositories, by subject



(Weighted total numbers of responses given in brackets)

Across all disciplines, a range of other approaches was adopted by researchers seeking articles, from using well-defined institutional channels to less formal peer-based channels. Workshop findings enabled the research team to draw a clearer picture of alternative options available to researchers when they cannot get access to a subscription-based source.

Inter-library loans, when available to researchers, represent the formal route to access material to which the institution does not subscribe, and seems well-used across disciplines; however it does not seem to be the preferred route. Asking the author directly for a copy of the article seems to be common practice across disciplines when one cannot get access to a journal article. Other ways to access articles through informal networks included asking peers on a mailing list whether somebody could provide the article or colleagues at other institutions. Participants also mentioned going to another institution to get access to articles – though it was not always clear whether this was meant as a double affiliation or as a walk-in user. Researchers seem to use informal networks frequently and easily to exchange articles, both as readers to overcome a lack of journal subscriptions and as authors for informal review of work in progress papers. The use of informal networks was unanimously thought to be a well-trialled and satisfactory practice by workshop participants. Although they suspected this may contravene publishers' copyright, they did not seem to attach much importance to this, often claiming this was justified by the (precedence of the) need to access the research. As an illustration of how well informal networks seem to work, a few workshop participants queried why we needed OAR or OAJ in the first place, as it was easy and simple for readers to ask the author for a personal copy of the paper if it happened they could not access it via their institution's journal subscription portfolio.

3.1.4 Disciplinary differences and similarities

There were disciplinary differences in the way in which readers arrived at repositories. Figure 3.11 above illustrates these differences according to the phase 2 survey results. The differences include:

- Readers from the Physical sciences & mathematics were more likely to have gone directly to an OAR than readers from other disciplinary groups, and were less likely to have accessed an OAR through a library portal or digital library than readers in other broad disciplinary groups.
- Readers from the Social sciences, humanities & arts are more likely to have accessed an OAR via a Google Scholar search, than readers from the other broad disciplinary groups.
- Readers from the Life sciences and the Medical sciences were more likely to have accessed an OAR through a general search engine other than Google/Google Scholar, than readers from the Physical sciences & mathematics or the Social sciences, humanities & arts.
- Readers from the Social sciences, humanities & arts were more likely to access an OAR through a link from a researcher's web page or through personal contacts than respondents from the Medical sciences.

Readers from the Medical sciences and the Social sciences, humanities & arts considered the article version found to be important, and were least likely to trust material, other than the published final version of journal articles, that they found in repositories *(Figure 3.3, page 16).*

There was a greater acceptance of pre-prints amongst readers in the Physical sciences & mathematics than in the other disciplines. For readers in the Medical sciences, author and journal reputation usually overcame their lack of trust in pre-prints or articles found in OAR. Social sciences, humanities & arts and Life sciences focus group participants placed great emphasis on the peer review system and journal reputation to evaluate the quality of research articles. Amongst Physical sciences & mathematics focus group participants, there was a dichotomy between those who used repositories and those who did not, and emphasis was placed on readers' critical judgement, rather than the peer review process.

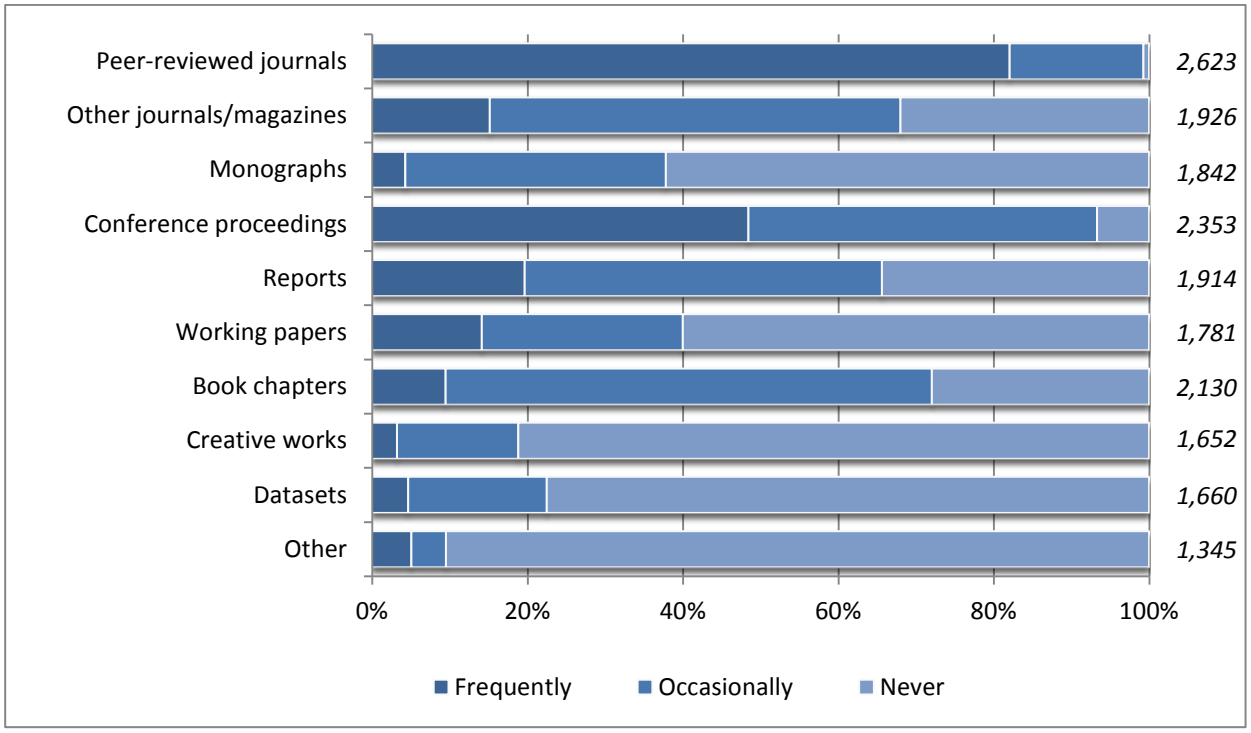
3.2 Authors' open access behaviour and self-archiving practice

3.2.1 Authors' dissemination practice

An overview of dissemination practices is given here to give some context to the use of OAR and authors' self-archiving behaviours.

The predominant mode for disseminating research was peer-reviewed journals, with over 99% of phase 1 survey respondents having published at least one journal article in the last five years. Figure 3.12 illustrates the frequency with which a variety of publication/dissemination formats were used. The formats are listed in order of importance, according to respondents' selection of the most important type of output in terms of their career.

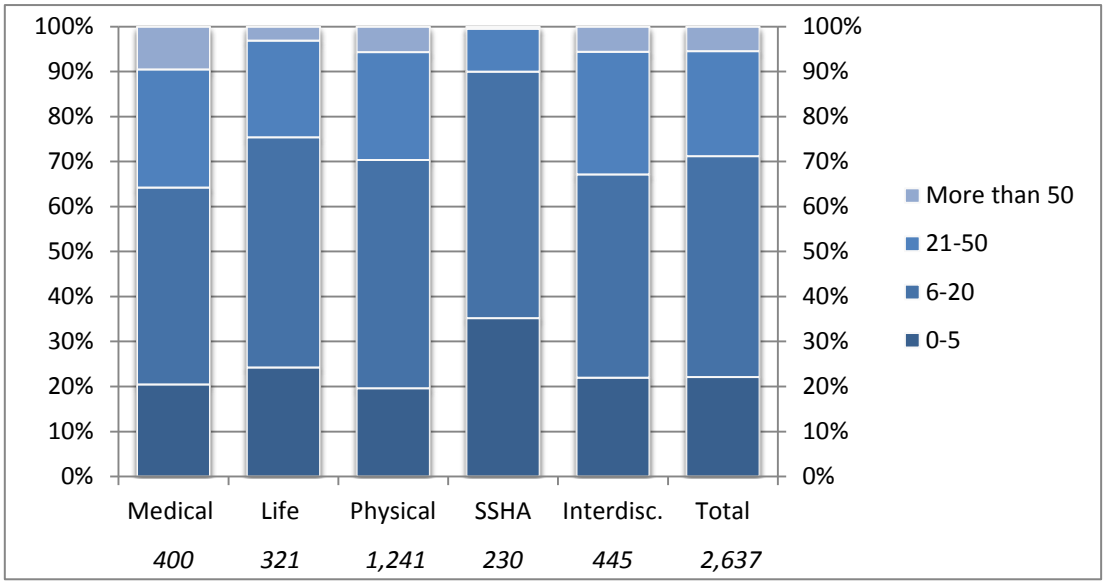
Figure 3.12 Preferred dissemination formats



(Weighted total numbers of responses given in italics)

Social science, humanities & arts researchers are arguably more likely to publish monographs and book chapters than researchers in other disciplines, and compared to researchers in the other broad disciplinary groups published the fewest articles on average, but nonetheless over 98% of respondents in this category had published a peer-reviewed article. Researchers from the Medical sciences were likely to have published the highest number of articles on average *(Figure 3.13).*

Figure 3.13 Number of journal articles published in the last five years



(Weighted total numbers of responses given in italics)

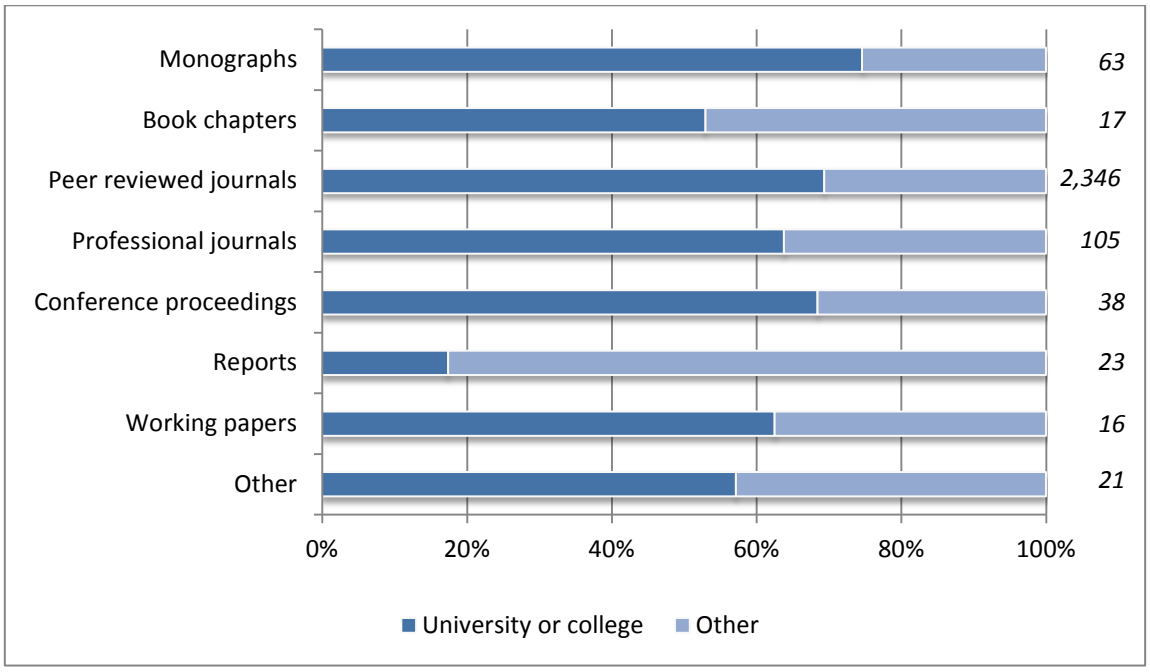
Whilst it could be argued that the sampling method used for the phase 1 survey was responsible for the high number of respondents who had published at least one journal article in the last five years, findings from related studies also confirm the continuing importance of journal articles, even in those fields where monographs have traditionally played a central role in the scholarly communication system (RIN, 2009).

The phase 1 survey findings indicated that career stage is an important influence for authors in choosing how and where they disseminate their research. This was also evident in the workshop findings, which highlighted the tensions between speed of publication, readership and journal prestige.

Focus group participants whose research areas demonstrated strong interdisciplinary characteristics highlighted that the limited availability of peer-reviewed journals in their areas was problematic, particularly in terms of submitting papers for publication. Institutional or subject based repositories were seen as a potential solution to this problem; however, participants emphasised that, despite the limited opportunities, publishing papers in reputable peer-reviewed journals was a priority when considering how and where to publish their research.

Although journal articles were the most common type of research outputs produced by workshop participants, differences were noted according to the extent to which research had industrial/commercial application. For instance, researchers conducting commercially-oriented research tended to produce reports as the primary research output, with journal articles being complementary to these reports. Findings from the phase 1 survey support this anecdotal observation, with over 80% of researchers who indicated that reports were an important dissemination output in terms of their career were from non-HE institutions, with just under 20% from universities or colleges *(Figure 3.14).*

Figure 3.14 Preferred dissemination formats, by institution



(Weighted total numbers of responses given in italics. Owing to the low number of respondents all non-HE institutional types (Hospital or medical, Research institute, Government and Industrial/commercial) have been included in the 'other' category here)

A mathematics workshop participant indicated that pre-prints and drafts were often put out as a forum for discussion in order to refine papers to be submitted to journals. This was notably the case in fields with a limited number of experts worldwide. This participant thought that informal communication mechanisms were working well and therefore OAR were not really perceived as needed or filling a gap in scholarly communication.

3.2.1.1 The underlying pressures in the conduct of research

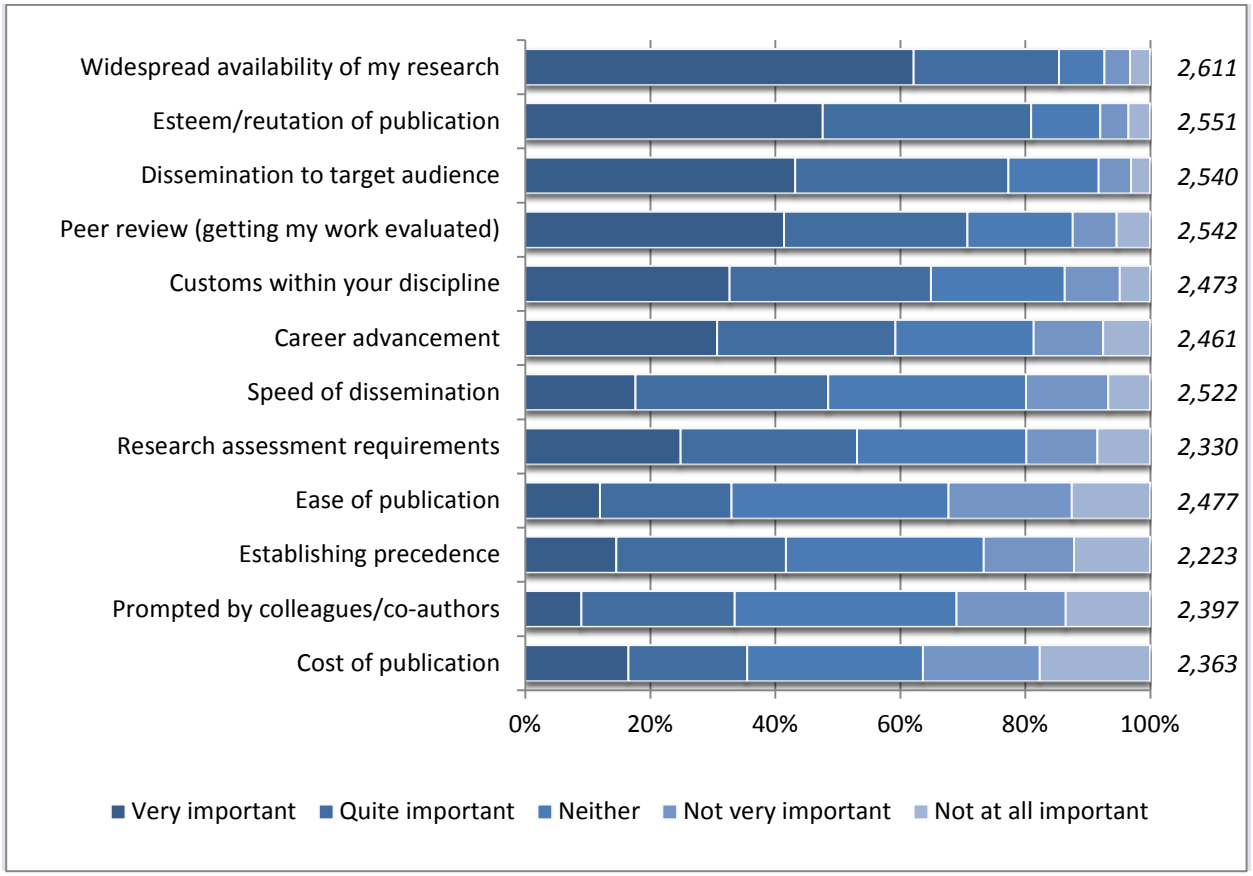
The workshop highlighted two important pressures faced by researchers in the conduct of their research, which have important consequences in the way they disseminate the findings of their research. These are detailed below:

- **Publish or perish state of play**

It was clear across the four disciplinary groups present at the workshop that publishing articles is an important part of researchers' work. The phase 1 survey found that peer-reviewed journal articles were the most important type of research output for almost 90% of researchers. Half of phase 1 survey respondents reported publishing between 6 and 20 articles in the previous five years, and a further 23% had published between 21 and 50 papers during that period. Although there were not enough survey responses from the non-Higher Education sector to give a statistically significant breakdown of preferred dissemination formats by detailed institution type, Figure 3.14 *(page 29)* supports the anecdotal evidence from the workshop indicating that this is likely to vary according to the type of institution a researcher belongs to. Commissioned reports, in particular, feature as a key output for non-Higher Education researchers, with peer-reviewed journal articles being a desired, but not necessarily essential, by-product of their research outcomes.

In the phase 1 survey, respondents were asked to rank the importance of a range of factors in their choice to disseminate in a peer-reviewed journal *(Figure 3.15)*. The most important factors were identified as the widespread visibility of the research, the reputation of the journal, and dissemination of the research to a specific target audience. The least important factors were the cost of the publication, being prompted by colleagues and co-authors, establishing precedence of ideas and the ease of publication. For the least experienced researchers (those with fewer than five years' experience) career advancement was considerably more important.

Figure 3.15 Factors affecting choice of peer-reviewed journals



(Weighted total numbers of responses given in italics)

In addition, workshop participants identified less obvious reasons, such as institutional pressure (i.e. research quality assessment) or, to a lesser extent, pecuniary inducement. It was reported that, at some institutions in some European countries, researchers are paid extra money for each article published – the inducement may vary according to the quality of the publication, (i.e. journal prestige). Some participants, notably in the Engineering group, felt that researchers were 'chasing' publications, mostly for recognition and reward, and this created a situation where there were too many articles published, thus 'diluting' research. It was mentioned that this was reinforced by the expectation of institutions for their researchers to produce a certain number of articles every year.

- ### Impact factors

In terms of journal impact factor, phase 2 workshop participants gave a clear indication that this played an important role in deciding where to publish their research. These participants felt that their reputation as researchers, career advancement and job security were influenced by the impact factors of the journals in which they publish. This qualitative evidence supports the results from the phase 1 survey, where respondents indicated that the esteem or reputation of a peer-reviewed journal (albeit more broadly interpreted than journal impact factor alone) was amongst the most important factors in deciding where to publish *(Figure 3.15 above).* There was also a perception amongst phase 2 workshop participants that journal impact factors were becoming increasingly more important, although this seemed to vary by country.

Researchers from the phase 1 focus groups and phase 2 participatory workshop indicated that, on occasion, researchers experience conflicting pressures in terms of where to publish their research. For example, an Engineering participant from the phase 2 participatory workshop expressed there was institutional pressure to consider high impact factor journals (and consequently more potential citations) on the one hand, and his own preference to make the outcomes of his research as widely available as possible on the other. OA was often perceived as an effective means by which to widely disseminate research outcomes, although discussions around this issue invariably highlighted confusion amongst researchers between OA journals, which may have an impact factor, and OAR, which do not.

3.2.1.2 Underlying influences for selecting journals for publication

In terms of dissemination, defined more broadly than publication alone, phase 1 focus group participants and phase 1 survey respondents indicated that the most important influence in their choice of journal was widespread dissemination of research, e.g. reaching their target audience(s) *(Figure 3.15 above).*

Phase 2 workshop participants also emphasised the importance that potential journal readership played in choices over where to publish. On the whole, participants seemed satisfied with their ability to reach targeted audience(s) through their publication choices, albeit more difficult for researchers in some interdisciplinary fields to always achieve this. The Chemistry group indicated that it is well known in their discipline which journals are well subscribed to and hence accessible to many peers, and therefore if one writes an article and wants wide dissemination, one would choose a journal to which everybody subscribes. There was also awareness across the other three discipline groups that each journal has its own readership and that knowledge of this readership influences the choice of journal for publication. There was a sense that, in any given field, readership was sufficiently covered by a set of journals and therefore OAR did not necessarily enhance reaching a specific readership.

Whilst the collective findings of phase 1 and 2 illustrate that authors place a high priority on targeting their publications at an appropriate readership (through careful selection of the most appropriate journal), authors do not seem concerned to the same degree about maximising the potential readership for their journal articles. Participants from the phase 2 workshop reinforced this preference for audience specificity over breadth of reach and were almost unanimous in their selection of a specific journal, or set of journals, on the basis of readership (peers in the same field). Very few of the phase 2 workshop participants expressed concerns about researchers or institutions not being able to afford subscriptions to the journals in which they publish, although a number of participants did voice concerns about library budget cuts in their own institutions. Phase 2 workshop participants generally recognised that researchers in countries with emerging and developing economies may not have the same level of access to scholarly literature and they recognised that OA may help in this regard. Likewise, participants acknowledged that access to journal literature may pose problems for researchers who were not affiliated with Higher Education Institutions and on this basis widening the readership through OA was perceived positively. In general, however, there was a sense that the responsibility of authors to disseminate their research ceased with the publication of their research, and that OA activities beyond the publication stage should be the responsibility of other stakeholders in the scholarly communication system.

3.2.1.3 Speed of dissemination

With regard to rapid dissemination of research, none of the phase 2 workshop participants suggested the use of OAR for this purpose; however the use of conferences for timely dissemination of research findings came up both in the Engineering group and in the Social sciences, humanities & arts group. An Engineering participant with a computer science background indicated that peer-reviewed conference presentations and proceedings were the forum of choice for the purpose of date-stamping ideas (owing to quick turn-around times), but also for rigorous peer review (some conferences can have quite high a rejection rate). Participants in the Social sciences, humanities & arts group shared this view and thought that conferences were appropriate to get important results out quickly. However, another Engineering participant indicated that conferences were not appropriate in his specific intellectual field as they do not have an impact factor, which is important in his particular sub-disciplinary research culture.

In the workshop, participants were presented with a scenario whereby the peer review process would take up to 6 months. They did not consider making a pre-print available as a particularly necessary course of action in this scenario. Participants across all four discipline groups would be inclined to submit their work to another journal, rather than make a pre-print available. A more viable alternative to making a pre-print available would be to submit a paper to a conference, as participants generally agreed that conferences allow rapid dissemination of research and help to raise the visibility of the research. The ability to gain immediate feedback, which in turn may help to further develop the paper, was also seen as a beneficial outcome of presenting a conference paper.

In terms of speeding up the process by which research can be disseminated, some participants noted that on occasion their institutions have discouraged them from making their research papers available in a pre-published version (i.e. on a webpage or in OAR) until they have been formally accepted by a journal.

Placing a copy in an OAR to mitigate the absence of advanced online publication was not at all on researchers' radar either; although a few hinted that they could use OAR if there was no advanced online publication when prompted by members of the research team, there was no strong feeling about it. A Chemistry participant indicated that some institutions require papers to be published before putting them on their institutional repository.

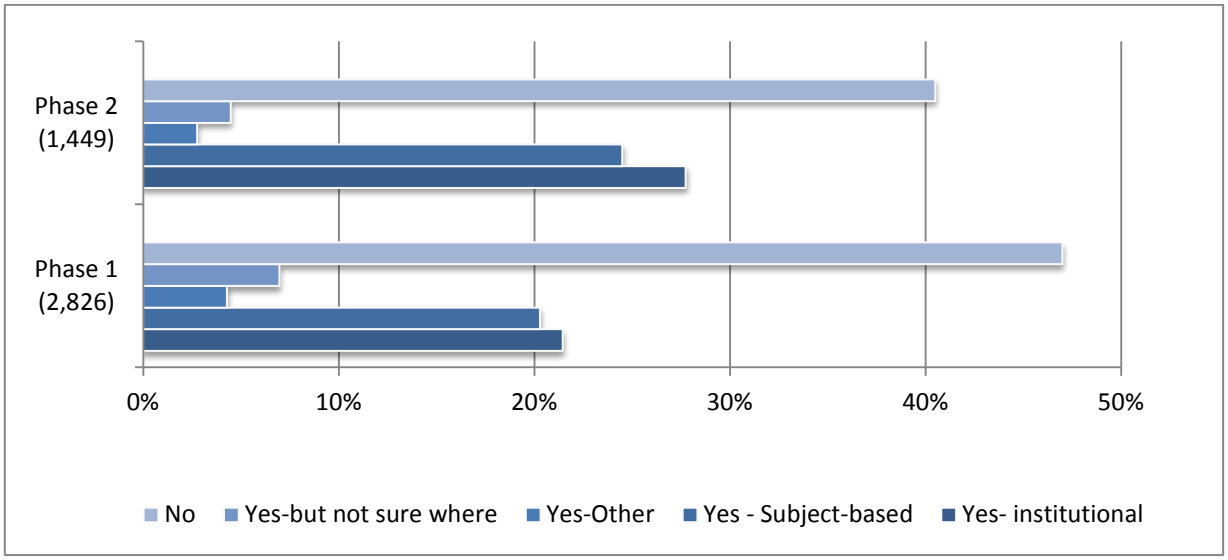
It became apparent in the plenary session of the workshop that high impact factor and speed of dissemination were important factors, alongside wide-spread dissemination and reaching a target readership, in influencing the choice of a journal. Interestingly, it was noted both in the Chemistry group and the Earth, marine & environmental sciences group that speed versus prestige was very much dependent on the researcher's own career stage. Although participants agreed that speed versus prestige would depend on career stage, they arrived at opposite conclusions: for the Chemistry group a young researcher may want to wait and publish in a prestigious journal, whereas a more senior and established researcher may favour broad and quick dissemination over prestige as their reputation is already established. However, participants seemed to indicate that speed was also crucial for young researchers, which makes it all the more difficult for young researchers to obtain a balance between speed and prestige. The Earth, marine & environmental sciences group also thought that there may be a conflict between speed and prestige; however, speed seemed to be more important for early career

researchers and they thought they would think twice about submitting anything in a journal with publication delays beyond six months.

3.2.1.4 OA Self-archiving behaviour

In both surveys, respondents were asked to indicate whether or not they had placed versions of any peer-reviewed articles in any OAR in the last five years. Figure 3.16 shows that the proportion of authors who had placed, or who had had placed on their behalf, one or more articles in an OAR over the previous five years increased from 53% in Phase 1 to 59% in Phase 2.

Figure 3.16 Have your peer-reviewed articles been self-archived?



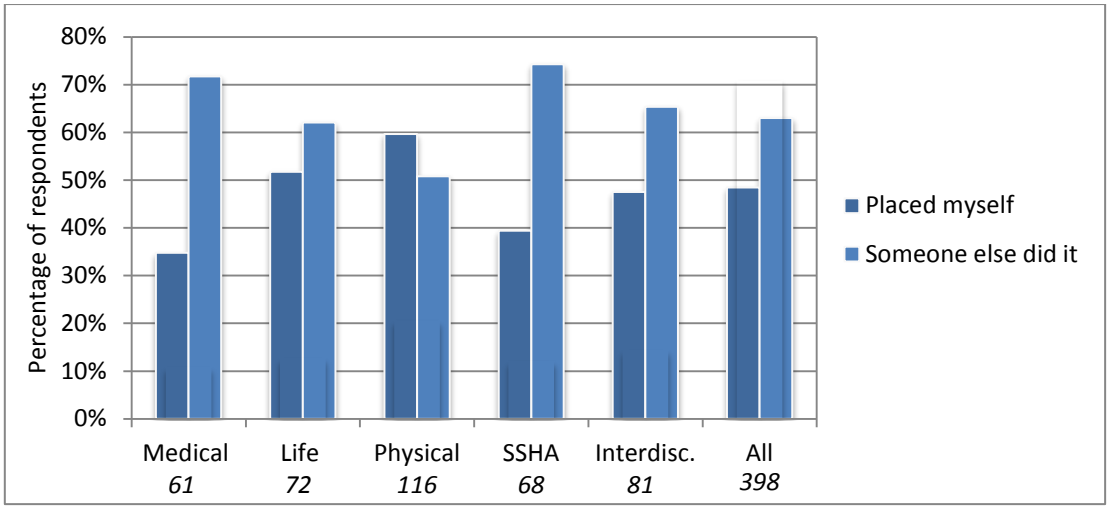
(Weighted total numbers of responses given in brackets)

There were disciplinary differences for this question in both the phase 1 and phase 2 surveys. The types of OAR used were broadly similar in both surveys, with respondents from the Physical sciences & mathematics most likely to use subject-based repositories, whilst those from the other broad disciplinary groups were more likely to use institutional repositories than subject-based repositories. This is particularly interesting in respect of the Medical sciences and the Life sciences, as both the physical and medical sciences communities have a number of esteemed subject repositories that are arguably central to each respective information landscape. In Phase 1, for those Physical sciences & mathematics focus group participants who routinely used repositories, both as authors and readers, such behaviour was expressed as the norm within their research community. A smaller number of Life sciences focus group participants also routinely used repositories and again it was expressed as the norm in their disciplinary community, partly driven by influential OA policies.

In Phase 2, respondents were asked to indicate whether they had deposited an article themselves, or whether someone else had done so on their behalf. Almost one-third of authors who had deposited an article indicated that someone else made their work available in the OAR, with 36% of these stating that this process was performed by library/repository staff, 18% by co-authors, 16% by administrative or secretarial staff, and just over 1% reported that this was undertaken by students or research assistants. All 35 respondents (15%) selecting 'other' provided details, with 29 indicating that the publisher or journal staff had placed the article in a repository for them.

Overall, the phase 2 survey found that 28% of authors had deposited an article in an institutional repository, with 48% of these reporting that they had done so themselves. Just under one quarter of researchers (24%) indicated that they had deposited in a subject-based repository, with 51% of these indicating that they placed an article themselves. Of those respondents indicating that they had deposited in an institutional repository, researchers from the Physical sciences & mathematics and the Life sciences were more likely to have placed the article themselves than respondents from the Medical sciences or the Social sciences, humanities & arts. The reverse also appears to be true, with respondents from the Medical sciences and the Social sciences, humanities & arts more likely to have had someone else place their article in an institutional repository than respondents from the Physical sciences & mathematics *(Figure 3.17).*

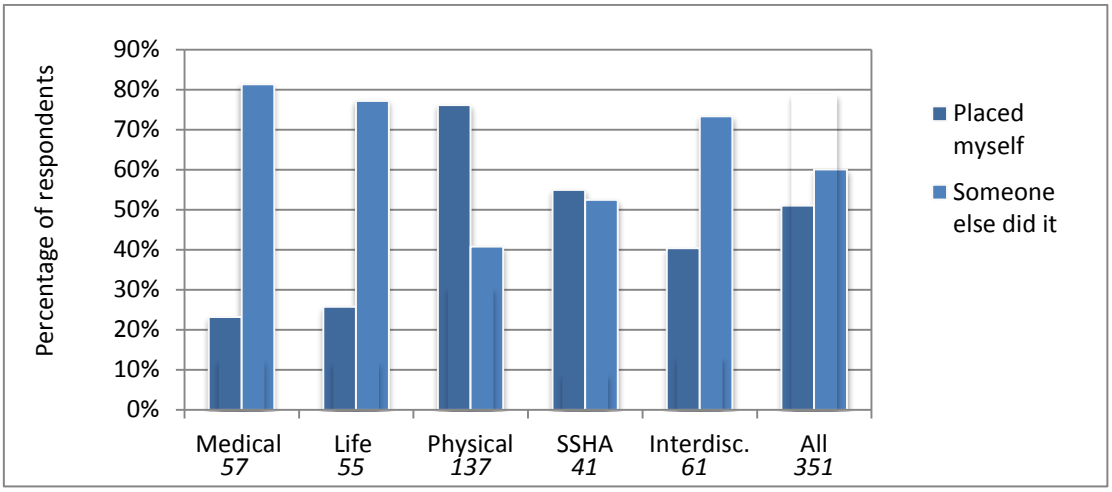
Figure 3.17 Deposit method in institutional repositories, by subject



(Weighted total numbers of responses given in italics.
Note that a small number of respondents in all disciplinary groups ticked both boxes)

There were also disciplinary differences apparent in who made the deposit of articles to subject-based repositories, with those respondents from the Physical sciences & mathematics more likely to have placed an article themselves, whilst those from the Life sciences and Interdisciplinary researchers were more likely to have had an article placed in a subject-based repository by someone else *(Figure 3.18).*

Figure 3.18 Deposit method in subject-based repositories, by subject

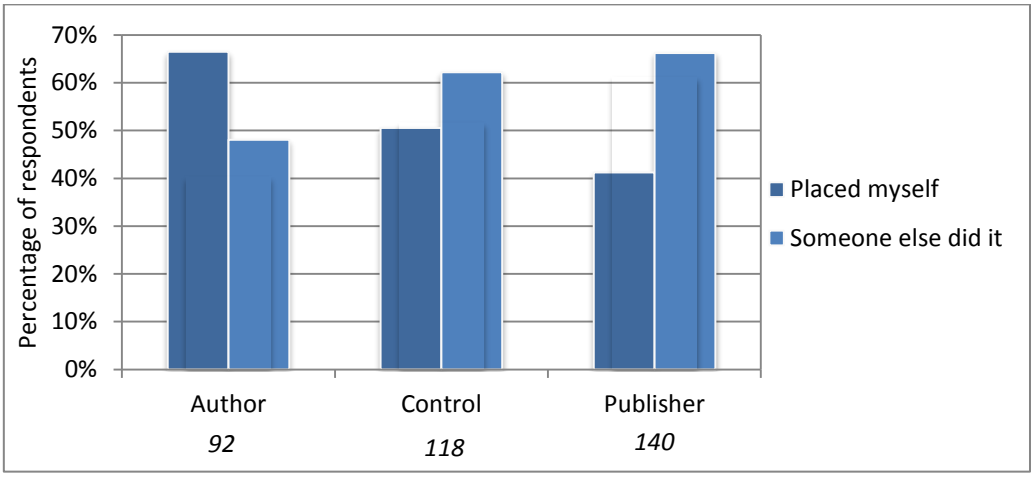


(Weighted total numbers of responses given in italics
Note that a small number of respondents in all disciplinary groups ticked both boxes)

An explanation for differences in the self-reported deposit rates between the Physical sciences & mathematics and the Medical sciences could be the way in which the deposit process is managed by the *de facto* centralised repositories in these disciplinary areas. For example, in the Medical sciences, journal publishers often deposit articles into PubMed Central on behalf of authors, which is in contrast with arXiv where authors often deposit articles themselves. Although many medical scientists tend not to deposit their papers or articles themselves in publicly accessible repositories, many do use PubMed Central for locating sources and the focus group and survey findings indicate that researchers appreciate its role as an important repository of material. Whilst the phase 2 survey findings did not indicate that the process of depositing was a major barrier to deposit, some focus group respondents from phase 1 of the research did outline that deposit involves an additional set of tasks that requires extra time and has to be a managed process (e.g. in terms of who within a group is responsible, which version is uploaded and at which stage in the research).

Differences were noted between the three PEER journal groups, with those respondents in the author group most likely to have placed an article themselves in a subject-based repository, which is perhaps what would be expected *(Figure 3.19).*

Figure 3.19 Deposit method in subject-based repositories, by journal group



(Weighted total numbers of responses given in italics)

In terms of the mechanisms by which authors deposit versions of their articles into either institutional or subject-based OAR, the findings of the phase 2 survey indicated that having someone else place a version of their article into a repository for them was, in general, more common than authors placing a version of their article(s) into an OAR themselves - of those who had reported that any version of their published journal articles had been placed in an institutional or subject-based repository in the last five years, 75% reported it had been placed by someone else, compared to 63% who had reported that they had placed it themselves[8]. There were some disciplinary differences, however, with authors from the Physical sciences & mathematics and the Life sciences more likely to place a version of their articles in an OAR, than authors in the Medical sciences or the Social sciences, humanities & arts.

The physicists who participated in the Physical sciences & mathematics focus group in Phase 1 were familiar with self-archiving versions of their papers in repositories. They expressed particularly strong feelings with regard to deposit workflows. In their experience, the phase of the workflow in which papers were deposited to a repository varied, with participants reporting colleagues who deposit papers at the same time that they submit a manuscript to a journal, compared with their own practice, which tended to be to deposit papers once the manuscript has been peer-reviewed and accepted by a journal. Experimental physics typically involves working in large collaborative teams where a high volume of papers are produced. A senior physicist explained that keeping track of different versions of papers, e.g. pre- and post- peer review, and whether or not they had been deposited to arXiv, for example, was time-consuming and not fool-proof. Workshop participants in Phase 2 of the project also described the process of depositing material in OAR as tedious, time-consuming, and complicated when it comes to sorting out copyright issues relating to self-archiving. This issue is discussed further in Section 4.1.3.

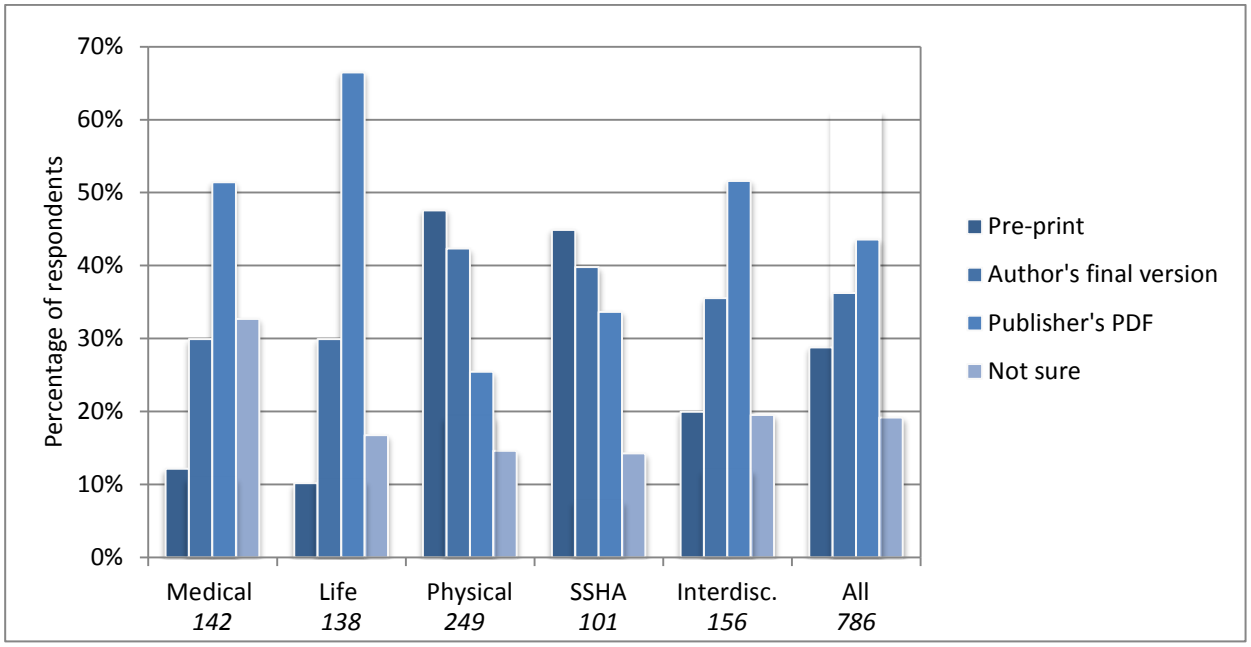
In short, self-deposit is perhaps feasible for individuals or small groups of authors producing a modest number of papers per year, but automatic deposit by publishers becomes more desirable for large-scale distributed groups of authors, such as those exemplified by international particle physics experiments.

3.2.1.5 Article versions self-archived

As for the versions of articles placed into an OAR, phase 2 survey findings indicated that the published final version (publishers' PDF file) was the most popular (44%), with the second most popular being the author's final peer-reviewed accepted version, also called stage-two accepted manuscript (36%). Preferences relating to the version deposited varied according to broad discipline group *(Figure 3.20);* authors from the Life sciences were more likely to have deposited the published final version than those from the Physical sciences & mathematics or the Social sciences, humanities & arts, whilst authors from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to have deposited a pre-print or author's final peer-reviewed accepted version, than authors from the Life sciences. Authors from the Medical sciences were most likely to be uncertain about which version of their article had been placed in an OAR.

[8] Note that respondents could select more than one answer to this question.

Figure 3.20 Version of article deposited, by subject



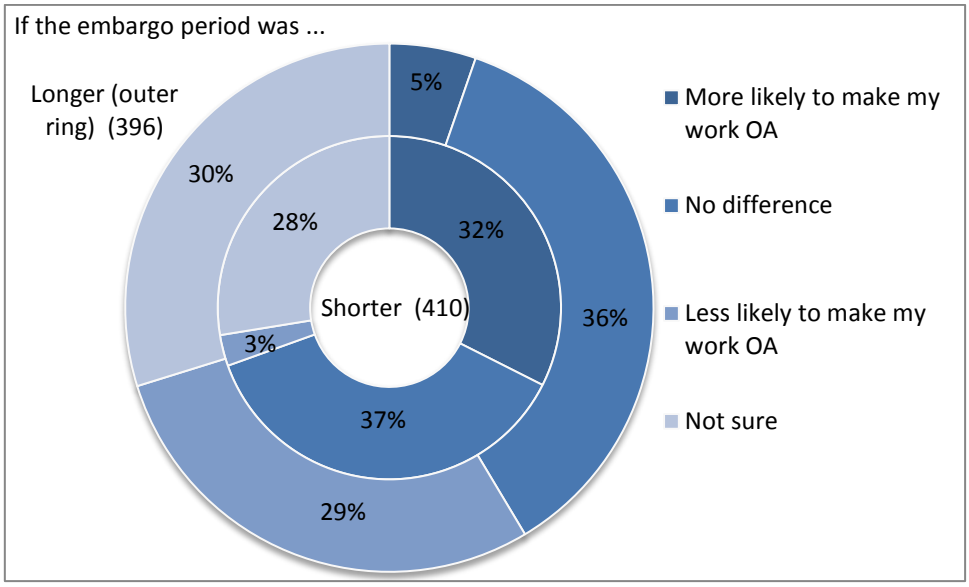
(Weighted total numbers of responses given in italics)

3.2.2 Embargo periods

Surprisingly, 29% of phase 2 respondents indicated that on placing their article in an OAR no embargo period was specified by the publisher, with 55% stating that they did not know or could not remember what, if any, embargo period there was. Six percent of respondents indicated that the embargo period was 6 months or less, with a further 6% indicating that the embargo period was 7 to 12 months. Respondents were asked to indicate how their self-archiving behaviour might have been influenced if the embargo period had been different – either shorter or longer – to that actually imposed, in relation to their inclination to self-archive, and their choice of publisher or journal for the publication.

Almost one-third of authors (32%) indicated that, if the specified embargo period was shorter than that imposed when they placed their article in an OAR, they would be more likely to make their work OA, while 37% stated that it would make no difference and just 3% thought they would be less likely to make their work OA in this scenario. These figures are illustrated in the inner ring of Figure 3.21. The outer ring of Figure 3.21 illustrates the effect of the opposite scenario, i.e. if the publisher had specified a longer embargo period than that had actually been imposed when they deposited in an OAR. In this case, just 5% of authors thought that they would be more likely to self-archive their article, while 36% claim that it would make no difference, and 29% felt that this would make them less likely to make their work OA.

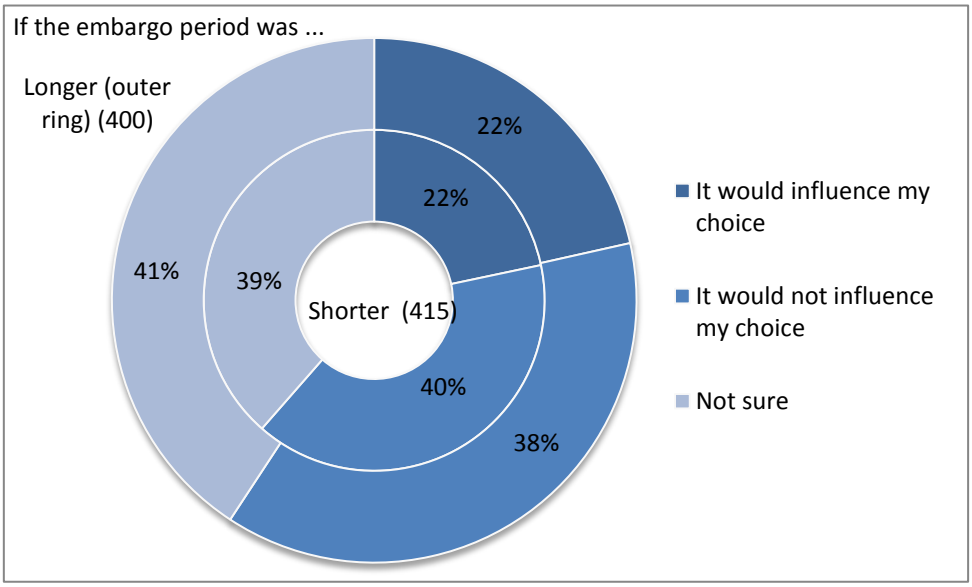
Figure 3.21 Impact of embargo periods on making work Open Access



(Weighted total numbers of responses given in brackets)

With regard to choice of journal, there was little difference in the patterns of responses whether the hypothetical embargo period was longer or shorter than that actually imposed. In each case 22% of researchers suggested that changing the embargo period – either longer or shorter - would affect their publication choices. Around two in five were unsure, while 40% claimed that a shorter embargo period would not affect their choice of publisher/journal, with 38% of researchers stating that a longer embargo period would have no effect. These percentages are illustrated in Figure 3.22, where the inner ring shows the percentages under the scenario of shorter embargo periods, with the outer ring illustrating the percentages under longer embargo periods.

Figure 3.22 Would changing the embargo period affect author choice of publisher/journal



(Weighted total numbers of responses given in brackets)

3.2.3 Disciplinary differences and similarities

Social science, humanities & arts researchers published the fewest articles on average, whereas researchers from the Medical sciences published the highest average number of articles *(Figure 3.13, page 28).* In terms of placing their article(s) into an OAR authors from the Physical sciences & mathematics were most likely to use subject-based repositories, whilst authors from the other broad disciplinary groups were more likely to use institutional repositories, than subject based repositories.

Authors from the Physical sciences & mathematics and the Life sciences who indicated that they had used an institutional repository were more likely to have placed the article themselves than authors from the Medical sciences or the Social sciences, humanities & arts. The reverse also appears to be true, with authors from the Medical sciences and the Social sciences, humanities & arts more likely to have had someone else place their article in an institutional repository than authors from the Physical sciences & mathematics *(Figure 3.17, page 35).*

There were also disciplinary differences apparent in who made the deposit of articles to subject-based repositories, with those authors from the Physical sciences & mathematics more likely to have placed an article themselves, whilst those from the Life sciences and Interdisciplinary researchers were more likely to have had an article placed in a subject-based repository by someone else *(Figure 3.18, page 36).*

In terms of the version of the article deposited, authors from the Life sciences were more likely to deposit the published final version than authors from the Physical sciences & mathematics or the Social sciences, humanities & arts. Conversely, authors from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to have deposited a pre-print or author's final peer-reviewed accepted version, than authors from the Life sciences *(Figure 3.20, page 38).*

4 Researchers' perceptions of the best way(s) of achieving Open Access

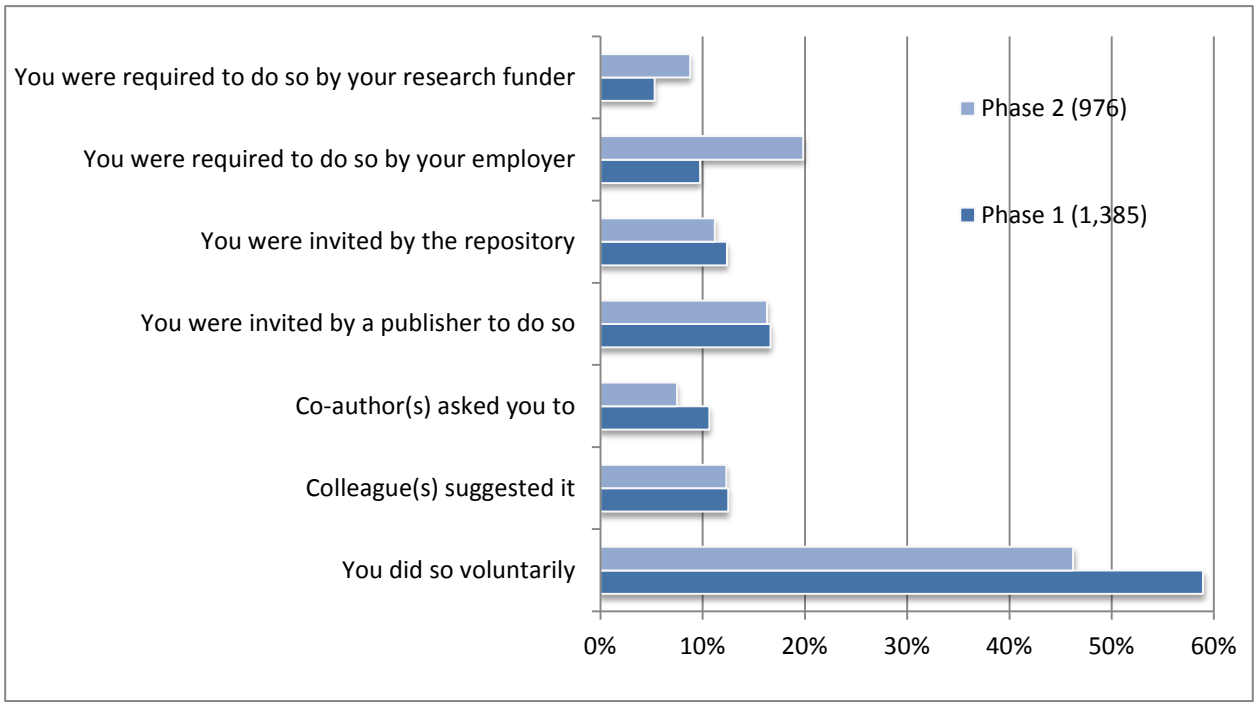
4.1 Motivations to self-archive and perceived benefits of OA

The two surveys, focus groups and workshop brought a wealth of information in regard to authors' perceptions of the benefits of OA and their perceptions of OA mandates. This is detailed in the following sections.

4.1.1 Reported motivations for repository deposit

Figure 4.1 shows the change in the proportions of respondents to the two surveys reporting various motivations for repository deposit. In phase 2, of those researchers who had deposited (or had had placed on their behalf) a copy of an article in an OAR, 46% reported that they did so voluntarily. This compares to 59% of researchers in phase 1 who had deposited voluntarily. This decrease might be explained by the notable increase in the percentage of authors depositing in response to mandates, from 10% of researchers in phase 1 to 20% of researchers in phase 2 indicating that they were mandated by their employer. Similarly, 9% of researchers in phase 2 stated that they were mandated by their funder (compared to 5% of researchers in phase 1).

Figure 4.1 Motivations for repository deposit

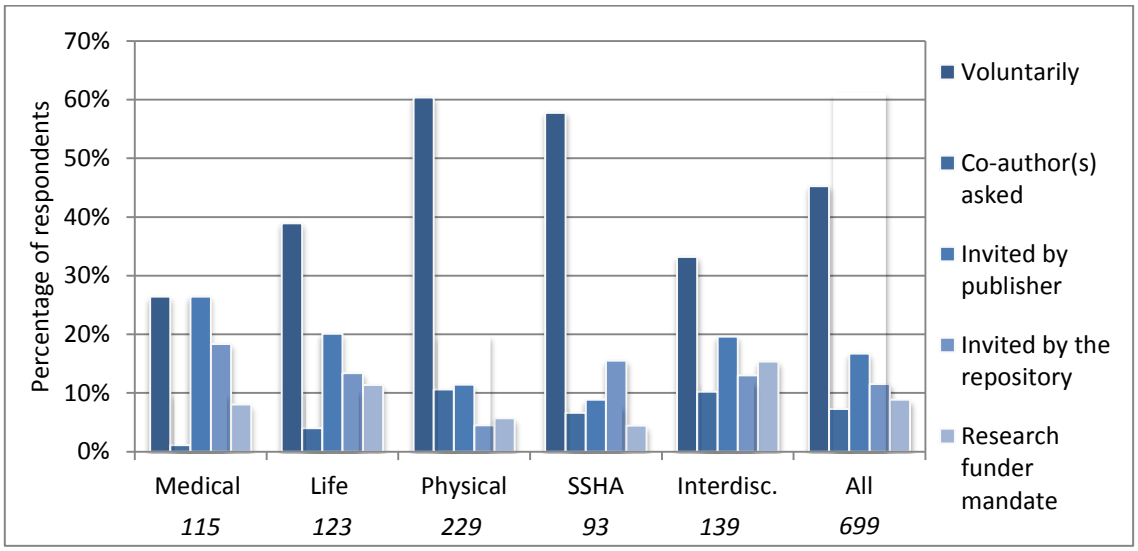


(Weighted total numbers of responses given in brackets)

There were some differences between the broad disciplinary groups in phase 2 *(Figure 4.2).* Researchers from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to have deposited their article in an OAR voluntarily, than those from the Medical sciences, and were less likely to be mandated by a research funder to deposit than researchers from the Life sciences or Interdisciplinary researchers. A request from co-authors was less likely to prompt researchers from the Medical sciences to deposit in an OAR than those from the Physical sciences & mathematics or Interdisciplinary researchers. Researchers

from the Medical sciences and the Life sciences were more likely to have been invited by a publisher to deposit than those from the Physical sciences & mathematics or Social sciences, humanities & arts. Researchers from the Physical sciences & mathematics were least likely to have been invited by the repository to deposit.

Figure 4.2 Motivations for repository deposit, by subject



(Weighted total numbers of responses given in italics)

There was a change (from the phase 1 survey to the phase 2 survey) in authors' behaviour with regard to repository deposit in response to OA mandates *(Figure 4.1)*. The reason may be that in recent years funding agencies in most disciplines have developed OA policies, or at least position statements on OA. Furthermore, where funding agencies do stipulate mandatory deposit or have position statements, they also encourage grant applicants to apply for the necessary funds to make articles publicly available via OA (e.g. financial support for Gold OA). However, it is of interest that funder, or indeed, institutional, mandates were still considered relatively unimportant as motivators for repository deposit by survey respondents. Nonetheless, it should be noted that, despite the limited importance of mandates, a difference was noted between respondents from the different PEER experimental journal groups, with those respondents in the publisher group more likely to have deposited as a result of a mandate by a research funder, than respondents in the author and control groups. It is of particular interest that respondents in the author group (i.e. those that had been invited to deposit in an OAR by a publisher) were not more likely than the other groups to have deposited for this reason.

Reflecting the views of scholars expressed by Thorn *et al.* (2009), workshop participants generally seemed to embrace the ethos of OA, and generally thought it a good idea to make work available to other researchers who may have problems accessing scholarly literature. However, there was a general feeling in the discussion that this is not the top priority for researchers. They wish first to satisfy the requirements they are subject to as authors and, despite the growing number of mandates (from none in 2003, to 213 worldwide in July 2011, with a further 20 proposed[9]), OA does not seem to be embedded in those requirements.

[9] http://roarmap.eprints.org [accessed 18.07.11]

It is clear that discipline norms influence the behaviours of authors in particular, not least as some disciplines have a long history of using subject-based repositories to disseminate their research, both as pre-prints and as authors' final manuscripts. The use of arXiv in physics and RePEc in economics, or more recently SSRN for the social sciences in general, are perhaps the best known of these. The focus groups found that only physicists were in the habit of depositing articles in OAR and both the surveys confirmed this, with more than half of respondents in Physical sciences & mathematics having deposited material in a repository in the last five years. Medical scientists were least likely to have deposited material.

4.1.2 Repository type preferences

As part of the phase 1 survey, respondents were asked to indicate whether they had a preference for the type of repository in which they chose to deposit their material, with 44% preferring a subject-based repository, and 23% an institutional repository, whilst 21% had no preference. A similar question was included in the phase 2 survey, asking which type(s) of OAR had been used by those depositing articles. Of those that had deposited one or more articles in the previous five years, 32% had used a subject-based repository and 36% an institutional repository, while 12% had used both. *(Table 4.1).*

Table 4.1 Authors' preference for type of OAR

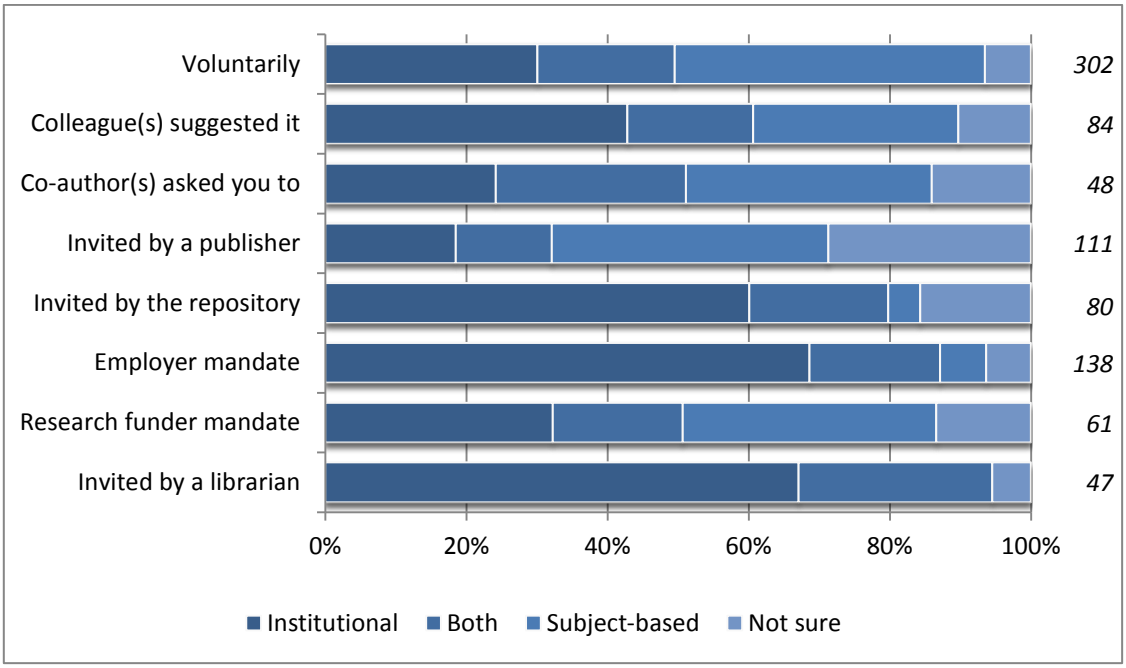
Type of OAR	Phase 1 - preferred	Phase 2 - used
Institutional	23%	36%
Subject-based	44%	32%
Both		12%
No preference	21%	

The apparent shift in repository type preferences between Phase 1 and Phase 2 of the research warrants further investigation, although it may be explained partly by the increasing importance of institutional mandates noted above. The findings from the phase 2 workshop provide anecdotal evidence that some researchers would prioritise institutional mandates over funder mandates.

There were some notable differences when comparing the reasons why phase 2 survey respondents had placed a copy of an article in an OAR and the type of repository they chose *(Figure 4.3):*

- Respondents that voluntarily deposited their article in an OAR, were asked by co-authors to do so, or were invited by a publisher to deposit were more likely to have chosen a subject-based repository
- Respondents that were invited by a repository or librarian to deposit, or were mandated by their employer were more likely to have chosen an institutional repository

Figure 4.3 Motivations affecting choice of repository type



(Weighted total numbers of responses given in italics)

A total of 22 respondents selected 'other' to the question regarding the type of OAR they had chosen to receive deposit of their articles, with all providing further details. Overall, ten respondents indicated that they had not made an active choice regarding the type of OAR they had deposited in - it was something that occurred automatically as a result of the journal they had published in, or had been decided by someone else. In addition, two respondents indicated that they had made their articles freely available on their own website; with a further two commenting that their articles had been placed on a publisher's website.

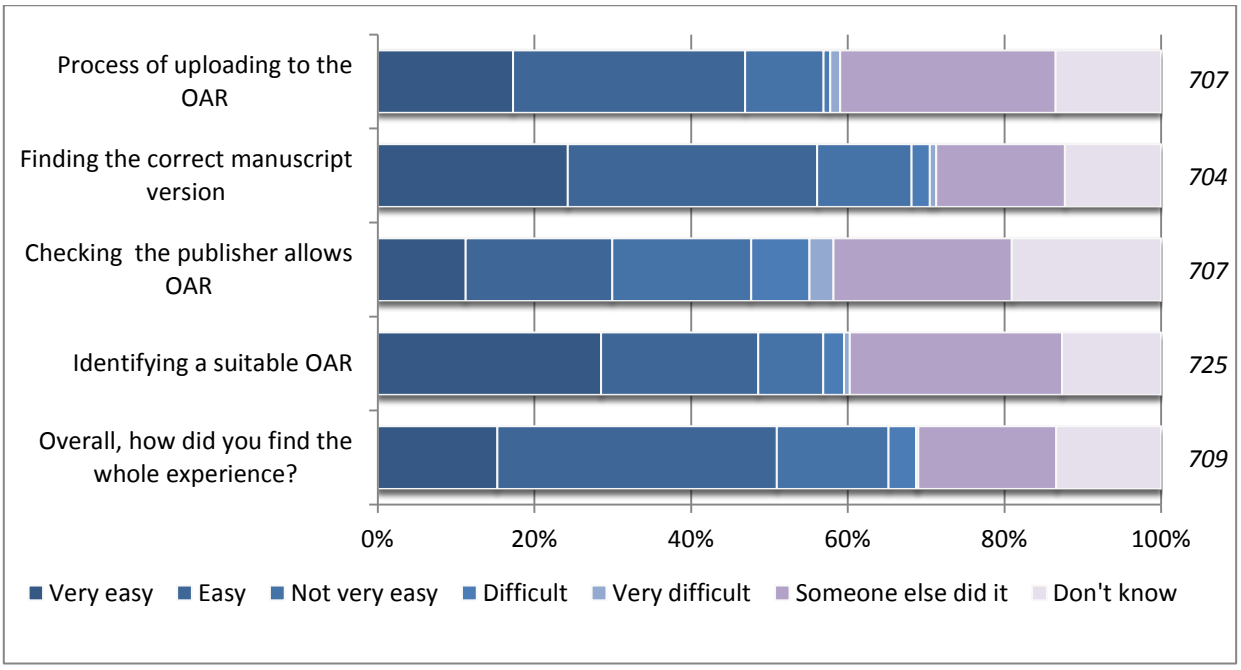
Focus group participants also expressed mixed preferences for repository type, with some strongly in favour of institutional repositories, as they were thought to support the institutional research strategy in the same way as staff profiles listing research interests and publications. The concept of service providers being able to harvest metadata from institutional repositories and provide subject-based services was not directly addressed in the surveys, and general levels of awareness implied from the focus groups and workshop participants indicated that this was not an area that had been considered by many participants. Focus group and workshop participants did, however, feel strongly that repositories should provide added-value in some way, and welcomed the introduction/availability of additional repository services such as download statistics and email alerts – though this issue was not elaborated on in the different groups owing to time constraints.

A few workshop participants across disciplines acknowledged the use of institutional repositories as a research management system, notably for periodic research quality assessment, though this was not perceived as a priority by most participants. However, it was noted that ultimately the distinction between subject-based repositories and institutional repositories may be irrelevant, as one participant commented, in the sense that it does not matter where material is deposited as long as it is well indexed, and hence picked up by the discovery tools that researchers use.

4.1.3 Authors' perception of the cost/benefit balance of repository deposit

In terms of authors' experience of depositing a copy of their article(s) into an OAR the overall results suggest that authors do not generally experience much difficulty. Figure 4.4 illustrates the reported ease of the various steps involved in uploading a copy of their article to an OAR. The overall experience of self-deposit reported in the phase 2 survey was rated as 'very easy' or 'easy' by 51% of authors. Finding the correct manuscript version proved to be the easiest task, with 56% of respondents indicating that this process was 'very easy' or 'easy'. Just 30% of researchers found the process of checking that the publisher allows the article to be placed in an OAR as 'very easy' or 'easy', with almost 11% indicating that this process was 'difficult' or 'very difficult'.

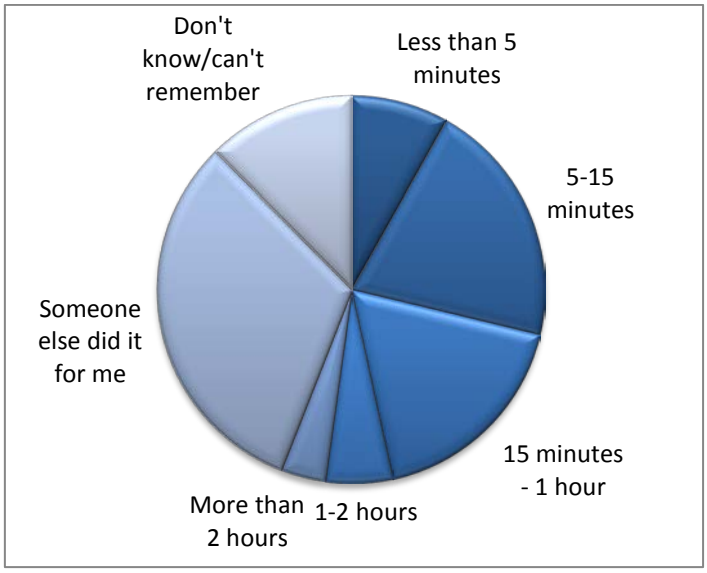
Figure 4.4 Ease of repository deposit procedures



(Weighted total numbers of responses given in italics)

Of course, individuals' perception regarding the ease of conducting these steps may depend on whether they deposited the article themselves or had someone else deposit the work for them. Overall, 21% of respondents indicated that it took just 5-15 minutes to make their work available in an OAR; however, almost 4% of researchers claim that this process took more than two hours *(Figure 4.9).*

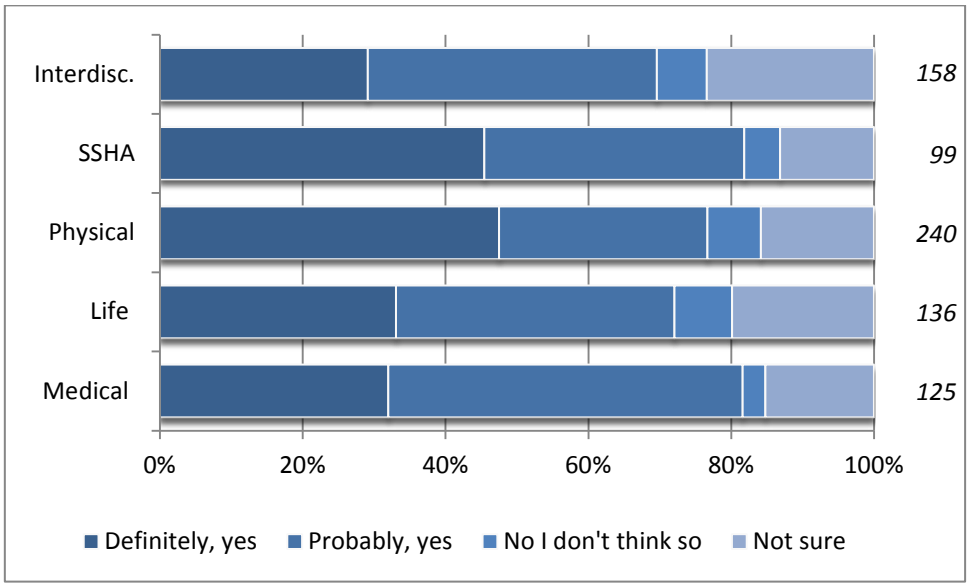
Figure 4.5 Time taken to deposit



(n=732)

In terms of weighing the perceived benefits of OA against the effort of depositing a copy of their article(s) in an OAR, 39% of phase 2 authors felt that the extra effort was definitely worth it, with around three-quarters feeling it was 'definitely' or 'probably' worth doing. Only 7% of authors indicated that they felt that the benefits were not worth the extra effort. Authors within the Physical sciences & mathematics, and the Social sciences, humanities & arts were most likely to feel deposit was 'definitely worthwhile', compared to respondents from the Life sciences and the Medical sciences *(Figure 4.6).* As an illustration, one Social sciences, humanities & arts workshop participant, familiar with RePEc and SSRN (both subject-based repositories), explained that by using OAR individually or in combination one could gain recognition for a body of work rather than for individual publications.

Figure 4.6 Is deposit in OAR worthwhile, by subject?

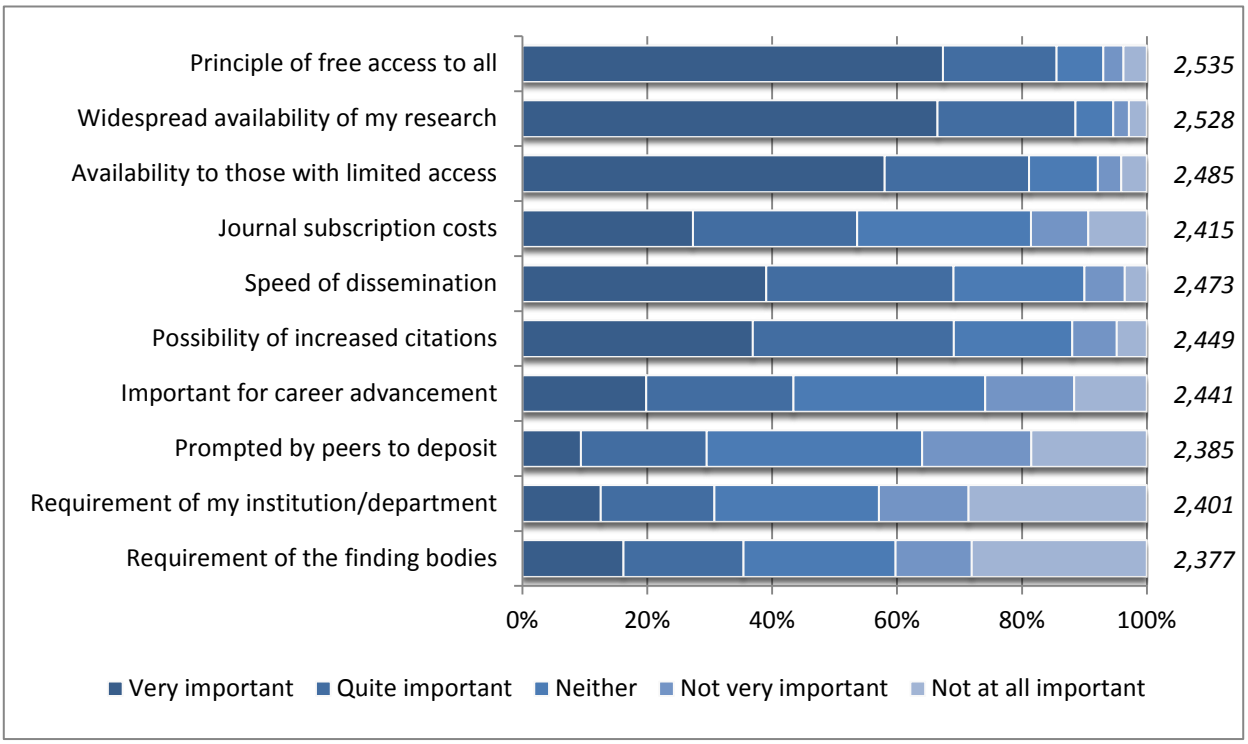


(Weighted total numbers of responses given in italics)

4.1.4 Enhanced scholarly communication

By accepting an author's final version (or a preprint) of an article, OAR exist alongside subscription journals and can provide authors with the ability to make their work accessible to everyone whilst still maintaining the authors' desire for quality control through peer review. Figure 4.7 illustrates that, in the opinion of the phase 1 survey respondents, the three factors that were most likely to encourage respondents to adopt OAR as a means of achieving Open Access were 'The principle of free access to all', 'Widespread availability of research' and 'Availability to those with limited access'. All of these factors are related to making work as widely accessible as possible. It is interesting that similar factors were also deemed important to the authors' choice of journal *(Figure 3.15, page 31)* where 'Widespread availability of my research' was the most important factor in journal choice. It would therefore seem that journal publication augmented by OAR deposit *should* be attractive to authors as a way of increasing availability while retaining the benefits of the existing scholarly publishing processes. However, the current inability of repositories to attract deposit from authors may be partially explained by other factors that influence an author's choice of journal *(Figure 3.15, page 31);* factors such as 'Esteem of publication' and 'Dissemination to target audience' are also deemed important (or very important) to a large percentage of authors. Comments about OAR in the phase 2 workshop imply that repository deposit does not add kudos or esteem to a publication. In addition, it seems likely that authors feel that dissemination to a *specific target* audience is what is really important. The indications are that, although authors state that they desire widespread dissemination of their work, what they really desire is widespread dissemination to an audience of peers within their own and related disciplines. Across the workshop groups, there was a feeling that widespread dissemination to everyone would be nice – but that it is not as important as widespread dissemination within their discipline. Choice of journal and acceptance into that journal would therefore appear to be of primary importance to many authors in comparison to repository deposit.

Figure 4.7 Factors encouraging OAR deposit



(Weighted total numbers of responses given in italics)

As well as improving availability, OAR can enhance scholarly communication in other ways. In some disciplines the publication lags associated with the most highly regarded subscription-based journals are so lengthy that the research community has developed complementary modes of dissemination that enable researchers to date-stamp important results in a timely manner, whilst they wait for manuscripts to go through the peer review and publication process. For example, in physics the arXiv pre-print OAR plays an important role in date-stamping ideas, and in economics working papers can be submitted to RePEc for the same purpose. In both disciplines workshop participants explained that they place their papers into these OAR at the same time, or almost the same time, as submitting them to a journal. Whilst participants from Chemistry-related sub-disciplines valued the advantages of rapid dissemination and date-stamping via pre-prints afforded by OAR, this same sentiment was not extended to authors' final drafts. It would seem that in these sub-disciplines of Chemistry, with which there was some overlap with biology in the nature of the research, OA journals play a role similar to that of arXiv in physics and RePEc in economics, because they offer rapid publication (i.e. these participants reported 3 months). This echoes anecdotal evidence that outlines the success of peer-reviewed OAJ, such as PLoS[10], and the availability of funding for open access publishing in life sciences (Björk, 2010). Figure 3.20, page 38, illustrates the importance of pre-prints as a mode of dissemination within the Physical sciences & mathematics and the Social sciences, humanities & arts, and the preference for published final versions of journal articles amongst Medical sciences and Life sciences researchers.

Supporting the fact that publication delays may be problematic to some researchers, the initial phase 1 survey showed that speed of dissemination was also an important factor in authors' choice of journal *(see Figure 3.15, page 31)* although other factors such as esteem of publication, peer review and career advancement were deemed important (or very important) by more respondents. In term of being a motivation for repository deposit, speed of dissemination was also considered important or very important by 70% of the respondents *(Figure 4.7).*

The phase 1 survey showed that the possibility of increased citations was nearly as important as improved speed of dissemination as a motivation for repository deposit, with 69% of respondents indicating that this was important (or very important) to them *(Figure 4.7)*. There were differences between disciplines - speed of dissemination was less important for respondents in the Medical sciences than those in the other broad disciplinary groups and respondents from the Social sciences, humanities & arts were most likely to rate the possibility of increased citations as 'very important'.

Other motivations for repository deposit were mentioned in the workshop in addition to those outlined in Figure 4.7. One workshop participant from the Social sciences, humanities & arts indicated that some Spanish institutions had set up financial rewards for deposits in institutional repositories. This seemed to be perceived positively, although the participant did think that financial incentives could affect researchers' priorities in choosing where to disseminate. Whether financial incentives could become a major factor (i.e. make it less likely for authors to consider submitting to prestigious journals if they prevent deposit in OAR) remains unclear.

Participants from the Social sciences, humanities & arts group also indicated that a motivation for self-archiving is that their articles are often extracts or adverts for forthcoming books, and in

[10] http://www.plos.org/ [accessed 29.09.11]

this regard making articles available on OA was seen positively as a way to encourage readers to get the related book.

Access to research information for readers in countries with emerging and developing economies is often cited as one advantage of OA, whether through OAJ or self-archiving in OAR or elsewhere (Swan, 2007). Participants in the Budapest focus group in Phase 1 commented that a lack of funding for research in their institutions also restricted journal subscriptions, so that OA sources were particularly helpful, although other methods of sourcing material were also used. Focus group participants also commented that more repository content in non-English languages would address a gap in dissemination technologies. Another aspect of material held in OAR was that it could be used by poorer institutions as a way to find an entry point to relevant articles/materials. Once articles had been identified it would then be possible to contact the author to get the final published version. More information on the use of formal and informal networks to access research articles is provided in Section 3.1.3.

4.2 Perceived barriers of Open Access Repositories

4.2.1 The self-archiving process

Although phase 2 survey respondents generally thought that deposit procedures were fairly easy *(see Figure 4.4, page 45)* and that deposit in OAR was worthwhile *(see Figure 4.6, page 46),* most workshop participants in the Chemistry group thought the whole deposit process was tedious, extremely time-consuming and somehow discouraging, especially when it came to sorting out copyright restrictions. It appeared that most of these authors did not know what rights they have once they have signed off Copyright Transfer Agreements (CTA) but they did not seem to care especially about this issue, as long as their moral right to be cited as the original creator was respected. As outlined in Figure 4.4, determining the copyright status of articles to be deposited does not come easily to authors. Besides copyright, the input of metadata was perceived as a tedious exercise; for instance it was indicated that in some disciplines, such as Chemistry, a paper can be co-authored by a great number of researchers, each with different affiliations, which can make the data input process particularly lengthy. This experience did not seem to be an isolated case and was generally shared by other groups, notably in the Engineering group, where a participant explained that their submissions to his institutional repository stopped in 2005 because, despite the fact that he was required to update his research profile, the "*messy confusing interface*" made the process difficult and tedious; consequently this participant only kept his own research webpage up-to-date. Although a minority viewpoint, this should be taken seriously, as it presents a significant barrier to depositing work in OAR. It can be mitigated by others (whether administrative staff, librarians, research assistants, etc.) taking responsibility for the deposit process, if not the initial decision.

Another barrier to greater adoption of OAR, and which links back to researchers' low levels of awareness of repositories and how they operate, is that researchers are not always fully involved in the self-archiving process. Researchers participating in the workshop were not always clear how material ended up in OAR. A few participants indicated they were surprised to find some of their articles in repositories, claiming they did not know how the material got there. Such experience may be particularly true for subject-based repositories such as PMC where publisher-mediated deposit is common, but may also be relevant to some institutional repositories where most of the deposit is mediated by library staff with minimal involvement from the researchers themselves - an example is an institution deciding to entrust library staff to deposit in the institution's repository on behalf of its authors in order to 'make it easy' for its

researchers to make their work available via OA, as well as boosting the repository's growth. Figures 3.17 and 3.18 *(pages 35 and 36)* show that in many disciplines articles are not directly placed into OAR by authors themselves but by other agencies. As a consequence researchers may feel 'alienated' with the deposit process itself, in that it is not a conscious step by the researcher to make papers OA.

4.2.2 The 'stigma' of open access

A key feature of focus group participants' uncertainty was what to expect in terms of quality, be it the reputation of a repository itself in the case of authors, or the quality of the articles deposited in the case of readers. This was in contrast to the phase 1 survey where respondents indicated that open access does not equate to diminished quality. There was a feeling amongst some focus group participants that, from an author perspective, non-refereed preprints of variable quality may not be appropriate alongside peer-reviewed authors' final manuscripts. The phase 1 survey identified that a reluctance to deposit articles where other material had not been peer-reviewed was a significant concern for survey respondents in all four of the broad disciplinary groups. Although this was not a concern raised within the Physical sciences & mathematics focus group, some of the Life sciences participants expressed concern over non-refereed papers appearing in the same repository as peer-reviewed articles.

OA started out with the sharing of pre-prints and it seemed to many workshop participants that OA often equates to pre-prints – though they did show some awareness that OA now includes author's final versions of published articles. Although the phase 1 survey findings indicated that only a small percentage of respondents equated OA with 'not peer-reviewed', the perception captured at the workshop in this respect did not equate to this. The idea that workshop participants tended to conflate OA with pre-prints was reinforced by the fact that participants (especially the Chemistry group) always seemed to refer to the pre-print version when discussing the different points at the workshop. Such an enduring conflation of OA and pre-prints seems to undermine trust in and adoption of OAR. Indeed, despite a growing awareness of OA and OAR, repositories seem to suffer from the stigma of including non-peer-reviewed material, and this may be also reinforced by the now wide availability of resources on the Internet, as it makes it even more difficult for readers to identify what sort of material they are accessing on the Worldwide Web.

The main concern regarding availability of work and OAR seemed to be that if repository content was fragmented in terms of quality, this will have a negative influence on the overall prestige of the repository and thus reduce the visibility of authors contributing to that content. This resonates with Borgman's (2007) argument that the 'legitimatization' of sources and resources becomes increasingly important in the context of emergent publishing paradigms, such as self-archiving, and the fuzzy boundary between traditional and emergent publishing paradigms. Furthermore, the negative correlation between a perceived low prestige of a particular OAR and willingness to deposit to it, suggests that studies (Swan and Brown, 2004; Watson, 2007) linking research evaluation mechanisms with disincentives for academics to disseminate in OA journals with a low impact factor are also relevant for OAR.

4.2.3 Ease of identifying the version accessed

Another area of confusion regarding the way that OAR present content is the clarity as to whether material accessed is in fact in an OAR. In part, this may be due to the success of providing transparent and easy access to subscription-based information resources. The

phase 1 survey found that only one-third of respondents reported that it was 'usually' clear that articles were held in a repository, and this was most likely to be so in the Physical sciences & mathematics. Seventy-four percent of phase 1 respondents overall felt that it was 'very important' or 'quite important' to know which version of the article they had found (i.e. a pre-print, author's final manuscript or publisher's PDF), but only 39% reported that they could 'usually' identify the version from the information provided by the OAR, whilst 37% reported that they 'always' trust the integrity of documents they find in repositories.

There are some interesting patterns in the phase 1 responses by subject area concerning the use of articles in repositories. Respondents from the Medical sciences and the Social sciences, humanities & arts are most likely to consider the version of an article important, and least likely to feel that they can identify this from the information provided. Consequently, they are least likely to trust the integrity of documents in repositories. A Social sciences, humanities & arts participant commented on points such as these by saying that "*serious scientists use serious sources. OA is thought to be of poor quality and not peer-reviewed*".

4.2.4 Visibility and usability

Visibility of an author's work was raised in the focus group discussions, and there was a perception that, unless a repository itself had visibility and kudos, then the motivation to deposit papers or articles in it would be low. This view is reinforced by the fact that availability and accessibility of articles to the authors' intended audience was one of the key factors driving their choice of journal and repository *(see Figures 3.15, page 31 and 4.7, page 47).*

Supporting the issue of visibility as a drawback to OAR, workshop participants thought that repositories were generally not easy to search and navigate when one goes directly to the repository to search for specific items, as one would do to search arXiv or PubMed Central. Indeed some Chemistry participants commented on the reader interface of arXiv, and to some extent PubMed Central, as being basic and devoid of the kind of sophisticated search interface found in commercial aggregators' or publishers' platforms. For those researchers, there was a feeling that one needs to use repositories on a very regular basis to understand how to navigate them properly. Furthermore, it is interesting to note that, to workshop participants, repositories were often seen as a resource where one goes directly rather than as suppliers of OA material to search engines for broader retrieval. Despite this, the general discussion about the importance and centrality of the indexation of OA material held in repositories showed that researchers are aware that content in OAR can be retrievable via search engines.

The Chemistry group at the workshop elaborated more than the other groups on the fact that it was very difficult to find material held in institutional repositories because they are not properly indexed and thus not easily retrievable by the discovery tools researchers use, notably search engines such as Google and Google Scholar. The Chemistry group thought that repositories did not really compete with publishers' commercial platforms for the reason that papers in OAR were difficult to find. Although some workshop participants felt that repositories were not indexed as well as they might be on Google and Google Scholar, the phase 2 survey found that searches via Google and Google Scholar were the most popular way of arriving at OAR *(Figure 3.2, page 14).*

Referring to the issue of unreliability, one workshop participant, who is mandated by his institution to upload material on the institution's OAR and, as such, regularly self-archives, indicated that repositories can sometimes be unreliable in terms of finding material. He

explained that he had been unable to retrieve some of the papers he had uploaded himself on his institution's OAR.

As well as a lack of visibility, other reasons suggested by workshop participants regarding the non-use of OAR by readers included the fact that pre-prints or authors' final draft manuscripts are not considered citable in a piece of scholarly work *(Section 3.1.2.4),* and the perceived unreliability of OAR. Researchers preferences regarding the suitability of citing OA material was supported by the phase 1 survey *(see Figure 3.7, page 21)*.

4.3 Open Access Publishing

OA publishing was not directly addressed in the surveys; however it was considered in the workshop as participants were asked to consider whether OA publishing was a better way to achieve OA than self-archiving in OAR. Workshop participants recognised that the reduction in publication delays enabled by Gold OA was a great advantage. It was felt that articles available via a Gold Open Access route are easily retrievable and discoverable as they sit on publishers' commercial platforms and this was perceived as a significant advantage. By contrast, participants thought material held in OAR were not easy to find. In this respect, subject-based repositories were seen as more attractive (and visible) than institutional repositories but needed a critical mass of content and use, i.e. they needed to be adopted and used by the whole community.

Despite the free and easy access to articles published in OAJ, researchers overall were not particularly supportive of OA publishing for two main reasons: the perceived lack of impact factor; and the high fees required to publish in OAJ. In terms of impact factor, participants acknowledged that these journals tend to be new journals and it was widely acknowledged that it takes time for a journal to build up a reputation and gain a good impact factor. In terms of cost, despite the fact that participants showed a good understanding of the business model for OA publishing, it was generally felt that OA fees were too expensive. Some clearly indicated they did not have the money for OA publishing, especially in the Social sciences, humanities & arts discipline. The majority thought paying to publish in an OAJ was not acceptable in principle, though a few disagreed, as there is a cost to publishing – appreciating that, although editors and reviewers often work for free, there is an administrative and publishing cost; however they could not understand why it was so high. Some added that the model was unfair – good research was published in high quality journals, which were able to charge the highest OA fees, with the implication that researchers only did good research if they could afford to pay high fees.

The general perception was that the current scholarly communication system seems to work reasonably well and workshop participants generally did not feel that making their work available through OAJ was worth the price asked by publishers. When confronted about building publishing costs into grant applications, Chemistry participants unanimously felt that there would be better use of the money somewhere else in the funded project. Some, notably in the humanities, indicated that the cost of OA publishing was sometimes equivalent to the grant money received for an entire research project. There were also concerns over the fact that not all research is funded by a grant. Only two out of the 22 workshop participants reported having already paid for publishing in OAJ.

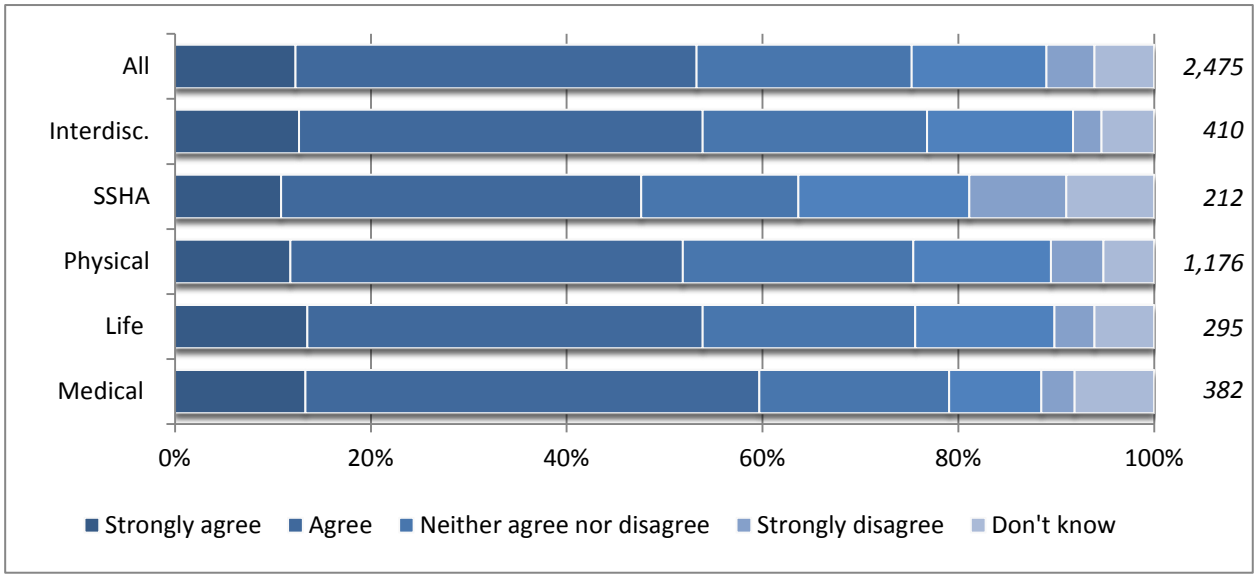
The vast majority of workshop participants indicated they did not have access to an institutional OA fund to support OA publishing, with only a few participants having an OA publishing support

mechanism set up at their institution. Some participants feared that the author-pays model was worse than the subscription model, as researchers with limited funds would not be able to publish at all. The issue of PhD student publications was also raised, as this group may not have the necessary money to pay for Gold OA.

4.4 Perceived influence of OAR on scholarly publishing

Focus group and workshop participants were appreciative of the needs of publishers, such as the organisation of the peer review process, the provision of citation-linking and the availability of back volumes, which have been previously identified in related studies (Swan and Brown, 2003). Furthermore, a substantial number of respondents to the phase 1 survey held the perception that OAR may affect subscription-based journals. The strength of this perception varied according to broad disciplinary groupings, with respondents from the Social sciences, humanities & arts least likely to hold this perception, as illustrated in Figure 4.8.

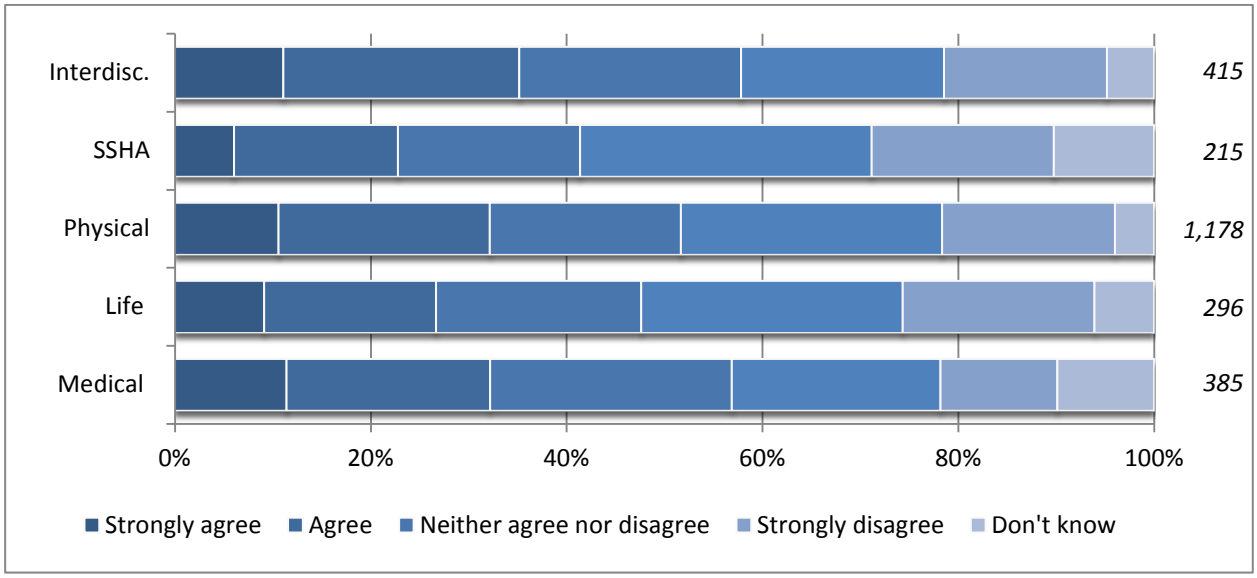
Figure 4.8 Does OA pose a challenge to subscription journals?



(Weighted total numbers of responses given in italics)

Looking more specifically at the peer review process the perception that OAR may challenge the peer review process was mixed, with a smaller, though not insignificant, number of researchers (31%) agreeing with this perception. Again, perceptions varied significantly according to broad disciplinary grouping *(Figure 4.9).* Researchers from the Social sciences, humanities & arts were least likely to hold this perception, which is consistent with their view regarding the perceived impact of OAR on subscription journals. Peer review was also acknowledged as being highly valued by the phase 2 workshop participants, a view that was consistent across the four disciplinary groups, despite a number of flaws in the system being highlighted, such as the fact that the quality of peer review varies from one journal to another, or peer review lags, for example.

Figure 4.9 Does OA pose a challenge to the peer review system?



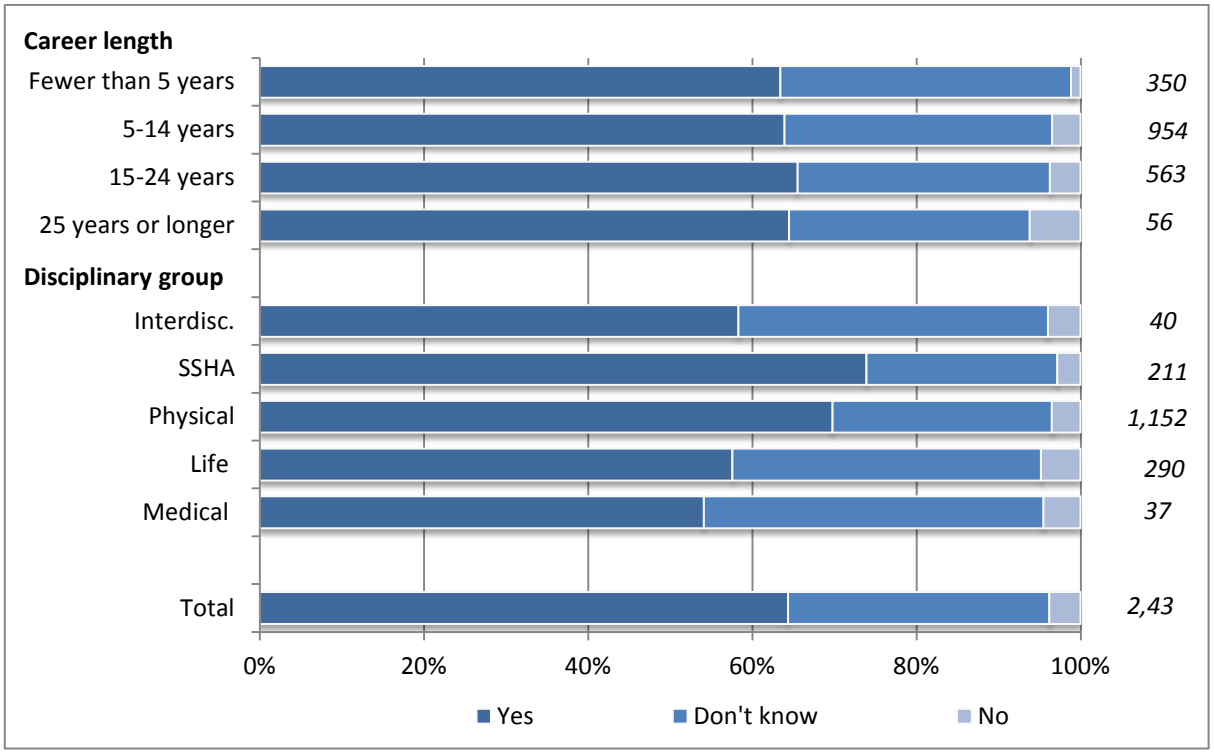
(Weighted total numbers of responses given in italics)

Opinions on the challenge of OAR to scholarly publishing were reinforced in the workshop, where participants also voiced concerns over the future of scholarly publishing if OAR were to become established. Their main concern was that journals may suffer from the availability of OA material and that OAR could potentially adversely affect the current scholarly publishing system. Researchers in the workshop also felt that in order for individual repositories to become an important part of the scholarly communication system they needed to gain a positive reputation. The most obvious way for OAR to gain researchers' trust about the quality of the material held would be to populate them with peer-reviewed articles. Researchers felt that as long as the peer review process was an integral part of the publication process, i.e. that the major mechanism for quality control within a discipline lies with journal editors and reviewers, OAR and peer-reviewed journals would need to co-exist in some mutually beneficial way.

The phase 1 survey found that there was a role for repositories in the scholarly communications landscape, with just 4% of researchers thinking that there was no such role for repositories *(Figure 4.10).* Perhaps unsurprisingly, researchers from the Social sciences, humanities & arts, and the Physical sciences & mathematics were more likely to think that there is a role for OAR; however, these are disciplines with a well-established tradition of OAR.

The main differences between research experience groups shown in Figure 4.10 are in the proportions who think that repositories do have a role in scholarly communications compared to those that have no opinion on the matter. These differences appear small, but are statistically significant ($p<0.01$). The more experienced respondents were, the more likely they were to have a view on this, and, in particular, the more likely they were to consider that there is no role for repositories.

Figure 4.10 Is there a role for OAR in the scholarly communication system?



(Weighted total numbers of responses given in italics)

Despite the fact that researchers in the phase 1 survey acknowledged problems in accessing scholarly literature *(see Section 3.1.3),* workshop participants felt strongly that, as authors, they needed scholarly journals, arguing that researchers depend on high quality peer-reviewed articles (though it is unclear whether this is for scholarly communication or for professional reward). Participants thought that OA may lead to fierce competition between publishers, resulting in fewer journals in which to publish, and, possibly, higher subscription/publication costs. Workshop participants thought that it would be difficult to persuade publishers to cooperate with OAR. The general feeling was that it was better for publishers to open their archive after a period, in the way PMC operates, because in this way readers could be sure about the quality of the papers accessed. The perceived downside of this mechanism was the effect on libraries and subscriptions; participants thought that libraries may not keep their subscriptions if it was known that material would be free after a period, and the whole scholarly publishing system could collapse. This issue is known to be a concern for libraries, although the effects of increasing OA on library practices are, at present, inconclusive (Creaser, 2010). There is no immediate evidence of libraries cancelling subscriptions because of OA, although this could play a part in future decisions.

Researchers identified a wider set of benefits of the peer review system beyond the reputation of an individual resource, be it a journal or OAR, in relation to its function as a quality control indicator for authors wishing to publish high quality articles. Participants in the phase 1 focus groups talked about the importance of having their research evaluated by peers in terms of gaining confidence in their research through reviewers' feedback. Some participants expressed reluctance, therefore, to disseminate their research in non-peer-reviewed forums owing to the perceived need for quality control of their own research. Participants in the Physical sciences & mathematics focus group, in particular, felt strongly that there was an interrelationship between feeling confident in the quality of one's own research or reputation and willingness to deposit

non-refereed papers in a subject-based open access repository, such as arXiv. Sub-disciplines that have a highly collaborative research culture, such as experimental particle physics where research is conducted in working groups or molecular biology where research is organised around laboratories, are likely to have some form of internal quality control (Fry, 2006). This is one factor that might explain why, in the phase 2 survey, respondents from the Physical sciences & mathematics were more likely to have deposited a pre-print into an OAR than respondents from the other broad disciplinary groups *(Figure 3.20, page 38).* Results from the phase 2 survey indicate that peer review or 'confidence in one's own research' is not the only factor, however, and may not be the most influential. Figure 3.20 *(page 38)* illustrates that researchers from the Social sciences, humanities & arts exhibited similar practice to the Physical sciences & mathematics in terms of being more likely to submit a pre-print to an OAR than those in the Medical sciences, the Life sciences or Interdisciplinary researchers.

Phase 2 workshop participants from the Social sciences, humanities & arts asserted that 'insecurity about how their research will be perceived by peers' was one of the reasons why they do not use pre-prints to disseminate their research, which seems to contradict the broader picture illustrated in Figure 3.20 and warrants further exploration. Researchers may be reluctant to disseminate research which has not been through a peer review process, be that internally (within a group, for example) or externally (through a blind peer review process typically associated with journal publication), and thus which has not been validated. Participants from various sub-disciplines in the biomedical sciences highlighted the 'competitive culture' that exists between research groups and explained that this was an influential factor in deterring them from disseminating their research before it had in effect been published.

Phase 2 workshop participants generally agreed that good science needs a rigorous system of peer review and that, in turn, diligent reviewers contribute to the quality and reputation of journals. The general feeling was that peer review is not only a way of weeding out bad papers, but also improving the quality of good papers as long as reviewers are carefully selected and perform their role responsibly. Good quality peer reviews were described as very important to the scholarly publishing system, but that diligent and knowledgeable reviewers were a scarce resource in some fields and were often inundated with requests. Participants from the Engineering group, in particular, described being over burdened by their reviewing activities and the sense of professional duty that makes it difficult to say 'no' to review requests. In the time pressured context of reviewing an article and being responsible for checking its bibliography, (which requires access to the articles cited), OA was seen as important. There is anecdotal evidence that some publishers provide reviewers with free subscriptions to facilitate their task of checking the accuracy of the literature referenced.

To a certain extent the findings from Phase 1 and 2 can be collectively interpreted as highlighting scholars' commitment to peer-reviewed journals and to the quality approval stamp they bring through the peer review system. A number of current developments in specific disciplines indicate that these findings are neither anecdotal nor isolated. For example, in the High Energy Physics community, innovative ways are being developed to maintain both the various value-added services (one of which is the peer review system) offered by the scholarly journal publishing industry and sound business models in an open access environment (Bianco *et al.*, 2007). The SCOAP model[11], whereby a consortium of private and public institutions (funding agencies, libraries, research laboratories etc.) aims to provide open access to

[11] SCOAP: Sponsoring Consortium for Open Access Publishing in Particle Physics

peer-reviewed literature while maintaining the integrity of journal publishers through the payment of a unique subscription, is an example of scholars' willingness to see the publishing industry developing alongside repositories.

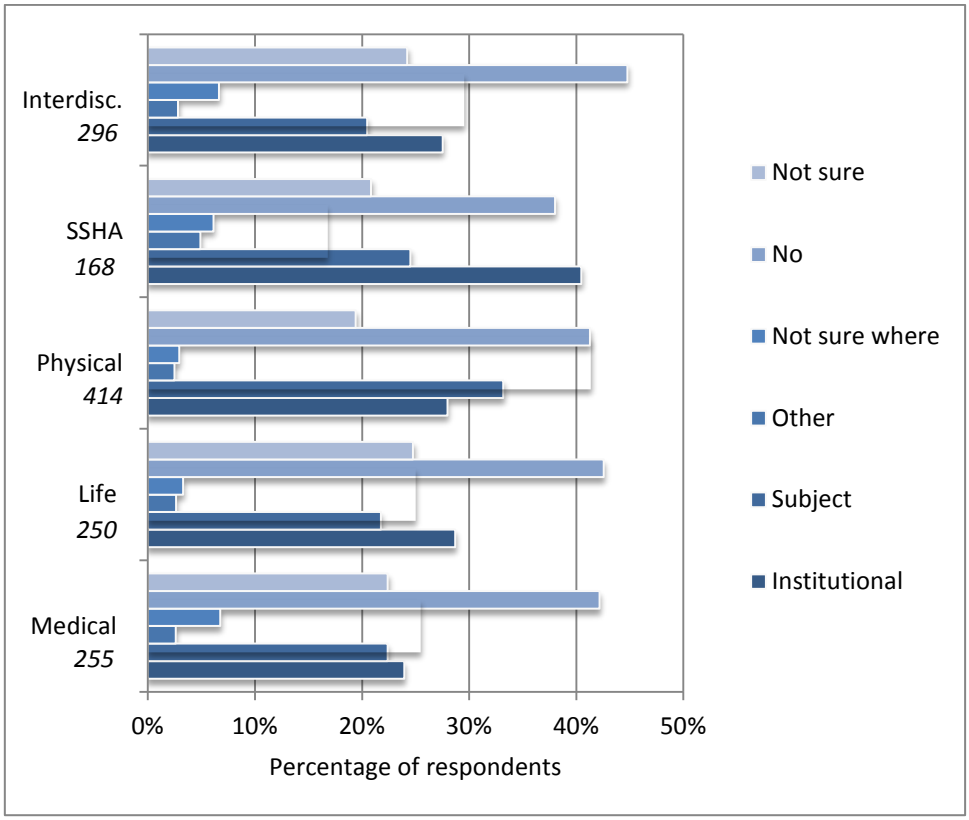
Despite the fact that both phase 1 focus group and phase 2 workshop participants recognised that current scholarly communication through peer-reviewed journal publications may have flaws (e.g. lengthy publication lags in some disciplines, cost of journal subscriptions etc.) the overall perception was that the system generally works well, and researchers did not express a strong desire to change it. However, it was noted that the current scholarly communication system was generally focused on authors' and publishers' needs, rather than the needs of readers. This was expressed most clearly in the participatory workshop, notably by the Engineering and Chemistry groups. Despite the fact that most readers are also authors, it was felt that readers' needs were not taken into consideration in the way scholarly communication operates. In other words, workshop participants expressed the feeling that the way scholarly communication currently operates does not serve the readers' community as well as it serves the authors' community, though the two are obviously intrinsically linked.

4.5 Disciplinary similarities and differences

A number of motivations and barriers to the adoption of OAR depend on the discipline of respondents. For instance speed of publication appears very important in some branches of Physics, with the ability to date-stamp research important to this community – the ability of OAR to accept deposit of pre-print material is therefore important. Other disciplines were much more focused on peer-reviewed material. Many focus group participants tended to conflate OAR with pre-prints, and this misconception may well act as a barrier to deposit in OAR.

Findings from the phase 2 survey presented in Figure 4.11 show that the decision as to the type of repository in which to place their article also seems to be influenced by broad disciplinary group. Authors from the Social sciences, humanities & arts were more likely to have deposited in an institutional repository than authors from the Medical sciences; however, authors from the Physical sciences & mathematics were more likely to have deposited an article in a subject-based repository than authors from the Medical sciences or the Life sciences.

Figure 4.11 Type of repository chosen for deposit, by subject



(Weighted total numbers of responses given in italics)

The general perception across the focus groups and participatory workshops indicated a slight preference for subject-based repositories over institutional ones, however this preference does not necessarily mean that subject-based repositories are the only way forward; in fact, focus group participants did appreciate the role of institutional repositories. Nonetheless, it would seem that subject-based services are important and should be developed if institutional, rather than subject-based, repositories become the most common type of OAR. At the workshop, opinions were very much divided across the four participating broad disciplines, and within each individual discipline, with regards to preferred repository types for research material. The Engineering group thought subject-based repositories 'made more sense' or were 'better', in that they provided better/greater visibility to research outputs – although it was ultimately agreed that this depended on how well outputs were indexed and made visible to search engines. The rationale behind the preference for subject repositories was that Engineering participants thought that researchers would not go and check what is new in an institutional repository whereas they would probably regularly check a subject repository. The visibility of research works seemed to be of particular importance to researchers in the Engineering and Chemistry groups. Visibility often meant being indexed by Google and Google Scholar. Conversely, the Earth, marine & environmental sciences group overall favoured institutional repositories over subject repositories, although this was not unanimous and, indeed, one participant in this group expressed a strong preference for a subject repository, on the basis that repositories tied to institutions may be biased (it was unclear what was meant by 'biased' but possibly that institutions have their own agenda, such as research management and showcasing institutional research). The Chemistry group also favoured institutional repositories over subject repositories despite the fact that they seemed to be regular users of PubMed Central. Interestingly, some Chemistry participants said that they did not need repositories because their papers were

already available/accessible in PubMed Central. When they were told that PubMed Central was also a subject-based repository, they indicated that they liked the procedures used to place material into PMC because '[they] don't do anything, it appears there!' Chemistry participants did not recognise PubMed Central as a repository *per se*. As mentioned earlier, there seems to be a disconnection between perception of what a repository is and the repositories that they actually use. Chemistry participants agreed that if all OAR worked like PubMed Central, i.e. publisher deposit of published articles, they would be happy to use repositories on a large scale.

5 Discussion

The overall goal of the Behavioural research project was to develop an understanding of the perceptions, motivations and behaviours of authors and readers towards stage-two manuscript repositories. In order to achieve this understanding a series of research objectives were addressed (given in the introduction) across two phases of research using a mixed methods approach (described in Appendix 1). The findings of the Behavioural research are discussed below and have been presented according to the specific research objectives. The different perspectives of authors and readers are highlighted where appropriate.

Choices authors and readers make in locating and selecting sources in the context of publication and dissemination, and information seeking behaviours and the major influences underpinning those choices

There is a close interrelationship between dissemination and information behaviours, with choices about how, where and when to disseminate research influencing ways in which researchers seek to locate and access that research. To obtain a more holistic picture of researchers' perceptions, motivations and behaviours, therefore, it is necessary to view these two processes in relation to one another.

The findings show that, despite the co-evolution of the scholarly communication system with advanced digital technologies, which has underpinned the emergence of new forums for disseminating and locating research outcomes in recent years, the central role of the peer-reviewed journal article is not diminishing. This confirms findings from related studies that identified scholarly articles as accounting for over 90% of scholars' information sources (Tenopir *et al.*, 2009; King *et al.*, 2009). It should be noted; however, that on average the volume of peer-reviewed articles read by researchers does vary according to broad-based disciplinary groupings, with researchers from the Medical sciences and the Life sciences collectively reading more peer-reviewed journal articles per year, than researchers from the Physical sciences & mathematics and the Social sciences, humanities & arts.

The creation of peer-reviewed journal articles takes up a considerable proportion of academic researchers' total written output and researchers reported that there is a tendency amongst universities to stipulate the minimum number of articles that should be published each year. Whilst career advancement was ranked by researchers as an important factor in choosing to disseminate their research via publication in a peer-reviewed journal, wide-spread visibility of their research and dissemination to specific target audience(s) were ranked as being more important. Results from the participatory workshop indicate that researchers have knowledge of the readership for each journal title and that this influences their choice of journal for publication. Researchers who were in the early stages of their careers (i.e. fewer than five years research experience) were more likely to rank career advancement as being the most important influence in choosing to publish in a peer-reviewed journal. The reputation or esteem of a journal was also an important influencing factor in terms of where researchers choose to publish and despite the documented limitations of the journal impact factor as a measure of reputation or esteem, anecdotal evidence from the qualitative research suggests that in some European countries journal impact factors are becoming more important in terms of evaluating the quality of researchers' outputs. The qualitative results also indicate that there is, at times, a tension between institutional-based motivations, e.g. reputation and career advancement, to publish in peer-reviewed journals and audience-based motivations to make research outcomes as widely

available as possible. OA was perceived in some circumstances as an effective solution to this tension, but focus group or participatory workshop discussions invariably highlighted confusion amongst researchers about the distinction between OA journals, which may have a journal impact factor, and OAR, which do not. Similarly, researchers vary in the extent to which they associate OA and OAR with the peer review process, and there are disciplinary differences in this association. Researchers from the Life sciences and the Medical sciences are least likely to associate 'non-peer-reviewed' with OA, whereas researchers from the Physical sciences & mathematics were more likely to associate OA with 'non-peer-reviewed', and 'not the published final version'.

Date-stamping ideas can also be an important function of researchers' dissemination and publication activities. Conferences were seen as the main forum for this purpose, particularly those conferences that peer review papers and publish proceedings. In disciplines where impact factors are important and conferences do not lead to published outputs, e.g. proceedings, then conferences are not seen as a solution for date-stamping, which was the case for the Engineering researchers in the participatory workshop.

Researchers typically select a narrow range of information resources that they use on a regular basis to locate research-based sources, and for most disciplines the search strategies deployed within the scope of these resources are highly individualised. Related studies have found a contrast to this norm in some disciplines, primarily within the biomedical sciences, where information seeking is a well-defined and structured process that follows a recognisable disciplinary pattern (Blake and Pratt, 2002). Traces of disciplinary differences in the information resources used can be identified in our findings, in that bibliographic literature databases are more likely to be used by researchers from the Life sciences, and subject-based portals/repositories are more likely to be used by researchers from the Medical sciences, whilst researchers from the other broad-based disciplinary groups are more likely to start their search strategy from a publisher's journal platform or Google/Google Scholar. Career length also seems to play an influential role, with less experienced researchers (i.e. fewer than five years) being more likely to initiate a search using Google Scholar and the most experienced (i.e. 25 years or more) being more likely to browse print-based journals.

In general, researchers are unlikely to go directly to a repository to search for journal articles, with Google and Google Scholar being the most likely route by which researchers locate material within OAR. The exception to this was those researchers who reported using well-established subject-based repositories, such as PMC, arXiv, RePEc or SSRN, and who went directly to their respective subject-based repository to locate journal articles. It was noted in the participatory workshop, however, that OAR tend to have poor search and navigational interfaces, with institutional OAR being particularly difficult to search and navigate. Informal scholarly networks, e.g. asking the author, peers at other institutions or close colleagues, were also an important channel for locating and accessing relevant journal articles.

Highlights

Authors

The central role of the peer-reviewed journal article in academic careers is not diminishing despite the emergence of digital scholarship and novel modes of dissemination.

Important factors influencing researchers' decisions to disseminate via peer-reviewed journal articles are wide-spread visibility of their research and dissemination to specific target audience(s), with career advancement being ranked as less important than these two factors.

Researchers who were in the early stages of their careers were more likely to rank career advancement as being the most important influence in choosing to publish in a peer-reviewed journal.

Researchers have knowledge of the readership for journal titles and this influences choice of journal for publication.

Readers

Researchers typically select a narrow range of information resources that they use on a regular basis to locate research-based sources.

Traces of disciplinary differences in the information resources used were found, with bibliographic literature databases more likely to be used by researchers from the Life sciences, whilst subject-based portals/repositories were more likely to be used by researchers from the Medical sciences. Researchers from other broad-based disciplinary groups were more likely to start their search strategy from a publisher's journal platform or Google/Google Scholar.

Less experienced researchers (i.e. fewer than five years) were more likely to initiate a search using Google Scholar, whilst the most experienced researchers (i.e. 25 years or more) were more likely to browse print-based journals.

Researchers are unlikely to go directly to a repository to search for journal articles, with Google/Google Scholar being the most likely route by which researchers locate material within OAR.

Researchers who reported using well-established subject-based repositories, such as PMC, arXiv, RePEc or SSRN go directly to their repository of choice to search. It was noted, however, that repository interfaces were difficult to navigate.

The influence of purpose of reading, e.g. current awareness, proposal and article writing, on readers' behaviours in relation to repositories

The main mediating factor between purpose of reading and repository behaviour appears to be the article version being sought and clarity of the article version found. In general, researchers perceived versions of articles found in an OAR differently depending on their purpose of reading.

Researchers seeking articles in order to cite them in their own work are most likely to try to locate the published final version, and whilst many would only cite the published final version, due in part to practical considerations such as the lack of page numbers and DOIs, a smaller number would be willing to cite any version of the article that they managed to access. There was also a sense amongst workshop participants that authors would have a preference for the

published final version of their articles to be cited, not least, because of increasing institutional/national policy emphasis on measuring the quality and impact of both individual researchers' and institutions' research outcomes and outputs. The mechanisms for measuring quality and impact often include citation counts based on resources such as the *ISI Citation Indexes*, which index published final versions of articles. The peer-reviewed journal article could be described, therefore, as being 'locked-in' to mechanisms for recognition and reward, which results in a self-perpetuating cycle that influences readers' perceptions of the role of OAR in their information seeking strategies.

Accessing the published final version of an article, however, appears to be less critical for writing research outputs other than peer-reviewed journal articles. This was also the case for current awareness activities, particularly in novel areas of research where there may be only a limited number of journal articles that have been published on the topic of interest.

There were notable disciplinary differences in the article version likely to be sought and its subsequent uses. Researchers from the Medical sciences and the Social sciences, humanities & arts were most likely to consider the article version important and least likely to 'trust' versions of articles held in a repository unless it was very clear to them that they had accessed the published final version. On the other hand, there was a greater acceptance of pre-prints by researchers from the Physical sciences & mathematics than researchers from the other broad disciplinary groups. Uncertainty about article versions located in repositories seemed to be mitigated to some extent by author or journal reputation.

Commonly held perceptions amongst authors and readers in relation to open access repositories, and the ways in which such perceptions influence publication and dissemination, and information seeking behaviours

Researchers' perceptions with regard to OAR are intricately interwoven with the notion of peer-reviewed journal articles being at the core of the scholarly communication system. For most researchers, in most disciplines, this is the baseline point of departure for engaging with new forums for dissemination and publication of research, such as OAR. This is perhaps to be expected, given the complex set of relations represented by publications in highly esteemed (by both authors and readers) peer-reviewed journals. For authors, there is the balance between reaching the target audience(s), wide-spread visibility of research, gaining a reputation in their area of research and meeting their employing institutions' targets in terms of outputs (with the quality of the output often being used as an indicator of the quality of the underlying research by evaluation mechanisms). For readers who are reading for the purpose of writing a peer-reviewed published journal article themselves, consideration needs to be given to the quality and authority of the sources being cited, e.g. the reputation of the author in combination with the reputation of the journal where the article has been published, the extent to which the source can be formally cited, e.g. which version of the article they have accessed and the norms in the field for citing pre/unpublished material, and the extent to which the sources they cite are aligned to the scope and requirements of the journal (as determined by editors, editorial boards and reviewers).

This collective process of evaluating where to publish and what to cite places a high value on peer review as being the primary mechanism for quality control in a discipline. Put another way, researchers linked good quality peer review with highly esteemed journal titles. This extends beyond contributing to the reputation or esteem of a journal title to include its function as

validating research and contributing to the quality of individual articles, where the reviewer is particularly knowledgeable, thorough and constructive in their feedback. The qualitative results indicate that this process is particularly important to researchers who do not routinely do research that is collaborative in nature.

The focus groups and participatory workshop enabled researchers' perceptions of OAR to be teased out in some detail. Researchers felt that in order for an OAR to become a central resource in their discipline then it would need to develop a positive reputation. It was felt that the most obvious way for OAR to gain researchers' 'trust' in the quality of material held would be to populate them with peer-reviewed journal articles. Although, conversely, there were some reservations amongst authors about having their peer-reviewed published journal articles held in an OAR with other content of variable quality, which they felt in some way might influence the perceived quality of their own articles. Researchers felt that as long as the peer-reviewed journal article was the major mechanism for quality control in their discipline, then OAR and peer-reviewed journals would need to co-exist in some mutually beneficial way.

There are two main types of OAR; subject-based and institutional, with each type having emerged in a different context and with a different set of goals. In some disciplines there are *de facto* centralised repositories, such as arXiv in physics, RePEc in economics and PMC in the medical sciences, while in other disciplines the repository landscape is less well established. There is an accompanying uncertainty about researchers' preferences regarding type of repository (both in terms of authors and readers). In the phase 1 survey, 44% of authors expressed a preference for placing their article(s) in a subject-based repository, with 23% of authors preferring an institutional repository, while 21% expressed no preference. This question was followed up in the phase 2 survey, with authors asked to indicate in which type of OAR a version of their article(s) had been placed; the results in the second phase show a distribution that is more equal across the two main types of repository (32% selected a subject-based repository and 36% selected an institutional repository into which to place their article(s),with 12% doing both). This apparent shift between what was preferred in phase 1 and actual practice reported in phase 2 may be due, in part, to the increase in mandates noted over the intervening period. There were disciplinary differences regarding preferences for type of repository. Researchers from the Social sciences, humanities & arts were more likely to have deposited an article into an institutional repository than authors from the Medical sciences; whereas researchers from the Physical sciences & mathematics were more likely to have deposited an article into a subject-based repository than authors from the Medical sciences or the Life sciences. Phase 2 workshop participants were divided with regard to preferred repository types for depositing material, with diverse opinions within each individual discipline and across the four broad disciplinary groups.

The nature of the motivation to self-archive seems to influence into which type of OAR authors are likely to place (or give permission to have placed) a version of their article(s). If the motivation is voluntary, requested by a co-author, or an invitation by a publisher then authors are more likely to choose a subject-based repository. If, on the other hand, authors are invited by repository/library staff, or were mandated by their employer, then they are more likely to choose an institutional repository. The pattern is less clear in relation to funder mandates, which appears to result in an even split between subject-based and institutional repositories, with the remaining one third of authors reporting that they had submitted to both or that they were unsure as to which type their article(s) had been submitted.

It might be expected that the most significant change in authors' behaviour with regard to repository deposit is likely to be in response to open access mandates, as funding agencies in most disciplines have developed open access policies, or at least position statements on open access in recent years. Increasingly, institutions are following suit by implementing mandates in relation to their institutional repository, although they are not necessarily enforced effectively. Where funding agencies do stipulate mandatory deposit or have position statements, they may also encourage grant applicants to apply for the necessary funds to make articles publicly available via open access e.g. financial support for 'author pays' mechanisms. It is of interest, therefore, that funder and institutional mandates were both considered to be relatively unimportant as motivators for repository deposit by survey respondents. Researchers' perceptions of how they would respond to the enforcement of multiple mandates was explored in the phase 2 participatory workshop and the anecdotal evidence seems to suggest that if institutions were to enforce mandates then researchers would feel compelled to prioritise their employer's mandate over other types of mandate.

Highlights

Authors

The process of evaluating where to publish places a high value on peer review as being the primary mechanism for quality control in a discipline.

Where reviewers are particularly knowledgeable and constructive in their feedback, the function of the current peer review system goes beyond contributing to the reputation or esteem of the journal, and plays an important role in validating research and improving the quality of individual articles.

There were some reservations amongst authors concerning having their peer-reviewed published journal articles held in an OAR with other content of variable quality, which they felt would have a negative influence on the perceived quality of their own articles.

In terms of authors' preference for type of repository into which to place a version of their article(s), there was an almost equal distribution between subject-based and institutional repositories, with a slight shift between phase 1 and phase 2 towards institutional repositories.

Researchers consider funder and institutional mandates to be relatively unimportant as motivators for placing their journal article(s) into an OAR.

Readers

- Researchers' selection decisions regarding what to cite places a high value on peer review as being an indicator of quality control in a discipline.
- The extent to which researchers perceive a source can be formally cited depends on the version of the article they have accessed and the disciplinary norms for citing anything other than the published final version.
- It was felt that the most obvious way for OAR to gain researchers' 'trust' in the quality of material they hold would be to populate them with peer-reviewed journal articles.
- Researchers felt that as long as the peer-reviewed journal article was the major mechanism for quality control in their discipline, then OAR and peer-reviewed journals would need to co-exist in some mutually beneficial way.

Researchers' green Open Access experience both as authors and readers

Across the phase 1 and 2 surveys, approximately half of the authors reported having placed (or having had placed on their behalf) a version of their journal article(s) in an OAR. Having an article placed in an OAR by someone else, e.g. a librarian, publisher or co-author, was more common than authors placing a version of their article(s) in an OAR themselves. This pattern varied according to broad disciplinary group; authors from the Physical sciences & mathematics and the Life sciences were more likely to place a version of their article(s) in an OAR themselves, whilst authors in the Medical sciences or the Social sciences, humanities & arts were more likely to have a version of their article(s) placed in an OAR by somebody else.

Once readers have identified a relevant journal article they will typically try to locate the published final version of it, although this varies according to purpose of reading and discipline. The clarity of repository metadata and readers' ability to distinguish between the different versions of an article seems critical to how the 'quality' (e.g. whether it is authoritative) of repository content is perceived. Researchers reported that, whilst the difference between a pre-print and a published final version was reasonably clear, it was much more difficult to distinguish between a submitted stage-one article and an accepted stage-two article. Owing to the diversity of types of sources held in OAR (i.e. published or unpublished and peer-reviewed or non-peer-reviewed) readers are often uncertain about the extent to which material accessed is authoritative, unless the author, journal, and version accessed is known to them. Therefore, there appears to be a general perception that material held in OAR is not citable. It should perhaps be noted here that repositories vary in the extent to which they hold full versions of journal articles or mainly provide only the metadata, which may result in different experiences with regard to readers being able to determine which article version they have identified.

Where readers are unable to access a published journal article that they have identified as potentially relevant, then they are likely to seek an open access version, but there are disciplinary differences. This course of action was more likely to be adopted by researchers from the Physical sciences & mathematics and, conversely, these researchers also seemed more likely to forego locating and reading an article if it was not easily accessible to them.

The article version being placed into OAR is most likely to be the published final version (publishers' PDF file), followed in popularity by the author's final peer reviewed accepted version, also called a stage-two accepted manuscript. There was variation in this pattern according to broad disciplinary group: authors from the Life sciences were more likely to place (or have placed on their behalf) the published final version, with authors from the Physical sciences & mathematics and the Social sciences, humanities & arts more likely to place (or have placed on their behalf) a pre-print or author's final peer-reviewed accepted version. Authors from the Medical sciences were most likely to be uncertain about which version of their article had been placed in an OAR, which correlates with the high proportion of Medical science researchers who reported that their article(s) had been placed in an OAR by someone else.

The phase 2 survey asked authors to specify whether or not they had placed their article(s) into an OAR themselves and almost one-third indicated that someone else had made their work available in an OAR on their behalf. In terms of authors' experience of depositing a copy of their article(s) in an OAR themselves, the survey results suggest that authors do not generally experience much difficulty. Breaking down the process of self-deposit into each component step, authors found that identifying the correct version was the easiest step and the process of checking that the publisher allows the article to be placed in an OAR the most difficult. The

amount of time required to place an article in an OAR as estimated by authors varied a great deal, with the most frequently reported estimates being between 5-15 minutes and 15 minutes to one hour, whilst lower (less than 5 minutes) and upper estimates (1-2+ hours) were also selected. The process of self-archiving was explored in more detail in the participatory workshop and was described by some participants as 'tedious' and 'time-consuming', with the need to check and resolve potential copyright issues adding a layer of complexity, and was perceived as a significant barrier.

It would appear that checking publishers' OA policies is the main barrier to authors managing the self-archiving process themselves. The qualitative results have also highlighted that authors are often unclear about publishers' OA policies and the nature of any copyright agreements that they have signed, with the perception that they are more constrained in terms of self-archiving than they actually are in practice. This indicates that authors might best be supported by repository/library staff and publishers by making self-archiving policies clearer. It also illustrates the importance of services such as SHERPA/RoMEO[12] that aim to clarify publishers' copyright and self-archiving policies.

Highlights

Authors

Across the phase 1 and 2 surveys approximately half of the authors reported having placed (or had placed on their behalf) a version of their journal article(s) in an OAR.

Authors from the Physical sciences & mathematics and Life sciences were more likely to place a version of their article(s) into an OAR themselves, whilst authors in the Medical sciences or the Social sciences, humanities & arts were more likely to have a version of their article(s) placed in an OAR by somebody else.

The published final version is the version most likely to be placed into an OAR. Authors from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to place (or have placed on their behalf) a pre-print or author's final peer-reviewed accepted version.

Authors from the Medical sciences were most likely to be uncertain about which version of their article had been placed in an OAR, which correlates with the high proportion of Medical science researchers who reported that their article(s) had been placed in an OAR by someone else.

Almost one-third of phase 2 survey respondents indicated that someone else had made their work available in an OAR on their behalf. In terms of authors' experience of depositing a copy of their article(s) in an OAR themselves the survey results suggest that authors do not generally experience much difficulty. None-the-less, findings from the phase 2 survey indicate that checking publishers' Open Access policies is the main barrier to authors managing the self-archiving process themselves.

[12] http://www.sherpa.ac.uk/projects/sherparomeo.html [accessed 29.09.11]

Readers

The clarity of repository metadata and readers' ability to distinguish between the different versions of an article seems critical to how the 'quality' (e.g. whether it is authoritative) of repository content is perceived.

Researchers reported that whilst the difference between a pre-print and a published final version was reasonably clear, it was much more difficult to distinguish between a submitted stage-one article and an accepted stage-two article.

Researchers' satisfaction with a version other than the published final version of a journal article is closely related to the purpose of their reading. Researchers seeking articles in order to cite them in their own articles are most likely try to locate the published final version.

Researchers from the Medical sciences and the Social sciences, humanities & arts were most likely to consider the article version important and least likely to 'trust' versions of articles held in a repository unless it was very clear to them that they had accessed the published final version.

Where readers are unable to access a published journal article, then they are likely to seek an open access version. Researchers from the Physical sciences & mathematics were more likely to follow this course of action than researchers from other disciplinary groups.

Researchers from the Physical sciences & mathematics were more likely to forego locating and reading an article if it was not easily accessible to them.

Authors' perceptions of the values/benefits of open access repositories in relation to the effort involved in making their work available via them.

The phase 1 survey looked at the relationship between various motivations typically associated with the dissemination and publication of research and their association with authors' motivations to place material in an OAR. As noted previously, reaching target audience(s)/widespread dissemination of research were very important considerations for choosing to publish articles in peer-reviewed journals and researchers also indicated that these were equally as important in terms of making journal articles available via OAR. Speed of dissemination and increased citations were ranked slightly lower as motivations for placing articles into OAR, but there were some interesting disciplinary differences in relation to these motivations. Speed of dissemination was less important for researchers in the Medical sciences, than researchers in other broad disciplinary groups, and researchers from the Social sciences, humanities & arts were more likely to rank the possibility of increased citations as very important than researchers from the other broad discipline groups.

In terms of weighing the perceived benefits of OA against the effort of placing a copy of their journal article(s) in an OAR, the majority of the authors surveyed in phase 2 felt that it was either 'definitely' or 'probably' worth the additional work involved. However, there was a small proportion (7%) of authors who indicated that they felt that the benefits did not outweigh the extra effort. There were also differences in opinion according to broad disciplinary group. Authors within the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to feel that placing article(s) in an OAR was 'definitely worthwhile', compared to authors in the Life sciences and the Medical sciences.

Participatory workshop discussions provided a different viewpoint from that of the wider author view represented in the phase 2 survey findings. Participants reached a consensus view that,

whilst in theory they support the ethos of OA, in practice the responsibility for achieving it lies with other stakeholders in the scholarly communication system. The primary reason given for this view was that the publication process itself is very time-consuming for researchers and that researchers' responsibility for publishing work and making it available should end at the point of publication.

Impact of authors' perceptions of publishers' open access embargoes on their open access repository behaviours

There appears to be a lack of awareness of publishers' open access embargo periods, amongst authors at least, with just over half of authors surveyed in phase 2 stating that they did not know or could not remember what embargo period, if any, was enforced by the publisher when they placed their article in an OAR. One third of authors indicated that no embargo period was specified by the publisher when they deposited their article in an OAR, with a further 6% stating that the embargo period was 6 months or less.

In terms of the perceived influence of length of embargo periods on authors' motivation to make their article(s) available on open access, almost one-third of authors who had placed their article(s) in an OAR indicated that if the publisher were to specify a shorter embargo period than the one that had originally been in place, then this would make them more likely to make their article(s) available via open access. An almost equal proportion of authors, however, stated that it would make no difference to them.

Where the publisher embargo period was longer than originally specified, again, almost one-third of authors perceived that this would make them less willing to make their article(s) available via open access, with just over one-third stating that it would make no difference to them.

Given the generally supportive attitude of researchers towards the ethos of OA, those authors who specified that shorter publisher embargo periods would make them more likely to make their article(s) available via open access, with longer periods making them less likely to do so, represent a perspective that might be expected. That an almost equal proportion of authors perceived that the length of publisher embargo periods would not influence their behaviours indicates the difficulty in making generalisations about the impact of publishers' embargo periods on authors behaviours in relation to OAR. This result could also be seen as reinforcing some of the fine-grained detail to come out of the qualitative results, indicating that there are those authors who feel that OA is important, but that its practical realisation is not their responsibility and there are other stakeholders in the scholarly communication system that they perceive to be better placed to take up this role.

The role played by open access repositories in the scholarly communication landscape as shaped by author and reader choices

The choices that authors and readers make with regard to dissemination and publication, and their information behaviours currently position OAR on the margins of these activities. There are, of course, disciplinary differences in the relative centrality of OAR to current research practices and these have been highlighted above and in the main findings sections of this report.

There are a number of factors that authors take into consideration when deciding what, when and how to disseminate and publish their research. These primarily revolve around reaching the target audience(s), recognition and reward (both at the level of their discipline/sub-discipline and within their employing institution), and date-stamping ideas.

Current perceptions of OAR amongst authors do not necessarily align with these underlying factors. Namely, OAR have an unknown readership, particularly institutional or broadly-based subject repositories, whereas authors are knowledgeable about, and influenced by, the readership of specific journals. OAR have the potential to maximise potential readership, yet authors are more concerned with reaching their target audience(s) than in maximising the readership for their journal articles, even where this might potentially lead to increased citations. In some disciplines, OAR have been used effectively as a mechanism for date-stamping ideas and certainly rapid dissemination of research is moderately important to researchers, but within many disciplinary communities OAR are not yet perceived as a viable option for date-stamping ideas. Similarly, in some disciplines the publication lags associated with subscription-based journals are deemed to be particularly problematic, and there are examples where OA journals and discipline-specific OAR have been successful at addressing this issue, but there was not a generally held (i.e. across disciplines) perception that such issues were pressing enough to change current dissemination and publication behaviours.

In terms of the accessibility of their published journal articles to researchers beyond Higher Education, i.e. those in the public, private and third sectors, or researchers in countries with emerging and developing economies, this was perceived by workshop participants as an area of concern that OAR could effectively address, but that responsibility for widening accessibility should be taken-up by other stakeholders in the scholarly communication system.

When searching for information sources readers are unlikely to go directly to an OAR, unless they are in disciplines where a subject-based OAR has been adopted to such an extent that it has become a core information resource. They are more likely to come across an OA source via general search tools, such as Google or Google Scholar, which means that researchers are not necessarily going to be aware of the extent to which they are actually using OAR. There appears to be a general perception that material held in OAR is not citable. Since the writing of journal articles seems to account for such a high proportion of reading-related activities, and given that readers prefer to cite the published final version of articles, readers' trust in OAR is yet to be earned. This seems to suggest that the visibility and reputation of OAR need to improve if they are to be positioned more centrally in readers' information behaviours.

This research has highlighted that whilst researchers are aware of OA, there are often misconceptions about the two main routes to achieving it. Researchers are becoming aware of OAR as a concept, but are often uncertain about what their potential role is in the scholarly communication system for their particular discipline. The focus groups and participatory workshop emphasised this observation, with a number of researchers expressing that they 'sensed' that OAR were going to become increasingly important, both in terms of research policy/evaluation and dissemination/information seeking, and felt that they needed to be better informed about the repository landscape.

Highlights

Authors

Current perceptions of OAR amongst authors do not fit well with the myriad factors that they need to take into account with regard to the dissemination and publication of their research. Such perceptions position OAR on the margins of dissemination and publication activities.

OAR have the potential to maximise the readership base for journal articles, yet authors are more concerned with reaching a target audience(s) than maximising the readership. OAR have an unknown readership, particularly institutional or broadly-based subject repositories, whereas authors are familiar with the readership of specific journals.

In some disciplines publication lags associated with subscription-based journals are problematic. There are examples where OA journals and discipline-specific OAR have been successful at addressing this issue, but the generally held perception (i.e. across disciplines) was that these issues were not pressing enough for researchers to change their current publication practices.

Accessibility of published journal articles to researchers beyond the Higher Education sector, or those in countries with emerging and developing economies, was perceived by authors as an area of concern that OAR could effectively address.

Readers

Unless researchers are in disciplines where a subject-based OAR has been adopted to such an extent that it has become a core information resource they are unlikely to go directly to an OAR to search for information sources.

If readers are accessing OA sources via general search tools they are not necessarily going to be aware of the extent to which they are actually using OAR.

Readers prefer to cite the published final version of a journal article and readers' trust in being able to cite material accessed in an OAR is yet to be earned.

The visibility and reputation of OAR need to improve in order for them to play a more central role in readers' information behaviours.

6 Key highlights and conclusions

Over the period of Phases 1 and 2 of the Behavioural research the increase in the number of researchers who reported placing a version of their journal article(s) in an Open Access Repository was negligible.

According to ROAR (Registry of Open Access Repositories) the number of institutional or departmental research repositories worldwide has grown exponentially year on year and in recent years a number of subject-based repositories have come to the fore as *de facto* centralised resources in some disciplines. These developments have been coupled with initiatives and policies to mandate, or at least encourage, researchers to deposit copies of their published journal articles into OAR. These collective developments represent the so called 'Green Road' to Open Access and it might be logically anticipated that an increased level of technical developments and policy activity towards this route to Open Access would be accompanied by a corresponding increase in levels of awareness amongst researchers. Our findings show that about half of all authors have placed, or given permission to have placed on their behalf, a version of their journal article(s) in an OAR, with only a negligible increase in this proportion between the phase 1 and phase 2 surveys.

Researchers who associated Open Access with 'self-archiving' were in the minority.

The research did not explicitly seek to measure any difference in levels of awareness of Green Open Access amongst researchers between Phase 1 and Phase 2 of the research, however, the phase 1 survey findings indicated that whilst there was a general awareness and positive ethos towards Open Access, researchers who associate Open Access with 'self-archiving' were in the minority, although there were disciplinary differences in the strength of this association. Furthermore, confusion amongst researchers is common with regard to the various types of OAR available and the different article versions held. In the phase 1 focus groups and phase 2 participatory workshop researchers also expressed confusion about what might reasonably be expected of them in terms of the array of institutional, funder, and publisher open access policies with which they may need to comply, highlighting that institutional, funder and publisher Open Access policies cannot be developed in isolation from one another.

Open Access is more likely to be associated with 'self-archiving' (Green Road) by researchers in the Physical sciences & mathematics and Social sciences, humanities, & arts than those in the Life sciences and the Medical sciences who are more likely to associate Open Access with Open Access Journals (Gold Road).

Levels of awareness and day-to-day practices in relation to 'self-archiving' vary according to discipline, so that aggregate-level observations need to be interpreted with sensitivity towards disciplinary research cultures. The phase 1 survey findings show that researchers in the Physical sciences & mathematics and some disciplines in the Social sciences, humanities & arts are more likely to associate Open Access with 'self-archiving' (Green Road) than researchers in the Life sciences and the Medical sciences who are more likely to associate Open Access with Open Access Journals (Gold Road), in-line with recent related studies (Björk *et al*, 2010; and Dallmeier-Tiessen *et al* 2010).

There is anecdotal evidence that some researchers consider making journal articles accessible via Open Access to be beyond their remit.

Evidence from the phase 2 participatory workshop suggests that roles and responsibilities in relation to achieving Open Access, regardless of route, are not clear to researchers and some perceive making journal articles accessible via Open Access to be beyond their current remit

vis-à-vis the scholarly communication system. It might be reasonable to anticipate that the deposit process itself could be a barrier to authors' use of OAR, and in this regard the phase 2 survey and participatory workshop gave a mixed picture. Authors from the phase 2 survey who had deposited articles in an OAR themselves and had found that process difficult were in a small minority. This suggests that the deposit process itself is not a barrier for most researchers and is encouraging, but warrants further investigation.

Authors tend to be favourable to Open Access and receptive to the benefits of self-archiving in terms of greater readership and wider dissemination of their research, with the caveat that self-archiving does not compromise the pivotal role of the published journal article.

The phase 2 survey findings show that the mode of activity in which a researcher is engaged, e.g. author 'deposit' or reader 'access' mode, to a certain extent shapes perceptions towards the role of OAR in the scholarly and scientific communication system, although disciplinary culture appears to be a stronger factor. None-the-less, within their respective disciplinary frameworks, researchers' perceptions of OAR, and stage-two manuscripts in particular, as valid and citable information resources/sources or as a viable means by which to disseminate their published research articles are likely to shift, according whether they are in author or reader mode. Evidence from the phase 1 focus groups and phase 2 participatory workshop suggests that, from an author perspective, researchers tend to be favourable towards Open Access and receptive to the benefits of self-archiving in terms of greater readership and wider dissemination of their research. This attitude, however, tends to be based on the caveat that self-archiving does not compromise the pivotal role of the published journal article, which has retained its position as one of the most important channels for contributing to knowledge, and to academic and professional careers, despite the emergence of new forms of digital information gate-keepers in recent years, e.g. digital pre-print archives and the web pages of individual researchers and research groups.

Readers have concerns about the authority of article content and the extent to which it can be cited when the version they have accessed is not the published final version. These concerns are more prevalent where the purpose of reading is to produce a published journal article, and are perceived as less of an issue for other types of reading purpose.

Findings from the phase 2 survey show that from a reader's perspective researchers have concerns about stage-two manuscripts, and all pre-cursor versions, as valid surrogates to the final published article. This is both in terms of trusting the authoritativeness of the article's content (e.g. whether the article has been peer-reviewed), and whether or not the version accessed can be cited if it is not the published final version (e.g. consideration of practicalities such as the inclusion of correct page numbers and the disciplinary norm for citing pre-published versions of journal articles). There was a perception amongst phase 2 workshop participants that readers who are not authors themselves or who are simply keeping up-to-date with developments in their discipline/sub-discipline (i.e. those who are seeking journal articles without the intention to subsequently cite them) are more likely to reap the benefits of being able to access stage-two versions of journal articles. Workshop participants then extrapolated this perception to researchers outside of academia, and it was felt that the benefits of Green OA might have a greater impact in non-Higher Education sectors. Findings from the phase 1 survey show, however, that researchers from hospitals, medical schools, government, industrial/commercial and other types of non-Higher Education institutions are less aware of OAR than researchers from universities, colleges and research institutes. If the benefits of OAR

and stage-two manuscripts are to be realised outside of the Higher Education sector, then there is a need for sector-specific initiatives to raise levels of awareness.

Academic researchers have a conservative set of attitudes, perceptions and behaviours towards the scholarly communication system and do not desire fundamental changes in the way research is currently disseminated and published.

Collectively, the findings point to a conservative set of attitudes, perceptions and behaviours amongst researchers towards the scholarly communication system, at least when an aggregate view is adopted. In the phase 1 focus groups and phase 2 participatory workshop it was often the case that whilst researchers identified flaws in the current journal publication system, e.g. publication lags, quality of peer review, costly subscription fees, there was no indication that they wished to see any fundamental changes in the way research is disseminated and published.

This is more likely to be the case for researchers within Higher Education, than those in other sectors, due to the centrality of the peer-reviewed journal article in academic recognition and reward systems and the complex set of institutional, disciplinary and stakeholder relations that function, intentionally or not, to maintain that centrality. The phase 1 survey findings showed that whilst peer-reviewed journal articles were important to researchers outside of the Higher Education sector the output type of central importance was the research report. When viewed at a finer level of granularity, therefore, there is a differentiation in attitudes, perceptions and behaviours, and subtle evolutions in the scholarly communication system with regard to OAR can be discerned. This rather fragmented co-evolution of the traditional and the novel is most notable at the level of the discipline and resonates with Borgman (2007) who noted the importance of constancy in some disciplines in the broader context of scholarly publishing. The phase 1 and 2 survey findings clearly indicate that this fragmented pattern of co-evolution holds for the uptake and use of OAR.

Open Access Repositories are perceived by researchers as complementary to, rather than replacing, current forums for disseminating and publishing research.

The role of stage-two manuscript OAR is likely to vary according to differentiating factors such as those highlighted above and discussed in the report. To a greater or lesser extent OAR are perceived by researchers as complementary to current systems for publishing peer-reviewed journal articles. If the use of OAR by authors and readers is dependent on this complementary role, then the collaboration of all stakeholders in the scholarly communication system will be required. Publishers in particular have the potential to play a key role in contributing to the reputation and prestige of OAR.

Bibliography

Björk, B., Welling, P., Laakso, M., Majlender, P., Hedlund, T., Guõnason, G. (2010) Open access to the scientific journal literature: situation 2009. *PLOS One*, 5(6).

Blake, C. and Pratt, W. (2002) Collaborative information synthesis. In Toms, E.G. (Ed.) *ASIST 2002: Proceedings of the 65th ASIST Annual Meeting*. Medford, NJ: Information Today.

Bianco, S. *et al*. (2007) *Towards Open Access Publishing in High Energy Physics Report of the SCOAP3 Working Party*. Available at: http://scoap3.org/files/Scoap3WPReport.pdf [accessed 13.07.11]

Borgman, C. L. (2007) *Scholarship in the digital age: information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.

Creaser, C. (2010) "Open Access to Research Outputs—Institutional Policies and Researchers' Views: Results from Two Complementary Surveys." *New Review of Academic Librarianship* 16(1), pp.4–25. Available at: http://www.tandfonline.com/doi/abs/10.1080/13614530903162854 [accessed 29.09.11]

Dallmeier-Tiessen, S., Darby, R., Goerner, B., Hyppoelae, J., Igo-Kemenes, P., Kahn D., Lambert, S., Lengenfelder, A., Leonard, C., Mele, S., Polydoratou, P., Ross, D., Ruiz-Perez, S., Schimmer, R., Swaisland, M., van der Stelt, W. (2010) *First results of the SOAP project. Open access publishing in 2010. 2010* [arXiv pre-print No arXiv:1010.0506v1– submitted to PLOS One]. Available at: http://arxiv.org/ftp/arxiv/papers/1010/1010.0506.pdf [accessed 07.07.11].

Duranceau, E.F., 2008. The "Wealth of Networks" and institutional repositories: MIT, DSpace, and the future of the scholarly commons. *Library Trends*, 57(2), pp.244-261.

EUA (European Universities Association) (2008) *Recommendation from the EUA Working Group on Open Access adopted by the EUA council on 26 March 2008*.

Fry, J. (2006) Scholarly Research and Information Practices: A Domain Analytic Approach. *Information Processing and Management*, 42, pp. 299-316.

Gentil-Beccot, A., Mele, S, Holtkamp, A., O'Connell, H.O. and Brooks, T. (2009) Information resources in high-energy physics. Surveying the present landscape and charting the future course. *Journal of the American Society for Information Science and Technology*, 60(1), pp.150-160.

Goodman, D., Dowson, S. and Yaremchuk, J. (2007) Open access and accuracy: author-archived manuscripts vs. published articles. *Learned publishing*, 20(3), pp. 203-215.

Harnad, S., Carr, L., Swan, A., Sale, A. and Bosc, H. (2009) Maximising and measuring research impact through university and research-funder open-access self-archiving mandates. *Wissenschaftsmanagement*, 15(4), pp.36-41.

King, D., Tenopir, C., Choemprayong, S. and Wu, L. (2009) Scholarly journal information-seeking and reading patterns of faculty at five US universities. *Learned Publishing*, 22(2), pp.126-144.

Organ, M. (2006) Download statistics – What do they tell us*? D-Lin Magazine*, 12(11). Available at: www.dlib.org/dlib/november06/organ/11organ.html [accessed 13.07.11]

Palmer, C.L. (2005) Scholarly work and the shaping of the digital access. *Journal of the American Society for Information Science and Technology*, 58(11), pp.1140-1153.

RIN (2009) *Communicating knowledge: How and why UK researchers publish and disseminate their findings*. A Research Information Network report, September 2009.

Rowlands, I., Nicholas, D. and Huntingdon, P. (2004) Scholarly communication in the digital environment: what do authors want?. London: CIBER. Available at: http://www.ucl.ac.uk/ciber/ciber-pa-report.pdf [accessed 29.09.09].

Swan, A. (2007) Open access and the progress of science. *American Scientist*, 95, pp.198-200.

Swan, A. and Brown, S. (2003) Authors and electronic publishing: what authors want from the new technology. *Learned Publishing*, 16(1), pp.28-33.

Swan, A. and Brown, S. (2004) Authors and open access publishing. *Learned Publishing*, 17(3), pp.219-224.

Swan, A. and Brown, S. (2005) Open access self-archiving: an author study. Truro, UK: Key Perspectives Limited. Available at: www.jisc.ac.uk/media/documents/themes/infoenvironment/oaselfarchivingauthorstudy.pdf [accessed 15.09.09].

Tenopir, C., King, D., Edwards, S. and Wu, L. (2009) Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proceedings New Information*, 61(1), pp.5-32.

Thatcher, S (2011). Copyediting's role in an open-access world. *Against the grain*, 23(2), pp. 30-34.

Thorn, S., Morris, S. and Fraser, R. (2009) Learned societies and open access: key results from surveys of bioscience societies and researchers. *Serials*, 22(1), pp.39-48.

Wates, E. and Campbell, R. (2007) Author's version vs. publisher's version: an analysis of the copy-editing function. *Learned publishing*, 20(2), pp.121-129.

Watson, S. (2007) Authors' attitudes to, and awareness and use of, a university institutional repository. *Serials*, 20(3), pp.225-230.

Appendices

Appendix 1: Methodology 79
Phase 1 80
Phase 2 81
Survey of authors 81
Demographics of the phase 2 survey sample 83
Repository exit survey 85
Demographics of the phase 2 repository exit sample 86
Participatory Workshop 86
Participants 86
Programme 87
Appendix 2: Representativeness of the survey samples 92
Appendix 3: Summary of Phase 2 survey responses 92

Appendix 1: Methodology

Disciplinary groups

One of the main aims of this study was to investigate the effects of subject discipline on researchers' behaviours vis-à-vis self-archiving and Open Access Repositories (OAR). Four broad disciplinary groups were defined, detailed in Table 1.

Table 1 Grouping of disciplines

Broad disciplinary grouping	Disciplines included:
Medical Sciences	Clinical medicine
	Clinical dentistry
	Anatomy & physiology
	Nursing & paramedical studies
	Health & community studies
	Pharmacy & pharmacology
Life Sciences	Biosciences
	Psychology & behavioural sciences
	Earth, marine & environmental sciences
	Veterinary science
	Agriculture & forestry
Physical Sciences & Mathematics	Chemistry
	Physics
	Mathematics
	General engineering
	Chemical engineering
	Mineral, metallurgy & materials engineering
	Civil engineering
	Electrical, electronic & computer engineering
	Mechanical, aero & production engineering
	Information technology & systems sciences & computer software engineering

Broad disciplinary grouping	Disciplines included:
Social sciences, humanities & arts	Architecture, built environment & planning
	Catering & hospitality management
	Business & management studies
	Economics
	Geography Social studies
	Media studies
	Humanities & language based studies
	History
	Archaeology
	Modern languages
	Design & creative arts
	Education & Sports

The four broad disciplinary groupings were selected based on the categorisation of journals included in the PEER Observatory. Disciplines were allocated to these broad groups based on the project team's experience, and data from Thompson ISI.

In the surveys for both Phase 1 and Phase 2 of the study, respondents were asked to indicate in which discipline(s) they carried out research, and were allocated to a broad disciplinary group accordingly. These broad disciplinary groups were used for all data collection in Phase 1, and for the survey analysis in Phase 2 (see below). For the phase 2 workshops, a more-fine-grained approach was desired, and participants were selected, as far as possible, from the individual disciplines within the broad groupings.

Phase 1

The first phase of this study comprised two primary data gathering elements, in addition to an extensive review of relevant academic and grey literature. A survey of European researchers gathered evidence on a broad basis covering many of the issues in the research questions. This was supported by a series of focus groups that ran concurrently and explored specific issues in greater depth. The detailed methodology for phase 1, together with copies of the instruments used and a summary of the demographics of survey respondents, is provided in the baseline report, available from the PEER website[13].

In brief, the survey was conducted between June and August 2009, and attracted 3,139 valid responses. The research team is grateful to the PEER participating publishers for distributing the survey on our behalf; the conduct of the survey was similar to that for phase 2 which is described in detail below.

Focus groups were held in the four broad disciplinary groupings, in different cities across Europe. Each was facilitated by two members of the research team, and comprised between three and seven participants. Participation was encouraged by the offer of a €20 Amazon voucher, and travel expenses were paid.

[13] http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf [accessed 29.09.11]

Phase 2

Survey of authors

An electronic survey of authors was conducted between January and March 2011. Comments and advice from the PEER Research Oversight Group and Special Advisor were sought and the questionnaire amended following the consultation process. The survey was then piloted in early January 2011, to a small number of European researchers, known to the research team, working across a wide range of disciplines, including several non-native English speakers. The objectives of the pilot were to ascertain that the questions were meaningful to researchers, to assess the smooth running and logical sequence of the questions, to ensure that the language used was plain and jargon-free, and to make sure that any unavoidable 'technical' terms were properly defined. As the survey was going out to European researchers it was extremely important to have it piloted by non-native English speakers. Pilot participants were specifically asked to comment on the language accessibility of the survey. Useful comments were received and incorporated into the final survey.

Invitations to complete the main author survey were distributed via the twelve publishers participating in the PEER Observatory. The intention was to restrict the distribution to EU-based corresponding authors, who had manuscripts published in those journals included in the PEER Observatory and the control group, since the start of the PEER experiment in 2008. This time scale was set to reduce the potential number of currently inactive researchers who might be approached. Such a restriction of the circulation to EU authors was not always possible, and the survey was distributed more widely by several publishers. Non-EU respondents were filtered out from the responses in the analysis. One publisher was not able to distribute the survey on our behalf; in this case the research team trawled the tables of contents of participating journals to extract the email addresses of EU-based corresponding authors. An incentive prize draw for Amazon vouchers was offered to encourage responses, and one respondent was selected at random from the completed responses received by the closing date.

Originally it was intended that the survey would be live for six weeks in total, which was the case for most potential respondents; however, due to some unforeseen delays in the sending out of emails to Nature authors, this was extended for a week. The questionnaire was designed to cover those issues addressed by the research questions relevant to this phase of the study, in so far as a broad survey instrument was appropriate, and was informed by both the background research and the outcomes of Phase 1 of this study. Whereas the survey in the first phase included questions addressed to scholars both as writers of research outputs and as readers or consumers of those outputs, and focused on research dissemination via journal articles; the phase 2 survey was more focused on researchers' use of OAR, both in their capacity as authors and readers.

For this phase of the study, the authors that were invited to participate in the survey were separated into three distinct journal groups, based on the PEER Observatory author categories, with each group being directed to a different version of the survey. The three groups were:

- Publisher deposit group: authors of articles published in those journals for which the publisher deposits the stage-two manuscript in a PEER repository
- Author deposit group: authors of articles published in those journals for which the publisher invites the author to deposit their stage-two manuscript in a PEER repository themselves

- Control group: authors of articles published in the control journals, where the publisher does not take any action regarding open access

All three surveys contained the same questions; however splitting the respondents in this way provided additional information in the context of PEER, and insight into the extent to which contact from publishers has influenced self-archiving practice. It is appreciated that this is not an exact science – prolific authors may have published in journals from all three PEER categories (publisher deposit, author invited to deposit, control) – and as such any conclusions drawn about the behaviour of the different groups have been based on high levels of statistical significance ($p<0.01$).

Survey responses were checked for completeness, coded, and quantitative responses analysed using the SPSS software package. All respondents answering more than just the first section of the questionnaire were included in the analysis. A total of 1,427 valid responses from researchers in the EU were received, although not all respondents answered every question.

As well as overall summaries of the responses, all questions were analysed by broad disciplinary group, and by the journal groups (author deposit, publisher deposit, control group, as defined above). Where sufficient data were available, differences between groups were identified using a χ^2 test. Results are noted as being statistically significant when the probability that the observed differences between the groups occurred by pure chance, and that in the wider population no such differences exist, is less than 5% ($p<0.05$) or 1% ($p<0.01$).

Demographics of the sample

Responses were received from all countries in the EU, except for Latvia, with more than 100 responses each from five countries (France, Germany, Italy, Spain, and UK) *(Figure 1).*

Figure 1 Survey response by country



Almost two-thirds of respondents are from a university or college, with 19% stating they are from a research institute and 10% reporting they are from a hospital or medical school *(Figure 2).* There was a broad range in the length of time for which respondents have been conducting research, although just 3% report being active in research for fewer than three years. In contrast to this, 23% report they have been conducting research for 15-24 years, with 22% stating they have been active in research for 25 years or longer *(Figure 3).*

Respondents were asked to indicate the areas of their research by ticking one or more of 39 distinct disciplines, grouped into four broad areas. Almost one-third of respondents were from the Physical sciences & mathematics *(Figure 4).* Respondents were allocated to an Interdisciplinary group if they ticked disciplines from two or more of the broad areas included; the majority of these had indicated disciplines in both the Medical sciences and the Life sciences.

Respondents were also asked to indicate how many articles they had had published in the last five years. Perhaps unsurprisingly given the nature of the survey, only two respondents indicated that they have not published in the last five years, illustrated by Figure 5. Almost 29% of respondents indicated that they have published more than 20 articles in the last five years.

.

Figure 2 Institution type

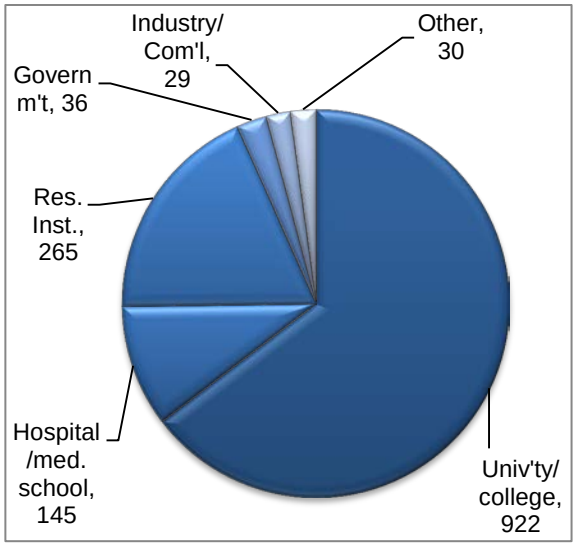


Figure 3 Research experience

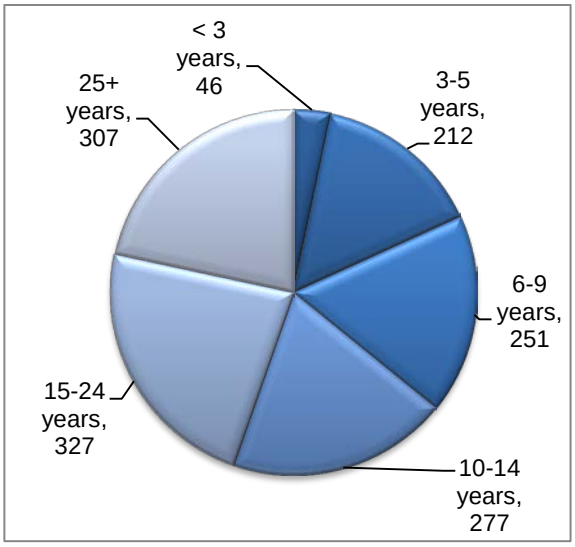


Figure 4 Subject area

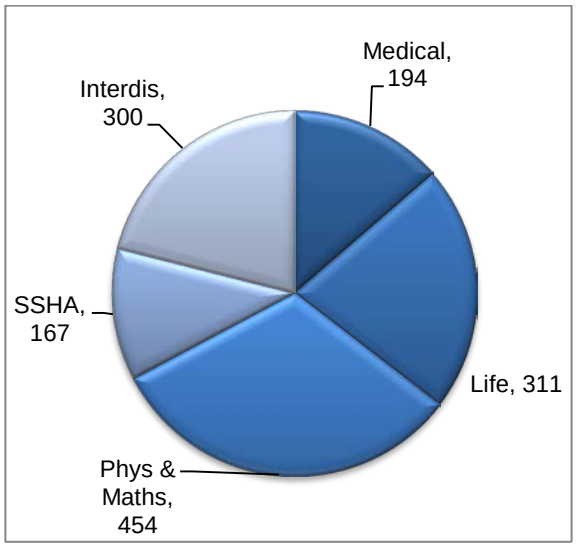
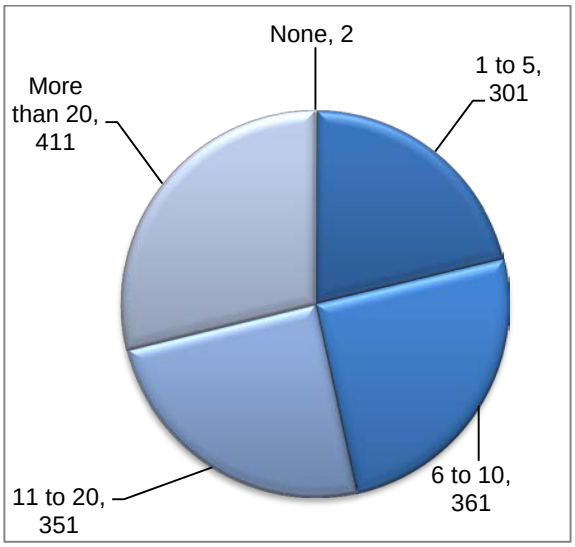


Figure 5 Articles published



Overall, just 8% of respondents indicated that they have been contacted by a publisher in connection with the PEER Observatory and were invited to place a copy of one of their articles in an OAR, with almost 4% stating that they had been informed that a copy of one of their articles has been placed in an OAR, whilst 58% stated that they have not been contacted by a publisher regarding the PEER Observatory.

Repository exit survey

The aim of the repository exit survey was to gather information from users of the PEER repositories, regardless of whether or not they were also authors. Most questions were a subset of those in the main survey, so that results from the two surveys could be compared and combined for the common questions. Three different ways of implementing an exit survey were trialled with mock surveys linked from three different PEER repositories. The INRIA-HAL PEER repository[14] investigated the implementation of a redirect link to a web-based mock survey; the Max Planck Digital Library (MPDL)[15] looked into the implementation of a homepage mock survey whilst the Göttingen PEER repository[16] (UGOE) trialled the implementation of a pop-up exit survey. It was generally felt, both by the Loughborough Behavioural research project team and the repository teams, that the redirection link, trialled by HAL, was the most efficient and easiest way to encourage repository users to take part in the survey. A visually appealing and distinctive button (mentioning the €50 Amazon voucher prize draw) was created and included on all participating repositories, both at the bibliographic metadata record level and on the repository homepage. While the implementation tests were taking place, the draft exit survey was also circulated to the PEER Research Oversight Group and Special Advisor for comments and advice. Feedback was incorporated into the final exit survey.

The six participating PEER repositories where the exit survey was implemented represent a mix of central, institutional and subject-based repositories, and are listed in Table 2. As with the main survey, the exit survey was live for a total of six weeks between January and March 2011. Only one repository failed to implement the link to the survey within the time allocated, and therefore the survey for this specific repository was open for less than one month with some minor problems (notably, the visibility of the link to the survey). Overall, the response rate was very disappointing with just 34 valid responses received from EU respondents. Because of this low response, it has not been possible to undertake any detailed analysis of this survey.

Table 2 Participating repositories where the exit survey was implemented

- eSciDoc.PubMan.PEER, Max Planck Digital Library (MPDL), Max-Planck-Gesellschaft zur Forderung der Wissenschaften e.V. (MPG), Germany
- HAL, CNRS & Institut National de Recherche en Informatique et en Automatique (INRIA), France
- Göttingen State and University Library (UGOE), Germany
- University Library of Debrecen, Hungary
- SSOAR - Social Sciences Open Access repository (GESIS - Leibniz Institute for the Social Sciences, Germany)
- TARA, Trinity College Dublin, Ireland

[14] http://peer.ccsd.cnrs.fr/index.php?halsid=7ngr0ohfe0ie5unjafc7jni803&view_this_doc=peer-00513032&version=1 [accessed 29.09.11]
[15] http://peer.mpdl.mpg.de/pubman/ [accessed 29.09.11]
[16] http://134.76.163.171:8080/jspui/ [accessed 29.09.11]

Demographics of the sample

Overall, 26 respondents to the exit survey reported that they were based in a university or college, with responses received from seven of the 27 EU countries only. A total of twelve respondents indicated that they were undergraduate or masters students, with fourteen respondents indicating that they had been involved in research for fewer than three years. Overall, 21 respondents indicated that they were from the Social sciences, humanities & arts, with 16 indicating that they were between 26 and 35 years of age.

Participatory Workshop

This one-day workshop was held at University College, London, in April 2011.

Participants

Participants were recruited from across Europe, from respondents to the phase 1 and phase 2 surveys who had indicated a willingness to participate further in the research. Invitations to participate in the workshop were initially sent to over 700 researchers who had expressed an interest in being contacted in the future in relation to this research, in their survey responses.

The aim for this final stage of the research was to drill down from the initial four broad disciplinary groupings to individual disciplines. Out of necessity recruitment was based both on selective sampling and practical considerations. Potential participants were recruited on the basis of the disciplines they had indicated in their survey response, and the final disciplines were: Earth, marine & environmental sciences (Life sciences); Chemistry (Physical sciences & mathematics); and Engineering (Physical sciences & mathematics). Within the Social sciences, humanities & arts, there were insufficient potential participants from any single discipline, so a broader approach was taken to ensure this broad disciplinary group was represented. There were insufficient volunteers from disciplines in the Medical sciences to include this broad disciplinary group in the workshop. Conversely, there was notable interest amongst potential participants from physics, but it had been decided by prior agreement with the PEER Executive, Research Oversight Group and Special Advisor, that since researchers from physics had been the object of numerous related studies it would be better to obtain a detailed picture from less-well understood disciplines. Details are given in Table 3.

Table 3 Participatory workshop sampling

Broad disciplinary group	Discipline	Institution types	No. of participants
Life sciences	Earth, marine & environmental sciences	University or College x5 Charity x1	6
Physical sciences & mathematics	Chemistry	University or College x2 Research institute x3	5
	Engineering	University or College x4 Research institute x2	6
Social sciences, humanities & arts	Mixed	University or College x3 Research institute x1 Private research institute x1	5
Total			22

No incentive was offered to participants, but reasonable travel and subsistence expenses were reimbursed. Table 4 summarises the demographics of participants, and shows that a wide range of institution types, countries and experience was represented overall.

Table 4 Demographics of workshop participants

Type of institution	No.
University or college	14
Research Institute	6
Charity	1
Private research institute	1
Total	22

Research experience	No.
fewer than 5 years	3
5-14 years	8
15-24 years	7
25 years or longer	4
Total	22

Country of origin	No.
Austria	1
Bulgaria	1
Czech Republic	1
France	4
Germany	3
Greece	1
Italy	2
Netherlands	2
Portugal	1
Spain	4
United Kingdom	2
Total	22

Programme

The workshop was facilitated by four members of the Behavioural research project team. The day was divided into three sessions, and included a short presentation by CIBER on the Usage research project, and a light lunch. Participants were provided with an information pack, including a glossary of open access-related terms likely to come up during the workshop.

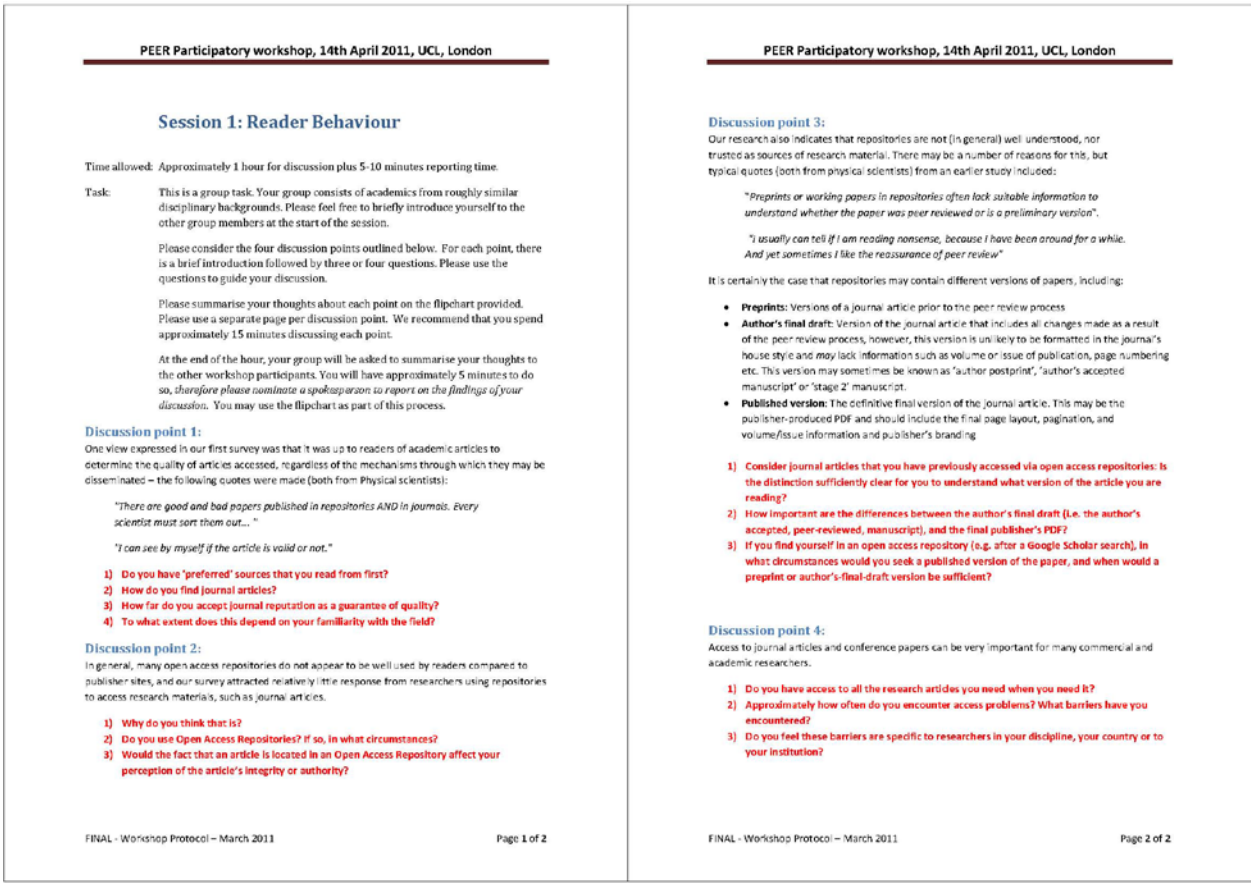
The approach chosen for the participatory workshop was a mix of general discussion points and scenarios to guide participants through the discussion. For sessions one and two, participants were divided into four groups, according to discipline. Each group was observed by a member of the project team. Following discussion in their disciplinary groups, a spokesperson from each was asked to report their results for wider discussion. The project team facilitated these cross-disciplinary reporting and discussion sessions. The third and final session was cross-disciplinary with all participants making a collective contribution to the discussion, which was facilitated by the project team.

Session 1: Reader behaviours

The session started with a presentation about finding and accessing material in OAR. As participants were thought to have different levels of knowledge of, and familiarity with, OAR we provided a demonstration of how material in an OAR might be located and accessed; illustrating a Google/Google Scholar search for a pre-defined article and how to access it from one of the PEER repositories. An article from a 'neutral' discipline was used as an example, i.e. a discipline that was not represented in the workshop and where the repository landscape was diverse. This allowed the demonstration to run without conveying prior assumptions about participants' use (or non-use) of OAR.

Each of the four discipline-based groups were given a copy of the discussion points *(Figure 6),* with 45 minutes allowed for discussion within the groups, which then re-convened for reporting back. All sessions were audio-recorded to aid analysis

Figure 6 Group hand-out for session 1



PEER Participatory workshop, 14th April 2011, UCL, London

Session 1: Reader Behaviour

Time allowed: Approximately 1 hour for discussion plus 5-10 minutes reporting time.

Task: This is a group task. Your group consists of academics from roughly similar disciplinary backgrounds. Please feel free to briefly introduce yourself to the other group members at the start of the session.

Please consider the four discussion points outlined below. For each point, there is a brief introduction followed by three or four questions. Please use the questions to guide your discussion.

Please summarise your thoughts about each point on the flipchart provided. Please use a separate page per discussion point. We recommend that you spend approximately 15 minutes discussing each point.

At the end of the hour, your group will be asked to summarise your thoughts to the other workshop participants. You will have approximately 5 minutes to do so, *therefore please nominate a spokesperson to report on the findings of your discussion.* You may use the flipchart as part of this process.

Discussion point 1:

One view expressed in our first survey was that it was up to readers of academic articles to determine the quality of articles accessed, regardless of the mechanisms through which they may be disseminated – the following quotes were made (both from Physical scientists):

"There are good and bad papers published in repositories AND in journals. Every scientist must sort them out... "

"I can see by myself if the article is valid or not."

1) Do you have 'preferred' sources that you read from first?
2) How do you find journal articles?
3) How far do you accept journal reputation as a guarantee of quality?
4) To what extent does this depend on your familiarity with the field?

Discussion point 2:

In general, many open access repositories do not appear to be well used by readers compared to publisher sites, and our survey attracted relatively little response from researchers using repositories to access research materials, such as journal articles.

1) Why do you think that is?
2) Do you use Open Access Repositories? If so, in what circumstances?
3) Would the fact that an article is located in an Open Access Repository affect your perception of the article's integrity or authority?

FINAL - Workshop Protocol – March 2011 Page 1 of 2

PEER Participatory workshop, 14th April 2011, UCL, London

Discussion point 3:

Our research also indicates that repositories are not (in general) well understood, nor trusted as sources of research material. There may be a number of reasons for this, but typical quotes (both from physical scientists) from an earlier study included:

"Preprints or working papers in repositories often lack suitable information to understand whether the paper was peer reviewed or is a preliminary version".

"I usually can tell if I am reading nonsense, because I have been around for a while. And yet sometimes I like the reassurance of peer review"

It is certainly the case that repositories may contain different versions of papers, including:

- **Preprints:** Versions of a journal article prior to the peer review process
- **Author's final draft:** Version of the journal article that includes all changes made as a result of the peer review process, however, this version is unlikely to be formatted in the journal's house style and *may* lack information such as volume or issue of publication, page numbering etc. This version may sometimes be known as 'author postprint', 'author's accepted manuscript' or 'stage 2' manuscript.
- **Published version:** The definitive final version of the journal article. This may be the publisher-produced PDF and should include the final page layout, pagination, and volume/issue information and publisher's branding

1) Consider journal articles that you have previously accessed via open access repositories: Is the distinction sufficiently clear for you to understand what version of the article you are reading?
2) How important are the differences between the author's final draft (i.e. the author's accepted, peer-reviewed, manuscript), and the final publisher's PDF?
3) If you find yourself in an open access repository (e.g. after a Google Scholar search), in what circumstances would you seek a published version of the paper, and when would a preprint or author's-final-draft version be sufficient?

Discussion point 4:

Access to journal articles and conference papers can be very important for many commercial and academic researchers.

1) Do you have access to all the research articles you need when you need it?
2) Approximately how often do you encounter access problems? What barriers have you encountered?
3) Do you feel these barriers are specific to researchers in your discipline, your country or to your institution?

FINAL - Workshop Protocol – March 2011 Page 2 of 2

Session 2: Author behaviours

This session was introduced with a brief presentation outlining some of the issues authors may need to take into consideration when choosing how to disseminate their research, and to which journal they might submit their work. The issues were based on results from the surveys of authors in both phases of the research, and were presented in the form of a diagram *(Figure 7),* using the case of a publication in information science as an illustration. The introduction also covered the concepts of funder and institutional mandates for OA, which are not uniform across Europe, and may have been more familiar to some participants than others.

In their groups, participants were asked to represent their own dissemination workflows and influences on a copy of the diagram shown in Figure 7, followed by discussion of how various scenarios for publication timings and open access mandates might influence this *(Figure 8).* The groups then re-convened for a general discussion.

Figure 7 Influences on research dissemination decisions

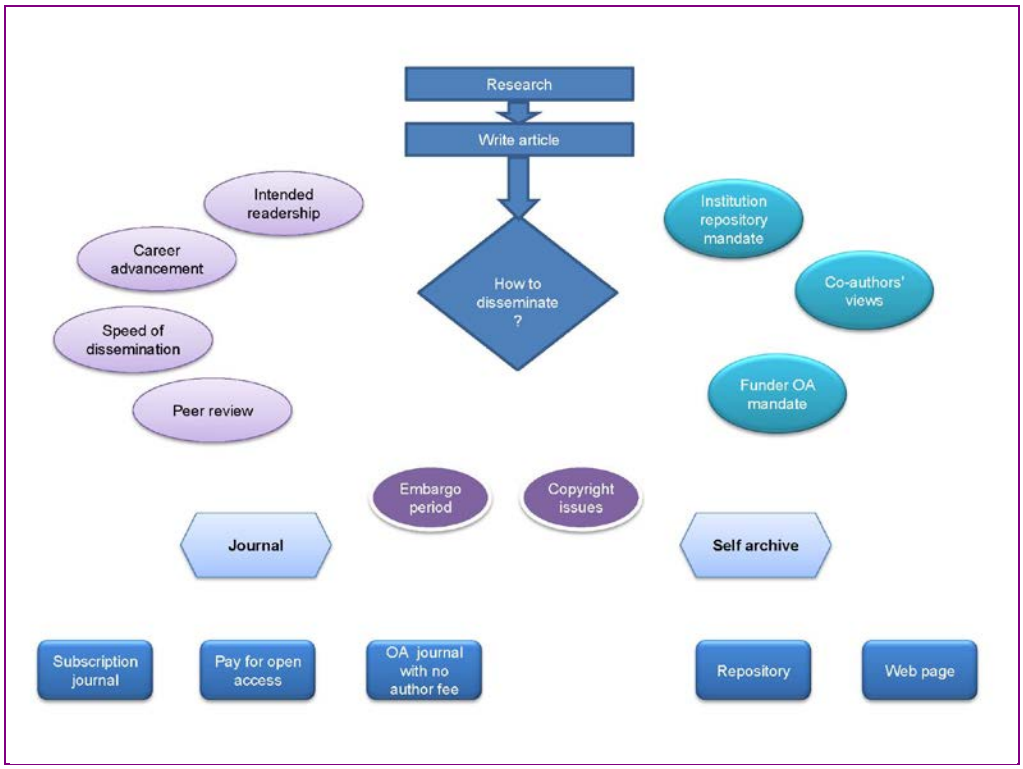
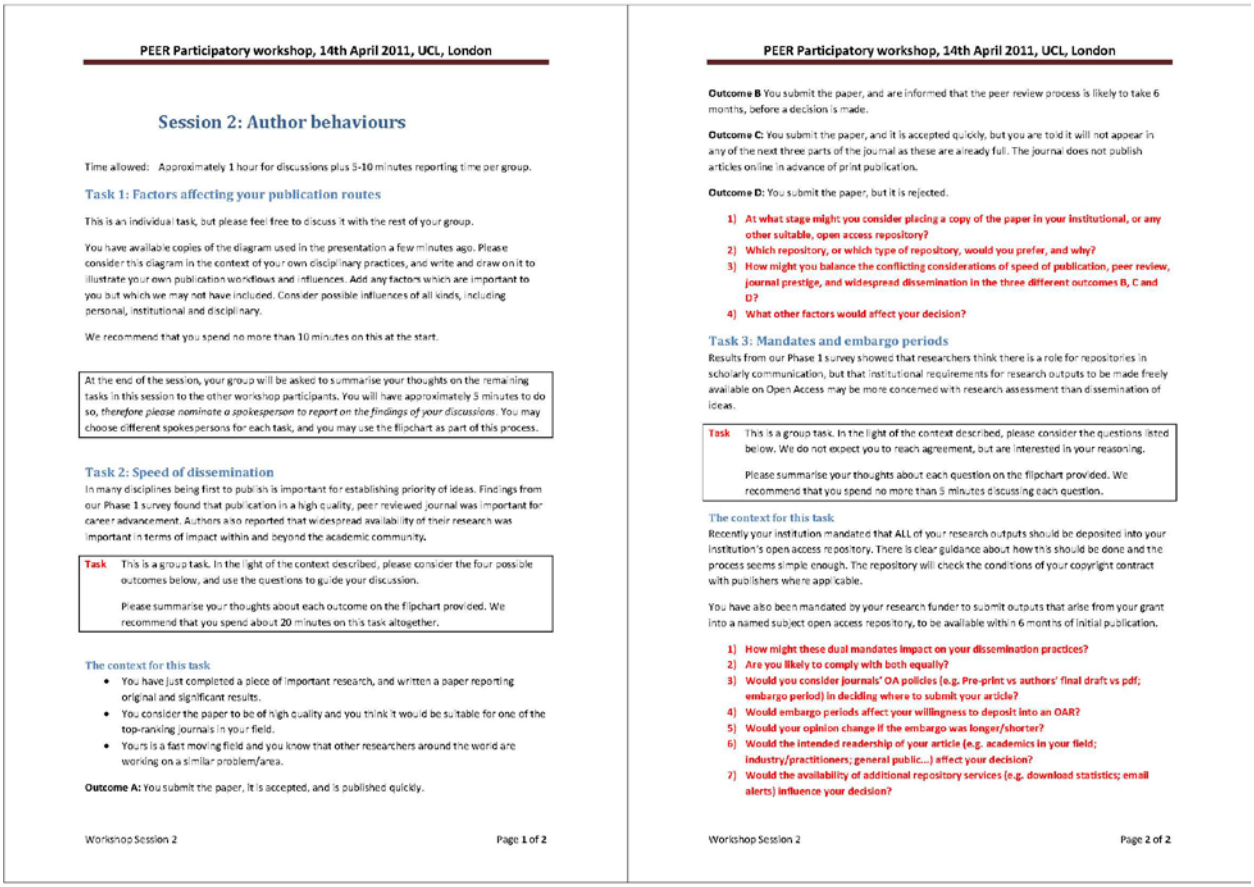


Figure 8 Group hand-out for session 2



PEER Participatory workshop, 14th April 2011, UCL, London

Session 2: Author behaviours

Time allowed: Approximately 1 hour for discussions plus 5-10 minutes reporting time per group.

Task 1: Factors affecting your publication routes

This is an individual task, but please feel free to discuss it with the rest of your group.

You have available copies of the diagram used in the presentation a few minutes ago. Please consider this diagram in the context of your own disciplinary practices, and write and draw on it to illustrate your own publication workflows and influences. Add any factors which are important to you but which we may not have included. Consider possible influences of all kinds, including personal, institutional and disciplinary.

We recommend that you spend no more than 10 minutes on this at the start.

At the end of the session, your group will be asked to summarise your thoughts on the remaining tasks in this session to the other workshop participants. You will have approximately 5 minutes to do so, *therefore please nominate a spokesperson to report on the findings of your discussions*. You may choose different spokespersons for each task, and you may use the flipchart as part of this process.

Task 2: Speed of dissemination

In many disciplines being first to publish is important for establishing priority of ideas. Findings from our Phase 1 survey found that publication in a high quality, peer reviewed journal was important for career advancement. Authors also reported that widespread availability of their research was important in terms of impact within and beyond the academic community.

Task This is a group task. In the light of the context described, please consider the four possible outcomes below, and use the questions to guide your discussion.

Please summarise your thoughts about each outcome on the flipchart provided. We recommend that you spend about 20 minutes on this task altogether.

The context for this task

- You have just completed a piece of important research, and written a paper reporting original and significant results.
- You consider the paper to be of high quality and you think it would be suitable for one of the top-ranking journals in your field.
- Yours is a fast moving field and you know that other researchers around the world are working on a similar problem/area.

Outcome A: You submit the paper, it is accepted, and is published quickly.

PEER Participatory workshop, 14th April 2011, UCL, London

Outcome B You submit the paper, and are informed that the peer review process is likely to take 6 months, before a decision is made.

Outcome C: You submit the paper, and it is accepted quickly, but you are told it will not appear in any of the next three parts of the journal as these are already full. The journal does not publish articles online in advance of print publication.

Outcome D: You submit the paper, but it is rejected.

1) **At what stage might you consider placing a copy of the paper in your institutional, or any other suitable, open access repository?**
2) **Which repository, or which type of repository, would you prefer, and why?**
3) **How might you balance the conflicting considerations of speed of publication, peer review, journal prestige, and widespread dissemination in the three different outcomes B, C and D?**
4) **What other factors would affect your decision?**

Task 3: Mandates and embargo periods

Results from our Phase 1 survey showed that researchers think there is a role for repositories in scholarly communication, but that institutional requirements for research outputs to be made freely available on Open Access may be more concerned with research assessment than dissemination of ideas.

Task This is a group task. In the light of the context described, please consider the questions listed below. We do not expect you to reach agreement, but are interested in your reasoning.

Please summarise your thoughts about each question on the flipchart provided. We recommend that you spend no more than 5 minutes discussing each question.

The context for this task

Recently your institution mandated that ALL of your research outputs should be deposited into your institution's open access repository. There is clear guidance about how this should be done and the process seems simple enough. The repository will check the conditions of your copyright contract with publishers where applicable.

You have also been mandated by your research funder to submit outputs that arise from your grant into a named subject open access repository, to be available within 6 months of initial publication.

1) **How might these dual mandates impact on your dissemination practices?**
2) **Are you likely to comply with both equally?**
3) **Would you consider journals' OA policies (e.g. Pre-print vs authors' final draft vs pdf; embargo period) in deciding where to submit your article?**
4) **Would embargo periods affect your willingness to deposit into an OAR?**
5) **Would your opinion change if the embargo was longer/shorter?**
6) **Would the intended readership of your article (e.g. academics in your field; industry/practitioners; general public…) affect your decision?**
7) **Would the availability of additional repository services (e.g. download statistics; email alerts) influence your decision?**

Session 3: Communication and dissemination practices

This final session was conducted with all participants together, to allow for greater cross-disciplinary discussion, and to tease out common themes from the earlier sessions. Each discussion point was introduced by a member of the project team, and open discussion was encouraged. Again, the session was audio-recorded to facilitate analysis.

Discussion point 1: Routes to Open Access

Some publishers allow authors to deposit in open access repositories, but not until a set period of time has elapsed (e.g. 6 months or 12 months after first publication is one stipulation).

1. What are your views on how this affects the usefulness of OA repositories?

Some publishers can make journal articles open access (i.e. available to everyone) as long as the author (or their institution) pays a fee. The fee (which for certain publishers is in the region of Euro 2000) may be paid by your institution/funder or covered by your research grant.

2. Is this a better way of achieving open access to journal articles compared to publishing in conventional journals and placing a copy in a repository?

3. What are the pros and cons of these two different approaches, both in terms of you as an author and also potential readers of your articles?

Discussion point 2: Usage and citations

Findings from our Phase 1 survey showed that more than two-thirds of respondents try to find the published final version of an article in order to cite this version.

1. As an author, how important is it to you which version of your paper is cited?

2. How important is the version you cite in your own work? Is citing a version other than the final publishers' version acceptable?

3. Do you consider the availability of the source material to others when researching with a view to writing?

4. If it was reported that benefits of Open Access may include:
    - more people download and view your research
    - more citations

    how important would this be to you and why?

    What increase in citations would make you more likely to make your work open access?

Discussion point 3: Copyright considerations

Our research has shown that over 40% of respondents are worried about infringing copyright when making their work open access. Services such as Sherpa-Romeo enable authors to identify the open access policy of specific journals (e.g. whether deposit in institutional repositories is allowed and any associated conditions and embargo periods).

1) How concerned would you be about infringing copyright?
2) Would you be prepared to use services such as Sherpa-Romeo?
3) Is this something that repository managers (or even publishers) should police?

Discussion point 4: Peer Review

Over 50% of respondents to our Phase 1 survey thought that open access repositories may challenge the existence of subscription-based journals and over 30% felt that open access may challenge the integrity of the peer review systems.

1) Are the existing models of publishing and peer review important to you?
2) Is change to be avoided or accepted (and at what level)?
3) What is important about traditional peer review?
4) Are other possibilities, such as peer commentary, viable alternatives?

Discussion point 5: Uptake of repositories

Open Access via repositories has been on the agenda for many years, yet the number of articles deposited has yet to meet the levels that some expect (in most disciplines), and this in turn limits the usefulness of the repositories.

1) What do you think OA repositories have to offer in terms of dissemination and access to articles?
2) What influences or factors would raise your adoption of self-archiving practices, and use of repositories?
3) Would features such as download statistics, citation links, peer commentary, and sharing functions be an enticement to deposit your research outputs into an OA repository?

Discussion point 6: The PEER experiment

Round up about the PEER experiment.

1) The large-scale deposit of author's final draft into open access repositories by publishers is a novel approach to OA publishing. What influence, if any, do you feel this approach will have on your current dissemination and information seeking practices?
2) If OAR were appropriated by a sufficient critical mass of researchers so as to become a central resource in your discipline, how do you think this might change the scholarly communication process?

Appendix 2: Representativeness of the survey samples

The number of individuals engaged in research and scholarly communication activities at any one time is not aggregated at a national or international level, although partial data are available for some sectors in some countries. It is, therefore, not possible to ascertain whether the sample of researchers included in our study, in either phase, is broadly representative of the wider population of European researchers.

Demographic data were collected from respondents covering their type of institution, the country in which this was based, their research discipline, and the length of time for which they had been engaged in research.

Data were also obtained from the participating publishers on the number of invitations to complete the survey that were sent out in each phase, for each journal title included in the PEER Observatory and the control group. Table 5 summarises these Figures, according to the four broad disciplinary groups covered. Table 6 gives the distribution of survey responses in each phase; there was a considerable number of respondents allocated to an interdisciplinary category in each case, where they had ticked subject areas which fell into two (or more) of the broader categories.

Table 5 Survey distribution, by discipline

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

Table 6 Survey responses, by discipline

	Phase 1		Phase 2	
Medical sciences	248	*8%*	194	*14%*
Life sciences	416	*13%*	311	*22%*
Physical sciences & mathematics	1,773	*57%*	454	*32%*
Social sciences, humanities & arts	259	*8%*	167	*12%*
Interdisciplinary	440	*14%*	300	*21%*
Total	3,136	*100%*	1,426	*100*

It is clear that researchers from the Medical sciences were under-represented in the sample compared to the invitations distributed, and those from the Physical sciences & mathematics were over-represented, particularly in phase 1. Discipline is hypothesised to influence researcher behaviour vis-à-vis open access and repositories, and this was confirmed by the survey analyses. Potentially, therefore, any comparisons between groups defined by other coarse-grained characteristics may be confounded by disciplinary differences where the sample is not fully representative of the population.

In order to investigate this for the analyses presented to PEER in the baseline report and the initial results of the phase 2 survey, weights were calculated by discipline for each survey separately. This was done by counting the number of respondents with a research interest in each of the four broad disciplinary groups, and comparing this proportion of the total with the proportion invited to complete the survey. Weights for interdisciplinary researchers were calculated as the average of the relevant individual discipline weights. The calculations for the phase 2 survey are presented in Box 1 as an illustration.

Box 1 Calculation of weights for Phase 2

Survey responses

Medicine	194	Medicine & Physical	12
Life	311	Medicine, Physical & SSHA	2
Physical	454	Medicine & SSHA	18
SSHA	167	Life & Physical	62
Medicine & Life	126	Life & SSHA	17
Medicine, Life & Physical	19	Life, Physical & SSHA	5
Medicine, Life & SSHA	14	Physical & SSHA	19
Medicine, Life, Physical & SSHA	6	Total	1,426

By broad discipline

	Medicine	Life	Physical	SSHA	Total
Publisher distribution	10,145	9,040	10,567	4,998	34,750
Proportion of total	29%	26%	30%	14%	
Survey responses	391[a]	560	579	248	1,778
Proportion of total	22%	31%	33%	14%	
Ratio of percentages	**1.33**	**0.83**	**0.93**	**1.03**	

[a] e.g. 194 + 126 + 19 + 14 + 6 + 12 + 2 + 18 = 391

Weights

Medicine	1.33	Medicine & Physical	1.03
Life	0.83	Medicine, Physical & SSHA	1.10
Physical	0.93	Medicine & SSHA	1.18
SSHA	1.03	Life & Physical	0.88
Medicine & Life	1.08[a]	Life & SSHA	0.93
Medicine, Life & Physical	1.03	Life, Physical & SSHA	0.93
Medicine, Life & SSHA	1.06	Physical & SSHA	0.98
Medicine, Life, Physical & SSHA	1.03		

The weighted overall distributions of responses from the two surveys were compared with the non-weighted Figures originally presented. Whilst these were frequently statistically significantly different, owing to the large number of researchers involved, in practice the differences were generally small, and did not change the conclusions of the initial analyses.

Appendix 3: Summary of Phase 2 survey responses

Note that the Figures presented in this appendix are as received from respondents, i.e. they have not been weighted by discipline. They may therefore differ from the Figures presented in the main body of the report, which have been weighted by discipline.

Appendix 3.1 Main survey

Section 1: About you

1 Which of the following best describes your institution?

	Frequency	%
University or college	922	*64.6*
Hospital or medical school	145	*10.2*
Research institute	265	*18.6*
Government	36	*2.5*
Industrial/commercial	29	*2.0*
Other	30	*2.1*
Total	1,427	*100*

2 In which country is your institution based?

	Frequency	%
Austria	28	*2.0*
Belgium	46	*3.2*
Bulgaria	5	*0.4*
Cyprus	2	*0.1*
Czech Republic	22	*1.5*
Denmark	65	*4.6*
Estonia	4	*0.3*
Finland	36	*2.5*
France	119	*8.3*
Germany	187	*13.1*
Greece	34	*2.4*
Hungary	21	*1.5*
Ireland	14	*1.0*
Italy	167	*11.7*

	Frequency	%
Latvia	0	*0.0*
Lithuania	5	*0.4*
Luxembourg	3	*0.2*
Malta	1	*0.1*
Netherlands	92	*6.4*
Poland	39	*2.7*
Portugal	38	*2.7*
Romania	12	*0.8*
Slovakia	3	*0.2*
Slovenia	2	*0.1*
Spain	108	*7.6*
Sweden	65	*4.6*
UK	309	*21.7*
Total	1,427	*100*

3 For how long have you been conducting research?

	Frequency	%
Fewer than 3 years	46	*3.2*
3-5 years	212	*14.9*
6-9 years	251	*17.6*
10-14 years	277	*19.5*
15-24 years	327	*23.0*
25 years or longer	307	*21.6*
Not applicable	4	*0.3*
Total	1,424	*100*

4 Approximately how many articles have you published in the last five years?

	Frequency	%
0	2	*0.1*
1-5	301	*21.1*
6-10	361	*25.3*
11-20	351	*24.6*
More than 20	411	*28.8*
Total	1,426	*100*

5 Which field best describes your research area?

	Frequency	%
Medical sciences	194	*13.6*
Life sciences	311	*21.8*
Physical sciences & mathematics	454	*31.8*
Social sciences, humanities & arts	167	*11.7*
Interdisciplinary	300	*21.0*
Total	1,426	*100*

6 Have you been contacted by a publisher in connection with the PEER Observatory in the last two years?

	Frequency	%
Yes, I was invited to place a copy of one of my articles in an open access repository	116	*8.1*
Yes, I was informed a copy of one of my articles would be placed in an open access repository	50	*3.5*
No	831	*58.2*
Not sure	457	*32.0*
Total	1,427	

Respondents could tick more than one option.
Percentages = % of the 1,427 respondents that answered any part of the question

Section 2: Authors - placing material in Open Access Repositories

7 Has a version of any of your published journal articles been placed in an OAR in the last five years?

	Frequency	%
a. Yes-institutional-placed myself	198	*14.2*
a. Yes-institutional-someone did it for me	251	*18.0*
b. Yes-subject based-placed myself	185	*13.3*
b. Yes-subject based - someone did it for me	210	*15.1*
c. Yes-other-placed myself	17	*1.2*
c. Yes-other - someone did it for me	27	*1.9*
d. Yes-not sure where - placed myself	13	*0.9*
d. Yes-not sure where - someone did it for me	55	*4.0*
e. No - placed myself	565	*40.6*
e. No- someone did it for me	377	*27.1*
f. Not sure- placed myself	125	*9.0*
f. Not sure- someone placed for me	267	*19.2*

Respondents could tick more than one option.
Percentages = % of 1,392 respondents that answered any part of the question

8 If 'Other' in Q7c, please specify:

40 responses

9 Which version(s) were placed in an OAR?

	Frequency	%
Pre-print (electronic version prior to peer-review)	232	*29.4*
Author's final version (electronic version after peer-review, including amendments following referees' comments, also known as post-print)	288	*36.5*
Published final version (publisher's PDF file)	344	*43.5*
Not sure	146	*18.5*

Respondents could tick more than one option.
Percentages = % of the 790 respondents that answered any part of the question

10 What prompted you to place (or have placed) a copy of your article in an OAR?

	Frequency	%
You did so voluntarily	326	*46.2*
Colleague(s) suggested it	87	*12.3*
Co-author(s) asked you to	53	*7.5*
You were invited by a publisher to do so	115	*16.3*
You were invited by the repository	79	*11.2*
You were required to do so by your employer	140	*19.9*
You were required to do so by your research funder	62	*8.8*
You were invited by a librarian	51	*7.2*
Other	63	*8.9*

Respondents could tick more than one option.
Percentages = % of the 705 respondents that answered any part of the question

11 Which type of OAR did you choose?

	Frequency	%
Institutional	277	*36.3*
Subject-based	246	*32.2*
Both	90	*11.8*
Not sure	128	*16.8*
Other	22	*2.9*
Total	763	*100*

12 How easy did you find each of the following processes when placing the copy of your article in an OAR?

	Very easy	Easy	Not very easy	Difficult	Very difficult	Someone did it for me	Don't know	Total
Identifying a suitable OAR	213	146	61	18	5	197	91	731
	29.1%	*20.0%*	*8.3%*	*2.5%*	*0.7%*	*26.%9*	*12.4%*	*100%*
Checking that the publisher allows the article to be placed in an OAR	80	135	127	54	23	160	135	714
	11.2	*18.9%*	*17.8%*	*7.6%*	*3.2%*	*22.4%*	*18.9%*	*100%*
Finding the correct manuscript version	174	226	86	16	6	115	87	710
	24.5	*31.8%*	*12.1%*	*2.3%*	*0.8%*	*16.2%*	*12.3%*	*100%*
Process of uploading to the OAR	125	213	71	6	9	195	95	714
	17.5	*29.8%*	*9.9%*	*0.8%*	*1.3%*	*27.3%*	*13.3%*	*100%*
Overall, how did you find the whole experience?	110	256	101	26	2	126	95	716
	15.4	*35.8%*	*14.1%*	*3.6%*	*0.3%*	*17.6%*	*13.3%*	*100%*

13 Please use this space to add any comments you may have about your experience with regard to Q12:

103 responses

14 How much time did it take to make your work available in the OAR?

	Frequency	%
Less than 5 minutes	60	*8.1*
5-15 minutes	155	*21.0*
15 minutes – 1 hour	132	*17.9*
1-2 hours	41	*5.6*
More than 2 hours	28	*3.8*
Someone else did it for me	231	*31.3*
Don't know/can't remember	91	*12.3*
Total	738	*100*

If someone else did it for you, please indicate who

	Frequency	%
Administrative staff/secretary	36	*15.6*
Library/repository staff	83	*35.9*
Student/research assistants	3	*1.3*
Colleague/co-author	41	*17.7*
Not sure/can't remember	33	*14.3*
Other	35	*15.2*
Total	231	*100*

15 Do you think that the benefits you expect to receive from your article being in OA form are worth the extra effort?

	Frequency	%
Definitely, yes	296	*38.6*
Probably, yes	282	*36.8*
No I don't think so	51	*6.7*
Not sure	137	*17.9*
Total	766	*100*

16 On placing your article in an OAR what embargo period was specified by the publisher?

	Frequency	%
None	225	*29.2*
Up to 6 months	51	*6.6*
7-12 months	49	*6.4*
13-18 months	12	*1.6*
19-24 months	3	*0.4*
Longer than 24 months	6	*0.8*
I don't know/can't remember	425	*55.1*
Total	771	*100*

17 If a publisher specified that the embargo period until your article was made freely available would be different to your answer to question 16, would this affect your willingness to make your work open access?

	More likely to make my work OA		Less likely to make my work OA		No difference		Not sure		Total	
Shorter embargo period specified	133	*32%*	12	*3%*	152	*37%*	113	*28%*	410	*100%*
Longer embargo period specified	21	*5%*	114	*29%*	143	*36%*	118	*30%*	396	*100%*

18 If a publisher specified that the embargo period until your article was made freely available would be different to your answer to question 16, would this affect your choice of publisher and/or journal?

	Yes		No		Not sure		Total	
Shorter embargo period specified	90	*22%*	165	*40%*	160	*39%*	415	*100%*
Longer embargo period specified	86	*22%*	151	*38%*	163	*41%*	400	*100%*

Section 3: Authors - Open Access policies as they relate to you

19 Does your institution require that you make a version of your published journal articles publicly available through your institution's repository?

	Frequency	%
Yes	196	*13.9*
No	848	*60.0*
Not sure	205	*14.5*
My institution does not have a repository	165	*11.7*
Total	1,414	*100*

20 Have any funders of your research required that journal articles resulting from that research are made openly and freely accessible?

	Frequency	%
Yes	174	*12.6*
No	1,067	*77.1*
Not sure	143	*10.3*
Total	1,384	*100*

If Yes, is this through:

	Frequency	%
Any repository	37	*21.3*
An institutional repository	40	*23.0*
A subject-based repository	39	*22.4*
An Open Access journal (i.e. author-paying fee journal)	53	*30.5*
Not specified	26	*14.9*
Not sure	14	*8.0*
Total	174	

Respondents could tick more than one option.
Percentages = % of the 174 respondents that answered any part of the question

21 Do you take into account the publishers' open access policies when choosing a journal to which to submit your work?

	Frequency	%
Always	95	*6.8*
Usually	188	*13.4*
Sometimes	328	*23.5*
Rarely	318	*22.7*
Never	469	*33.5*
Total	1,398	*100*

22 Approximately how many of your articles published in the last 5 years are freely available in some form (pre-print or post-print) for download from the following digital platforms?

	None	1-10%	11-25%	26-50%	51-75%	More than 75%	Don't know	Total
An institutional repository	430	97	88	82	59	148	275	1,179
	36.5%	*8.2%*	*7.5%*	*7.0%*	*5.0%*	*12.6%*	*23.3%*	*100%*
A subject-based repository	366	96	90	63	49	130	354	1,148
	31.9%	*8.4%*	*7.8%*	*5.5%*	*4.3%*	*11.3%*	*30.8%*	*100%*
Other open access repositories	380	116	102	54	38	39	369	1,098
	34.6%	*10.6%*	*9.3%*	*4.9%*	*3.5%*	*3.6%*	*33.6%*	*100%*
Your webpage	613	47	46	41	29	150	143	1,069
	57.3%	*4.4%*	*4.3%*	*3.8%*	*2.7%*	*14.0%*	*13.4%*	*100%*

Section 4: Readers - accessing material held in OAR

23 Have you accessed journal articles held in an open access repository (e.g. institutional, subject-based) in the last year?

	Frequency	%
Yes	940	*67.6*
No	226	*16.3*
Can't remember/don't know	224	*16.1*
Total	1,390	*100*

24 How did you reach the repository site?

	Frequency	%
Went directly to the repository	355	*35.1*
Via Google Scholar search	495	*48.9*
Via BASE/OAIster/OpenDOAR/other specialist search engine	41	*4.1*
Via Google search	470	*46.4*
Via other general search engine	115	*11.4*
Library portal, digital library	279	*27.6*
Copy-pasted a link from a citation/reference	170	*16.8*
Link from a researcher's webpage	266	*26.3*
Personal contacts	119	*11.8*
Don't know/can't remember	50	*4.9*
Other	37	*3.7*

Respondents could tick more than one option.
Percentages = % of the 1,012 respondents that answered any part of the question

25 For what purpose did you visit the OAR?

	Frequency	%
Current awareness, keeping up	465	*46.8*
Exploring a new topic	462	*46.5*
Writing an article	638	*64.2*
Writing a research report	218	*21.9*
Writing a research proposal	271	*27.3*
Professional development, continuing education	272	*27.4*
Can't remember	61	*6.1*
Other	25	*2.5*

Respondents could tick more than one option.
Percentages = % of the 994 respondents that answered any part of the question

26 Were you looking for:

	Frequency	%
A specific article?	675	*69.0*
Work by a particular author/research group?	393	*40.2*
Articles about a particular topic (keyword search)?	703	*71.9*
Other	7	*0.7*

Respondents could tick more than one option.
Percentages = % of the 978 respondents that answered any part of the question

27 What version(s) of the article(s) were you hoping to find?

	Frequency	%
Pre-print	111	*11.2*
Author's final version (i.e. post-refereeing but before publisher's copy-editing)	236	*23.9*
Published final version (publisher's PDF file)	611	*61.8*
Not important - any version	320	*32.4*

Respondents could tick more than one option.
Percentages = % of the 989 respondents that answered any part of the question

28 What version did you actually get?

	Frequency	%
Pre-print	117	*11.8*
Author's final version (i.e. post-refereeing but before publisher's copy-editing)	181	*18.2*
Published final version (publisher's PDF file)	501	*50.5*
Don’t know/can’t remember	194	*19.5*
Total	993	*100*

29 To what extent was the version of the article you accessed satisfactory for the purpose of your task?

	Frequency	%
Very satisfactory	395	*39.7*
Quite satisfactory	516	*51.9*
Not very satisfactory	24	*2.4*
Not at all satisfactory	1	*0.1*
Don't know	58	*5.8*
Total	994	*100*

30 How difficult was it to evaluate/assess whether the version accessed would be suitable for the purpose of your task?

	Frequency	%
Very easy	226	*22.7*
Easy	408	*40.9*
Neither easy nor difficult	261	*26.2*
Difficult	30	*3.0*
Very difficult	2	*0.2*
Don't know	70	*7.0*
Total	997	*100*

31 If you did not find the version you were looking for, did you try another method to access it?

	Frequency	%
Yes	439	*45.0*
No	143	*14.7*
Not sure	98	*10.0*
Not applicable	296	*30.3*
Total	976	*100*

32 Please comment on any aspects of your Open Access Repositories experience:

141 responses

Appendix 3.2 Summary of phase 2 exit survey responses

Section 1: About you

1 Which of the following best describes your institution?

	Frequency	%
University or college	26	*76.5*
Hospital or medical school	0	*0.0*
Research institute	6	*17.6*
Government	0	*0.0*
Industrial/commercial	1	*2.9*
Other	1	*2.9*
Total	34	*100*

2 In which country is your institution based?

	Frequency	%
Austria	1	*2.9*
Czech Republic	1	*2.9*
Germany	27	*79.4*
Ireland	2	*5.9*
Italy	1	*2.9*
Romania	1	*2.9*
UK	1	*2.9*
Total	34	*100*

3 Which of the following best describes your role?

	Frequency	%
Undergraduate or Masters student	12	*35.3*
PhD student	8	*23.5*
Postdoctoral researcher	2	*5.9*
Academic staff	9	*26.5*
Other researcher	3	*8.8*
Total	34	*100*

4 For how long have you been involved in research?

	Frequency	%
Fewer than 3 years	14	*41.2*
3-5 years	9	*26.5*
6-9 years	3	*8.8*
10-14 years	0	*0.0*
15-24 years	4	*11.8*
25 years or longer	2	*5.9*
Not applicable	2	*5.9*
Total	34	*100*

5 Which field best describes your research area?

	Frequency	%
Medical sciences	0	*0.0*
Life sciences	1	*2.9*
Physical sciences & mathematics	2	*5.9*
Social sciences, humanities & arts	21	*61.8*
Interdisciplinary	10	*29.4*
Total	34	*100*

6 Please tell us your age:

	Frequency	%
25 years or under	5	*14.7*
26-35 years	16	*47.1*
36-45 years	4	*11.8*
46-55 years	6	*17.6*
56-65 years	2	*5.9*
Over 65 years	1	*2.9*
Prefer not to say	0	*0.0*
Total	34	*100*

Section 2: Information seeking

7 Is this the first time you have visited an OAR?

	Frequency	%
Yes	11	*32.4*
No	22	*64.7*
Don't know	1	*2.9*
Total	34	*100*

8 How did you arrive at this repository?

	Frequency	%
Went directly to the repository	5	*14.7*
Via Google Scholar search	10	*29.4*
Via BASE/OAIster/OpenDOAR/other specialist search engine	1	*2.9*
Via Google search	15	*44.1*
Via other general search engine	0	*0.0*
Library portal, digital library	2	*5.9*
Copy-pasted a link from a citation/reference	2	*5.9*
Link from a researcher's webpage	2	*5.9*
Personal contacts	3	*8.8*
Don't know/can't remember	1	*2.9*
Other	2	*5.9*

Respondents could tick more than one option.
Percentages = % of the 34 respondents that answered any part of the question

9 For what purpose did you visit the OAR today?

	Frequency	%
Current awareness, keeping up	4	*11.8*
Exploring a new topic	13	*38.2*
Writing an article	2	*5.9*
Writing a research report	2	*5.9*
Writing a research proposal	7	*20.6*
Literature review	17	*50.0*
Completing a student assignment	4	*11.8*
Professional development, continuing education	4	*11.8*
Can't remember	0	*0.0*
Other	2	*5.9*

Respondents could tick more than one option.
Percentages = % of the 34 respondents that answered any part of the question

10 On this visit were you looking for:

	Frequency	%
A specific article?	15	*45.5*
Work by a particular author/research group?	7	*21.2*
Articles about a particular topic (keyword search)?	16	*48.5*
Other	1	*3.0*

Respondents could tick more than one option.
Percentages = % of the 33 respondents that answered any part of the question

11 For the last article you viewed, what version were you hoping to find?

	Frequency	%
Pre-print	0	*0.0*
Author's final version (i.e. post-refereeing but before publisher's copy-editing)	0	*0.0*
Published final version (publisher's PDF file)	19	*55.9*
Any version after peer review	5	*14.7*
Not important - any version	10	*29.4*
Total	34	*100*

12 Which version have you actually accessed?

	Frequency	%
Pre-print	0	*0.0*
Author's final version (i.e. post-refereeing but before publisher's copy-editing)	7	*21.2*
Published final version (publisher's PDF file)	21	*63.6*
Not sure	5	*15.2*
Total	33	*100*

13 To what extent was the version of the article you accessed satisfactory for the purpose of your task?

	Frequency	%
Very satisfactory	17	*50.0*
Quite satisfactory	11	*32.4*
Not very satisfactory	0	*0.0*
Not at all satisfactory	2	*5.9*
Don't know	4	*11.8*
Total	34	*100*

14 How difficult was it to evaluate/assess whether the version accessed would be suitable for the purpose of your task?

	Frequency	%
Very easy	8	*23.5*
Easy	10	*29.4*
Neither easy nor difficult	12	*35.3*
Difficult	1	*2.9*
Very difficult	1	*2.9*
Don't know	2	*5.9*
Total	34	*100*

15 If you have not found the version you were looking for, do you intend to try another method to access it?

	Frequency	%
Yes	17	*53.1*
No	2	*6.3*
Not sure	1	*3.1*
Not applicable	12	*37.5*
Total	32	*100.0*

16 Would you access material held in an OAR again?

	Frequency	%
Yes	29	*90.6*
No	2	*6.3*
Not sure	1	*3.1*
Total	32	*100.0*

17 Would this experience with open access material encourage you to place copies of your own articles in OAR?

	Frequency	%
Yes	24	*70.6*
No	4	*11.8*
Not sure	6	*17.6*
Total	34	*100.0*

18 Please comment on your experience of using Open Access Repositories:

15 responses