Self-Archiving and Journal Subscriptions: Co-existence or Competition?

An International Survey of Librarians' Preferences

Chris Beckett and Simon Inger
Scholarly Information Strategies Ltd

The Publishing Research Consortium (PRC) is a group representing publishers and societies supporting global research into scholarly communication, with the aim to provide unbiased data and objective analysis. Our objective is to support work that is scientific and pro-scholarship. Overall, we aim to promote an understanding of the role of publishing and its impact on research and teaching.

Outputs from work supported by the PRC are available from the website: www.publishingresearch.org.uk.
Hard copies are available from The Publishers Association.

The founding partners are The Publishers Association, the Association of Learned and Professional Society Publishers, and the International Association of Scientific, Technical & Medical Publishers. Corresponding partners include The Association of American University Presses and the Professional/Scholarly Publishing Division of the Association of American Publishers.

Executive Overview

1 A major study of librarian purchasing preferences (*Self-Archiving and Journal Subscriptions: Co-existence or Competition? An International Survey of Librarians' Preferences*[1]) has shown that librarians will show a strong preference for Open Access (OA) materials as they discover that more and more learned material has become available in institutional repositories. The tipping point at which journal subscriptions start to be cancelled may be closer than is currently thought.

2 The study shows that librarians are very sensitive to quality, content cost, the version of the content and how immediately the content is made available.

3 Overall the survey suggests that a significant number of librarians are likely to substitute OA materials for subscribed resources, given certain levels of reliability, peer review and currency of the information available. This last factor is a critical one – a preference for OA resources diminishes significantly if access is embargoed for a significant length of time.

4 The key attributes identified in this study, apart from the universal requirement for content quality and sensitivity to cost, were what version of the content (author's preprint, refereed manuscript, final published version, etc.) is made available and how up-to-date the content is (the embargo period).

   a. There is a strong preference for content that has undergone peer review. The authors' unrefereed original manuscripts were seen as a poor substitute for any post-refereed version of an article. Librarians showed an insignificant shift in preference between any versions of an article once it had been refereed, irrespective of the inclusion of editorial changes such as copy editing.
   b. The change in the librarian's preference for the subscribed journal over the same content in an OA archive is greatest, in favour of the subscribed journal when the only version of the content available in the OA archive is the unrefereed author's submitted manuscript.
   c. How soon content is made available is a key determinant of librarians' acquisition behaviour; delay in availability reduces the attractiveness of a product offering. A significant impact on librarians' preference for OA, and licensed database content, was seen when embargoes were set to 12 and 24 months. A 6-month embargo has little impact.

5 Perhaps unsurprisingly librarians show a strong preference for content that is made freely available, all other factors being equal. Even as librarians were asked to trade off price considerations against other factors such as the version of the content and the immediacy of its availability, there remained a significant pull towards free content or content whose cost had been greatly reduced.

6 The great majority of librarians surveyed welcomed the challenge that Open Access presents to established publishing models.

7 While many disagree, there is a high level of confidence in the reliability of content on Open Access archives.

8 Only 38% believe that publishers should not worry about libraries cancelling subscriptions because of Open Access repositories, and as many disagree (or think that publishers should worry).

9 As many as 40% believe that libraries are wasting their money subscribing to journals when almost the same content is available for free on repositories; but a similar proportion disagree.

Context

The debate on new publishing models has suffered from a marked lack of data as was clear from the report analysing data on scholarly journals publishing commissioned by the Research Information Network (RIN), Department of Trade and Industry (DTI) and Research Councils UK (RCUK) and carried out by EPS.[2] The Publishing Research Consortium (PRC) aims to identify where more data on scholarly communication is needed and commission research work to provide such data.

An obvious need was the question of whether self-archiving could lead to librarians cancelling subscriptions. Some commentators felt there is no risk, citing the example of pre-print servers in physics, and several funders were considering or implementing new policies for mandating self-archiving within six months of publication of the article.

PRC felt that our study[1] produced convincing evidence that self-archiving could put journals at risk and that the issue is of sufficient importance to merit a Summary Paper for wider dissemination of our findings.

Channels of communication for scholarly content

Prior to the advent of the online electronic journal in the mid 1990s the channels by which scholarly journal content was distributed were limited and distinct. Print versions were obtained by libraries either directly from the publisher or via subscription agents and electronic versions were only available from the providers of CD-ROM databases of journal articles. The journals in these databases were then, and still are, licensed from the primary publisher. The print journal and the CD-ROM licensed database represented two very distinct products, with different functionalities, addressed largely separate markets and were delivered via two very different interfaces – the printed page, and a PC screen only available on a dedicated workstation in a library. The CD-ROM database version was limited in its functionality compared to today's electronic journals due to technical limitations of that time. Specifically they rarely had images and were largely limited to ASCII files of the article's content. Researchers obtained content as a result of libraries purchasing the content in whichever form was most appropriate to their institutional remit and requirements.

The last 10 years has seen the web become the dominant form of delivery for most scholarly materials and has led to a plethora of other ways in which content can be delivered. Specifically of relevance to this study is the possibility of researchers accessing scholarly articles via the author's self-archived copy rather than via the subscribed journal. The Budapest Open Access Initiative[3] (BOAI) was signed in February 2002 and since then there has been considerable debate and some progress towards enabling and achieving author self-archiving[4]; the goal of strand 1 of the BOAI. Alongside this has been the growth of a market for institutional archive software and also the means to aggregate the scholarly article content distributed across dispersed institutional archives via the Open Archives Initiative Protocol for Metadata Harvesting (OAIPMH).[4] Via this mechanism search engines, dedicated subject gateways, and Abstracting and Indexing services can harvest the metadata from all archives making the metadata visible in a form appropriate to the target user group and providing links to the full text on the distributed repositories. The net effect of this is to provide, at least in outline, the beginning of a new delivery infrastructure that constitutes an alternative channel by which the researcher can access scholarly article material.

A further impact of the web on the scholarly information chain, in addition to the growth in the number of delivery channels, is that the distinction between these channels has become blurred. In 1995 there was a distinct difference between a print journal and an ASCII representation of the text available on a CD-ROM both in terms of content and format. Today the author's self-archived, peer-reviewed and PDF'd article (post-

print), the final published article PDF delivered from the publisher's server and the article PDF delivered from an online aggregated database all share many commonalities. They all can share the same PDF format and all can have the relevant text and images. There are small variations in content, the result of copy editing, and page formatting that are likely to distinguish the author's self-archived peer-reviewed copy of an article from the final published version.[5] In addition the final published version is likely to have additional functionality such as live hypertext reference links enabling the reader to navigate further through the literature. The licensed database version of the article and the publisher's version are essentially identical – from the user's perspective only the means of accessing them varies.

This convergence of format, the rise of ubiquitous search and retrieve mechanisms such as Google, and the universality of web-based delivery all combine, from the researcher's perspective, to obscure where an article comes from, whether or not it is the final published version, and the price, if any, that has been paid for it. From the content owner's perspective the combination of these factors increases the possibility of one source of a document substituting for another. In addition paid-for content is often delivered through third party interfaces such as Google where the source of the content and the fact that it has been paid for via a library licence is often hard to recognize. Much paid-for content appears free to the researcher.

Alternative acquisition methods

Currently most content is delivered to researchers as a result of an acquisition decision made by a librarian. The librarian today has multiple choices as to how the same article can be obtained electronically; via an aggregated database, via a journal subscription, and potentially, as the volume of self archived material grows, via an institutional or central repository of author self-archived content.

Given this, scholarly publishers have an interest in discovering how these alternative acquisition choices are perceived by those librarians responsible for selecting and acquiring content. Publishers need a clearer understanding of the main drivers behind selection decisions as they apply across these alternative ways of obtaining content, and therefore whether or not these alternative acquisition routes are likely to affect their current core business of selling journal subscriptions to libraries. Specifically they need to develop models which can predict behaviour. Publishers have a vested interest in confirming or otherwise whether the relationship between content on institutional or subject archives and subscriptions is one of mutualism or parasitism. Does the availability of Open Access articles on an institutional repository threaten the journal subscription and what are the key factors that determine when and if that threat becomes significant? In particular both they and librarians (although for different reasons) have an interest in identifying at what point switching from one means of acquiring content to another is likely to make sense. Identifying the tipping point (if there is one) is of importance to libraries because it provides a marker as to when they can begin changing their acquisition behaviour and for publishers because it indicates when subscription revenues are likely to be significantly threatened and alternative revenue streams need to be developed, if funding bodies policies remain unchanged.

Previous work in this area sponsored by ALPSP[6] generally concluded that content on OA repositories is not currently seen by librarians as a substitute for properly managed journal holdings. However it also concluded that 53% (rising to 81% in the next five years) saw the availability of content via OA archives as an important or very important factor in determining cancellation.

More broadly a better understanding of the factors that generally determine acquisition behaviour in different sectors and different geographical markets will assist publishers in developing more effective editorial, product development and sales and marketing strategies, and inform their debate with other stakeholders.

[5] Wates, E. and Campbell, R. Are there real differences between the author's version of an article and the publisher's version? An analysis of the copy-editing function. *Learned Publishing*, 2007. In press.

[6] Ware, M. ALPSP survey of librarians on factors in journal cancellation, 2006. http://www.alpsp.org/publications/pub12.htm

Purpose

The objectives of this study were as follows:

To identify and prioritize the main factors that determine librarians' acquisitions decisions.

To model how these factors interact and therefore identify the combination of product factors with the most appeal.

To look at librarians' intentions in respect of whether or not author self-archiving is likely to result in cancellation of library subscriptions.

To identify at what point substitution of one product for another (author self-archiving or aggregated database access, for journal subscription) is likely to take place and how these interact, i.e. identify the tipping point.

To investigate librarians' general attitudes to self-archiving.

Given that a key objective of the research was to predict the choices that decision-makers would make based on a competing set of alternatives (i.e. acquiring scholarly information from sources other than directly from the primary publishers), it was decided to utilise for the project a form of conjoint analysis (specifically the latent class, maximum-differential approach). By presenting respondents with a series of anonymous product configuration scenarios and asking them to select the one they preferred the most and the one they preferred the least, it is possible to infer the relative importance of different attributes of scholarly content in driving their acquisition choices. It also enables the creation of a 'share of preference' model that predicts the likely uptake of different product configurations.

The following analyses were prepared to support the main report:

Conjoint analysis of the most important factors that librarians will take into account affecting purchasing, renewal and collection development.

Identification of librarian segments based on their purchasing needs.

Attitudes towards Open Access archiving and repositories.

Predicted behaviour based on the impact of different scenarios using a simulator that was created as part of the analysis.

Detailed sub-group analysis to identify how sub-groups differ in terms of purchasing behaviour.

The influence and interaction of some of the main factors affecting acquisition decisions

Conjoint analysis was used to test six *attributes*. These attributes were identified in discussion with senior decision-making librarians as being some of the most important factors that influence acquisition decision making.

Respondents were presented with different *levels* of each *attribute* to identify the point at which they would trade one *attribute* off against another and thus established their relative importance. The attributes tested were:

Version of Article

Percentage of a journal's articles that are available

Reliability of Access

How up-to-date is the content

Quality of the content

Cost

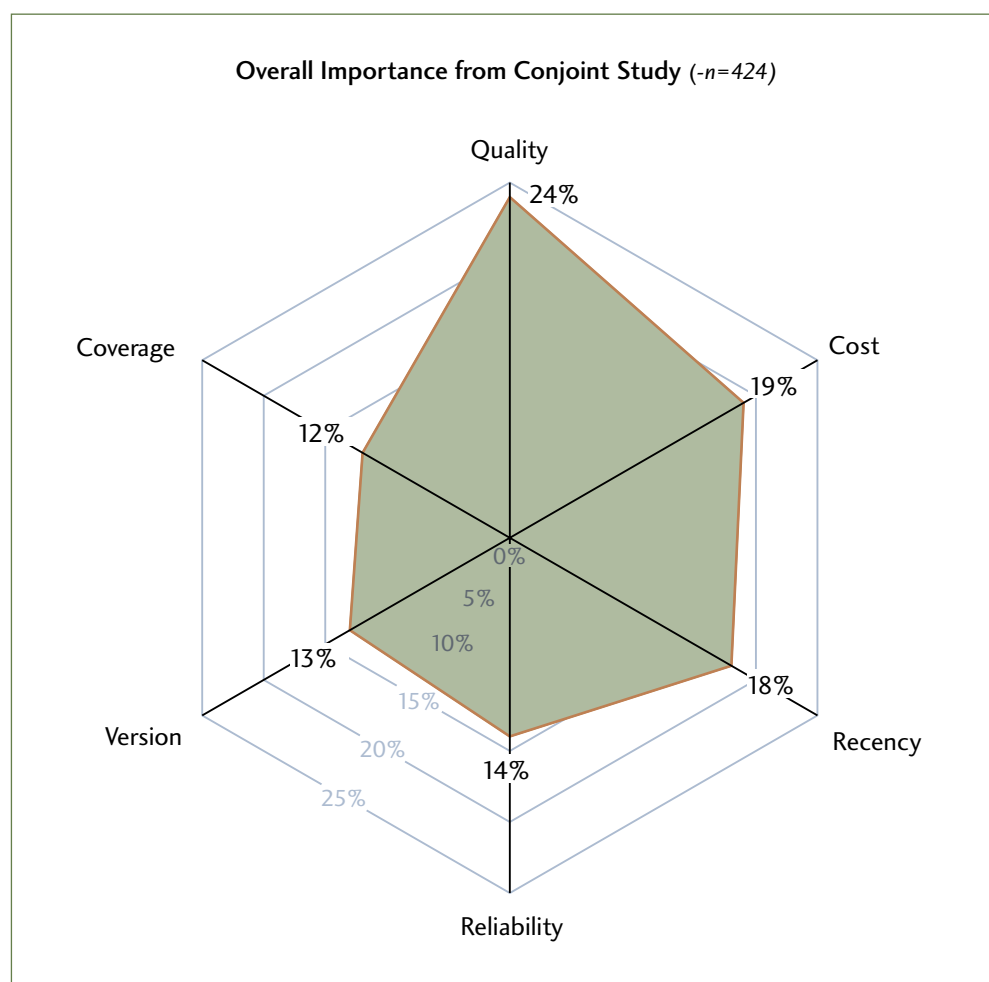## Acquisition preferences in different market scenarios

Overall Importance from Conjoint Study *(-n=424)*

Figure 1
Overall importance of attributes as a radar plot

Figure 1 shows the overall importance of the attributes tested in the conjoint analysis. The percentages indicate the relative 'pull' of each of these attributes in librarians' decision making.[7] The chart shows that librarians' primary driver in making acquisition decisions is, as would be expected, the quality of the content. It has the strongest influence, taking priority over cost and other attributes.

This is followed by *Cost* and *Recency of publication* as the next most important and these are the main factors that would be considered for a given quality of journal.

Amongst the other attributes there is little difference in values so they have a similar, and lesser, influence on decision making. They are, in order of importance: *Reliability of access*, *Version availability*, and *Percentage of articles available*.
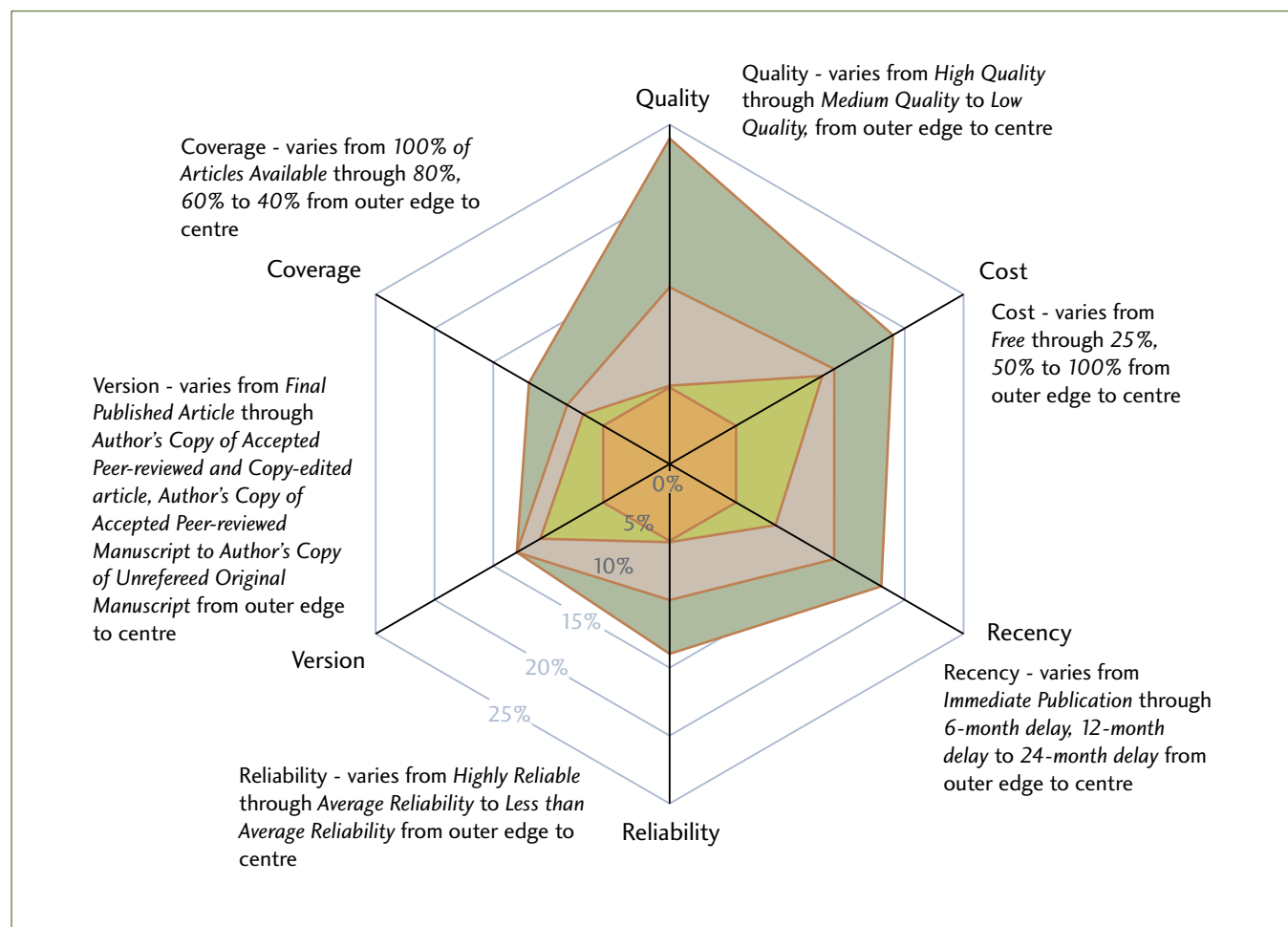

Quality - varies from *High Quality* through *Medium Quality* to *Low Quality,* from outer edge to centre

Coverage - varies from *100% of Articles Available* through *80%, 60% to 40%* from outer edge to centre

Version - varies from *Final Published Article* through *Author's Copy of Accepted Peer-reviewed and Copy-edited article, Author's Copy of Accepted Peer-reviewed Manuscript* to *Author's Copy of Unrefereed Original Manuscript* from outer edge to centre

Cost - varies from *Free* through *25%, 50%* to *100%* from outer edge to centre

Recency - varies from *Immediate Publication* through *6-month delay, 12-month delay* to *24-month delay* from outer edge to centre

Reliability - varies from *Highly Reliable* through *Average Reliability* to *Less than Average Reliability* from outer edge to centre

Figure 2

Overall importance of attributes and their levels – shown as a radar plot

Figure 2 shows the relative importance of each of the *attributes* and within each *attribute*, the influence of each of the *levels*.

The green shape is the same as shown in Figure 1 and represents the influence of the *attribute* overall, while the concentric polygons contained within it represent the levels tested within the *attributes*. The closer together that any two apexes sit, the smaller the difference between the two levels. For example in the case of *Recency* there is a significant and clear difference between each polygon, whereas the data points and therefore the polygons within *Version* are much less highly spaced, with the *Published version* and the *Author's copy of the accepted, peer-reviewed and copy-edited article* being held in the same favour. What this indicates is that in the case of the attribute *Recency* – which seeks to measure the impact of embargoes – the

impact of moving from a 6-month, to 12-month , to 24-month embargo period has an evenly stepped (negative) impact on the appeal of content restricted in that way.

In the case of the attribute *Version* – which seeks to measure the impact of different versions of the article as it moves through the editorial process – the impact of the additional editorial changes as the article moves from the *Author's copy of the accepted, peer-reviewed manuscript* to the *Author's copy of the accepted, peer-reviewed and copy-edited* article and on to the *Final published article* is much more muted. There is a clear positive increase in preference for articles that have been peer reviewed over the non-peer-reviewed *Author's copy of the unrefereed, original manuscript*. Once peer review has taken place, however, librarians do not show the same step change in preference based on the subsequent editorial effort as occurs between different levels of embargo.

In terms of policy this suggests that increasing the delay [*Recency*] before an author can self-archive on an institutional repository (or equally before a licensed database can release journal issues in its database) has a fairly predictable and stepped impact on the appeal that such content incarnations have in terms of librarians' preferences.

On the other hand, in the context of the *Version* of the article that is available on Open Access institutional repositories the preference of librarians for one version over another is not as equally stepped. The *Author's copy of the unrefereed, original manuscript* is deprecated but once peer review has occurred there is little impact on preference as a result of subsequent editorial input.

The importance of peer review

This study confirmed the importance of content quality and cost as key factors in determining acquisition choices. This suggests that producing high quality content at reasonable costs to the library is going to be the most successful long-term strategy for any publisher. Unsurprising though this is it is worth emphasizing the importance of these two factors; library market penetration – visibility for authors on the user's desktop – is dependent upon strong editorial development allied to pricing policies that translate to strong value-for-money propositions for libraries. Apart from these the two other significant factors identified in this study, that influence acquisition preference, were what version of the content (author's preprint, etc.) is made available and how up-to-date the content is (the embargo period).

There is a strong preference for content that has undergone peer review. The author's unrefereed original manuscript was seen as a poor substitute for any post-refereed version of an article. This perhaps predictable finding was accompanied by the more interesting finding that librarians showed an insignificant shift in preference between any version of an article once it had been refereed. The inclusion of editorial changes such as copy-editing had little impact on preference share.

Figure 3 shows that the change in the librarian's preference for the subscribed journal over the same content in an OA archive is greatest (in favour of the subscribed journal) when the only version of the content available in the OA archive is the author's submitted manuscript.
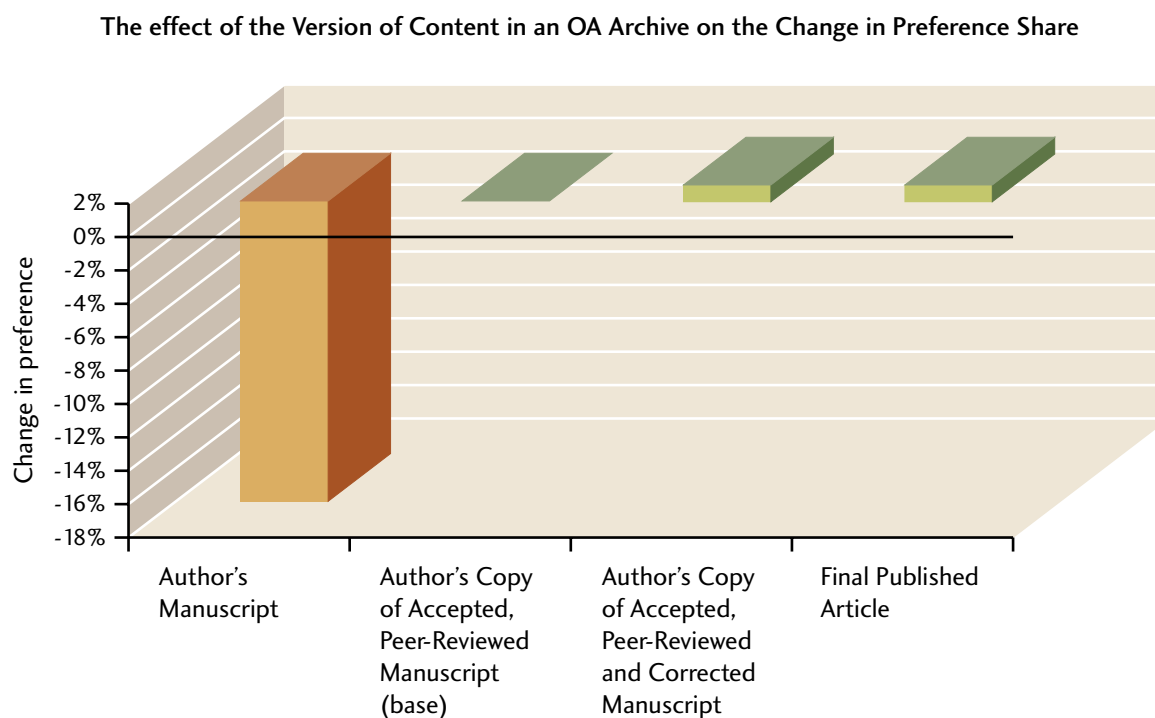
The effect of the Version of Content in an OA Archive on the Change in Preference Share

Change in preference

2%
0%
-2%
-4%
-6%
-8%
-10%
-12%
-14%
-16%
-18%

Author's Manuscript | Author's Copy of Accepted, Peer-Reviewed Manuscript (base) | Author's Copy of Accepted, Peer-Reviewed and Corrected Manuscript | Final Published Article

Figure 3
The effect of version of content on the change in preference share

The implications of this are potentially significant for publishers. It suggests that the value added by copy-editing and other editorial changes including the inclusion of reference links and other value-added features is not a strong defence against substitution by content from institutional repositories.

The impact of access delays

How soon content is made available is a key determinant of content model preference in librarians' acquisition behaviour; delay in availability reduces the attractiveness of a product offering. The survey tested the effect of access embargoes on OA and licensed database content set at 6, 12 and 24 months. There was a significant impact on librarians' preference for OA, and licensed database, when embargoes were set to 12 and 24 months. On the other hand, a 6-month embargo has little impact.

Figure 4 shows the share of preference for degrees of embargoed and non-embargoed content in an institutional repository versus paid-for journal articles, assuming 100% of content is available in the archive. Only when the embargo is extended to 24 months in this model, does the final published article obtain a greater than 50% share of preference.

Figure 4
The share of preference for a paid-for final-published article versus an Open Access article
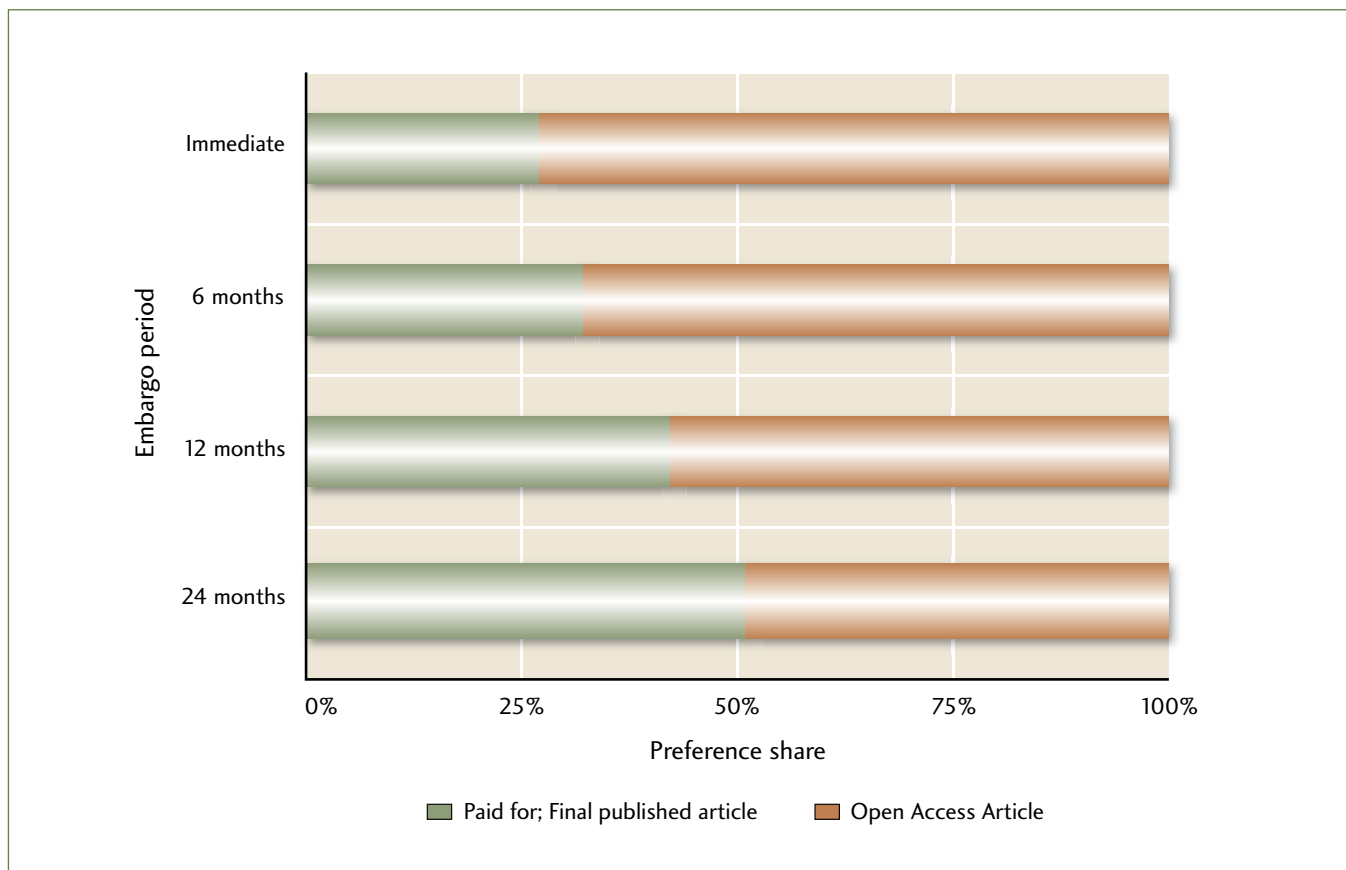
This suggests that generally a 6-month embargo offers only a limited defence against erosion of subscriptions. This is true both in the context of embargoes on author self-archiving on institutional repositories and in the context of licensed databases. This presents publishers with a dilemma. The presumption that a 6-month embargo offers a defence against subscription erosion is a common assumption that under-pins the current de-facto application of this limit to author self-archiving. This study suggests that as the percentage of self-archived material grows this embargo period will be less effective in preventing substitution of self-archived content for subscribed content. In the extreme case of 100% availability of content on the insti-tutional archives and a 24-month embargo, still nearly half the market for subscrip-tion journals has disappeared. In addition, presumably a 24-month embargo period on self-archiving would have a significant impact on the attractiveness of a particu-lar journal to an author.

Lastly and perhaps unsurprisingly librarians show a strong preference for content that is made freely available, all other factors being equal. Even as librarians were asked to trade off price considerations against other factors such as the version of the content and the immediacy of its availability, there remained a significant pull towards free content or content whose cost had been greatly reduced.

Share of Preference between Journals and Open Access Content

Influence of availability
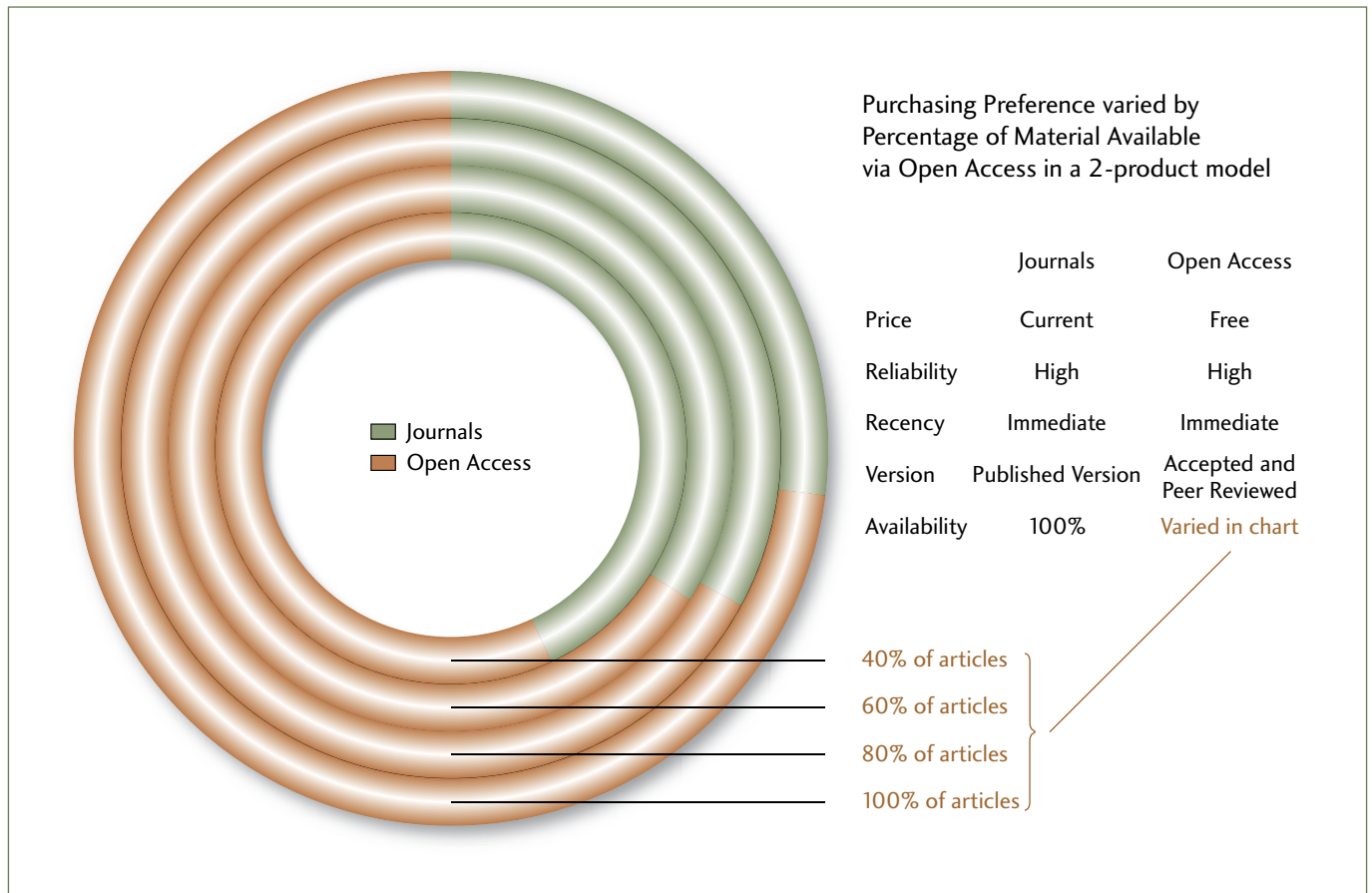
Purchasing Preference varied by
Percentage of Material Available
via Open Access in a 2-product model

	Journals	Open Access
Price	Current	Free
Reliability	High	High
Recency	Immediate	Immediate
Version	Published Version	Accepted and Peer Reviewed
Availability	100%	Varied in chart

40% of articles
60% of articles
80% of articles
100% of articles

One of the other major areas of current and future interest, perhaps because it is one of the more simply measured, is what effect the level of content availability might have on the appeal of the subscribed journal versus content on archives versus content in licensed databases. Both the journal and the licensed database have 100% of the articles in any given journal issue or volume. However, institutional archives contain only a percentage – currently on average around 15% of the content of any journal. The conjoint survey asked respondents to make choices between the varying percentages of articles available in a variety of product offerings. From this we are able to look at the effect on choice of varying just this 'Availability' aspect, with all the other parameters fixed.

To create Figure 5 we have set those other parameters in a the way that best describe Open Access today[8] and those that best describe published journals today and then varied just the percentage of the journal's articles available via Open Access in order to study the impact of availability on librarian choice.

The data show and the Figure 5 illustrates a significant shift of preference towards Open Access as increasing percentages of journal articles become available via Open Access.

Where 40% of articles are available by Open Access, only 43% of librarians' preference is for journals, falling to just 27% of preference when 100% of articles are available by Open Access.

While this might not be unexpected these results do not concur with the current observed behaviour and choices, especially within physics where it is widely accepted that a very high percentage of content is available via ArXiv. This may be due to a number of factors including:

1. That the content 'Version' available in ArXiv can vary between preprint and peer-reviewed post-print which is not the 'Version' modelled here.

2. That widespread librarian awareness of these archive resources as a potential source of supply of content is low.[9]

3. That librarians are used to having content presented to them by publishers and intermediaries as pre-packaged products (such as journals, journal databases, and licensed databases) and require a similar degree of productisation and associated support including looking for a gateway-style, common-overlay to Open Access content before the shift in behaviour that the model predicts becomes apparent.

4. That librarians are looking for compatibility of Open Access gateways with Link Resolvers and other library technologies.

5. That there are other more emotive factors or conservatism that stand in the way of the predicted behaviour.

This disconnection between predicted behaviour and observed behaviour is one of the more interesting areas for exploration, and one that now has been widely discussed. In the view of the authors the most important likely explanation is to do with item 3 (above). The chances of OA self-archived content representing a significant challenge to the journal subscription will depend upon a factor not explicitly assessed in this paper which is the degree of organization – or packaging – of the OA content. Currently self-archiving is at a relatively low level. While the adoption of various mandates by funding bodies may accelerate the rate of deposition significantly, of itself this may not be sufficient to provide the library market as a whole with a realistic alternative to the journal subscription. If commercial or not-for-profit third parties, however, provide a coherent level of overlay to the self-archived content and offer it for acquisition then we believe the uptake would be significantly accelerated. What can publishers do to defend against such a development? Perhaps encouraging or requiring authors to deposit on institutional archives using a Creative Commons Attribution-NonCommercial-ShareAlike 2.5 licence offers some breathing space.

[9]Anecdotally, the authors perceive that awareness of Open Access repositories is still rather low, and is borne out by their experience in working in the library technology training arena.

Attitudes to Open Access

In the second part of the study we undertook an attitudinal survey. In that survey, librarians were asked to state the level to which they agreed or disagreed with a series of statements concerning Open Access.
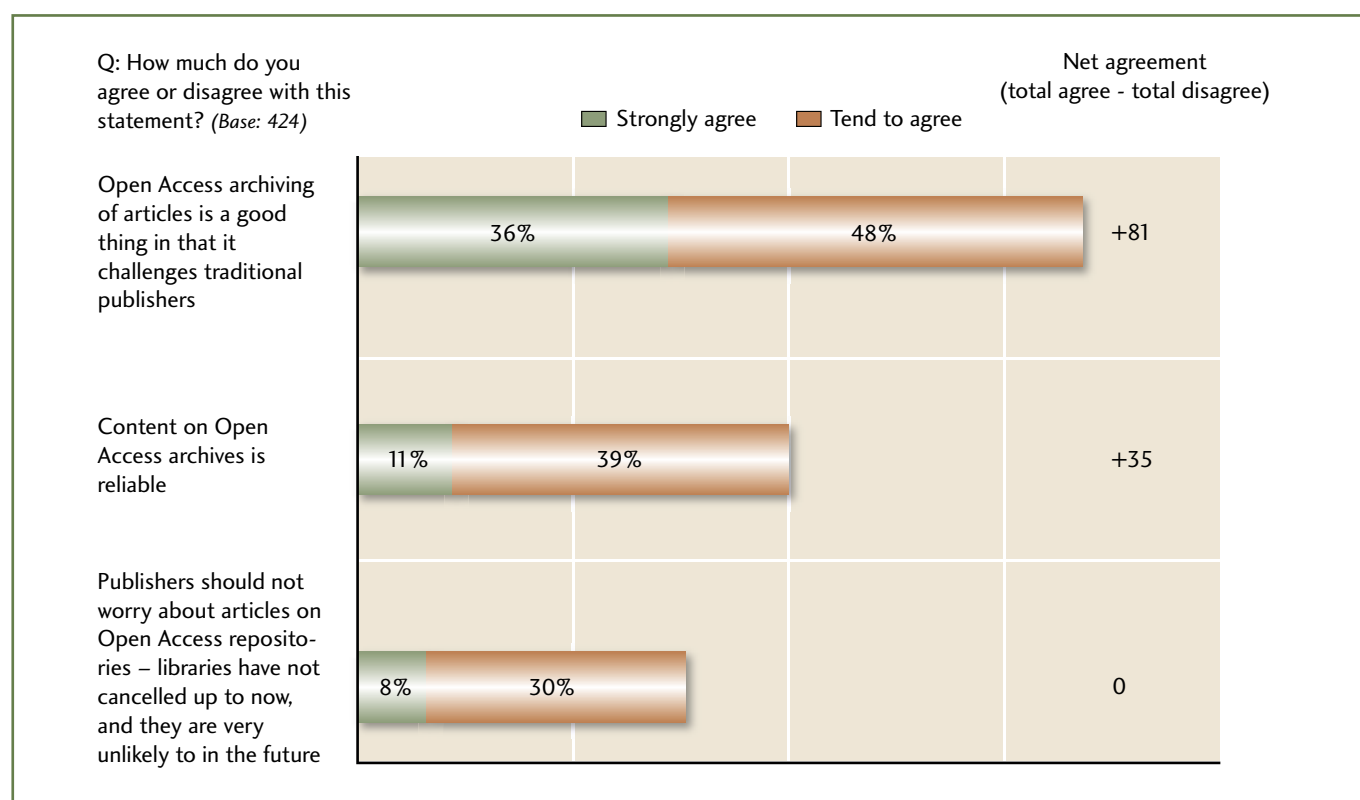
The purpose of the attitudinal survey was to identify if, with specific reference to the issue of Open Access (which until this point had not been mentioned at all in the previous conjoint component of the survey as a whole), librarians' attitudes would shed any light on the likelihood or otherwise of them substituting content in an Institutional repository for content supplied via a subscription journal. The data and Figure 6 show that:

The great majority of librarians surveyed welcomed the challenge that Open Access presents to traditional publishers.

While many disagree, there is a high level of confidence in the reliability of content in Open Access archives.

Only 38% believe that publishers should not worry about libraries cancelling subscriptions because of Open Access repositories, and as many disagree (or think that publishers should worry).

Q: How much do you agree or disagree with this statement? *(Base: 424)*

Net agreement (total agree - total disagree)

Legend: Strongly agree, Tend to agree

Statement	Strongly agree	Tend to agree	Net agreement
Open Access archiving of articles is a good thing in that it challenges traditional publishers	36%	48%	+81
Content on Open Access archives is reliable	11%	39%	+35
Publishers should not worry about articles on Open Access repositories – libraries have not cancelled up to now, and they are very unlikely to in the future	8%	30%	0

Clearly librarians welcome the challenge that Open Access generally offers to publishers. This is unsurprising given that the Open Access challenges focuses mainly on journals and journals have taken an ever increasing proportion of acquisition budgets over the last few years. Any relief from the budget pressure that they experience would presumably be welcome and the high level of net agreement (+81) with the general statement that '*Open Access archiving of articles is a good thing in that it challenges traditional publishers*' is perhaps best understood in that context.

More specifically the fact that only 38% either agreed or strongly agreed with the statement '*Publishers should not worry about articles on Open Access repositories – libraries have not cancelled up to now, and they are very unlikely to in the future*' (and more particularly as many disagreed) suggests that earlier theories that suggested there is no threat of substitution are overstated. This result suggests a willingness on behalf of a substantial number of librarians to substitute content from Open Access repositories for subscriptions.
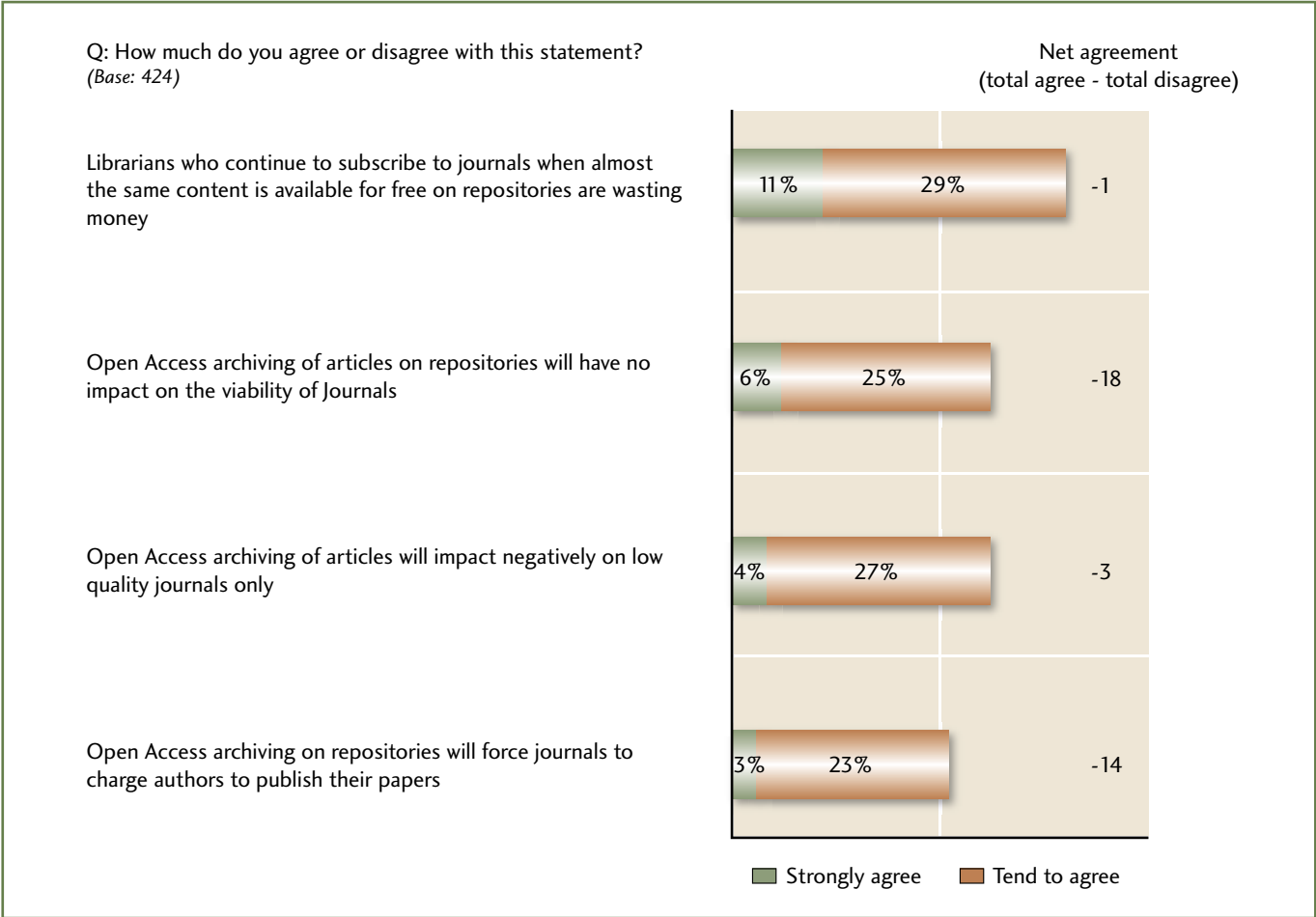
Figure 7 shows that:

As many as 40% believe that libraries are wasting their money subscribing to journals when almost the same content is available for free on repositories; but a similar proportion disagree.

There is concern about the impact repositories will have on journals' viability, though 31% believe it will have no impact.

Just a third agree that Open Access will impact negatively on low-quality journals only, implying that it will also impact negatively on high-quality journals.

A minority (just 26%) believe that journals will be forced to charge authors and more believe this won't be the case.

Q: How much do you agree or disagree with this statement?
(Base: 424)

Net agreement
(total agree - total disagree)

Statement	Strongly agree	Tend to agree	Net
Librarians who continue to subscribe to journals when almost the same content is available for free on repositories are wasting money	11%	29%	-1
Open Access archiving of articles on repositories will have no impact on the viability of Journals	6%	25%	-18
Open Access archiving of articles will impact negatively on low quality journals only	4%	27%	-3
Open Access archiving on repositories will force journals to charge authors to publish their papers	3%	23%	-14

■ Strongly agree   ■ Tend to agree

Opinions are clearly split over whether libraries are wasting their money subscribing to journals when almost the same content is available for free in repositories (40/41). This result fails to provide any reassurance to publishers that there is no threat of substitution of content in an OA archive for the journal subscription.

Conclusion

In some sense the question of Open Access content in an institutional repository as a substitute for journal subscriptions is simply a facsimile of the long-standing channel management challenge that publishers have confronted for many years with reference to licensed databases. The crucial difference is that in the case of licensed databases the publisher is free to choose to proceed or not, once they have come to a balanced decision on the risks and benefits of licensing. They have control.

In granting permission for self-archiving publishers have, although it was not apparent at the time, potentially set up an alternative channel, and one over which they have a lower degree of control. While a publisher can rescind a redistribution licence to a third-party commercial aggregator if they feel that they are suffering from channel conflict and substitution, sometimes called cannibalization, the situation with authors self-archiving is more complex. The publisher's relationship with the author is not a commercial one, so they are unlikely to be able to rescind or retreat from a self-archiving commitment and they fear losing authors and editors as a result. In addition, for some disciplines, they would be closing off the opportunity to publish significant research output because of funder mandates requiring grantees to self-archive.

The authors of this report consider that self-archiving is here to stay – and likely to grow – a result of both a general network effect and funder mandates. We also believe that this research begins to sketch out the specific conditions that will need to be in place for self-archiving to co-exist with the subscription model. Funders will need to decide whether they want to achieve Open Access through policies of mandated self-archiving that minimize the threat to journals (e.g. 12 months or more, depending on the subject) or through further sponsorship of the 'pay-to-publish' model which is now offered widely by Open Access publishers and subscription-based journals (the so-called 'hybrid' model). We believe our research demonstrates that mandating self-archiving within 6 months or less of publication will undermine the subscription-based peer review journal. This, we presume, cannot be in the long-term interest of funding bodies.