First Monday, Volume 24, Number 7 - 1 July 2019

PEER-REVIEWED JOURNAL ON THE INTERNET

The status quo bias and the uptake of open access

by Melissa H. Cantrell and Lauren B. Collister

Abstract

In this paper we argue that the framing of open access through language adopted by a variety of stakeholders serves to inhibit the uptake of open access publishing through the mechanisms of complexity and cognitive load. Using both quantitative and qualitative methods, we analyze both the language and tiers of decisions that confront authors seeking information online about open access. We conclude that this information is for the most part prohibitively complex and introduces contradictory interpretations and executions of open access that act to motivate a phenomenon known as the status quo bias. The only reliable method of counteracting this status quo bias in order to bolster the uptake of open access is to re-frame the language that is commonly employed in association with open access and to minimize the tiers of decisions expected of authors, which create a barrier rather than a gateway to open access engagement.

Contents

Introduction

At the most basic level, open access is a simple concept: the free availability of scholarly work to read and reuse. Open access addresses a problem for anyone who wants to read scholarly research papers: the "paywall", or a cost-to-access charge that a user must pay to read the articles published in many scholarly journals that operate under a subscription pricing model. Open access removes the paywall, making these journal articles (and many other products of research) accessible to anyone with an Internet connection. Open access is disruptive to the economic model of scholarly publication that traditionally relied on the subscription charges to either cover costs or make a profit (Tennant, *et al.*, 2016). As open access publishing has become a mature concept, its complexities and nuances have increasingly become a topic of discussion and deliberation. Advocates, scholars, and publishers discuss questions like the following: How does copyright intersect with availability, and what is the best way to use permissive licenses like Creative Commons licenses to make work available? Is it truly open access if the author has to pay for an article processing charge (APC) in order to get an article published? Is it open access if it is restricted to journal subscribers only through a one year (or longer) embargo?

While we agree that discussing these complexities of open access is a worthwhile endeavor, actual uptake of open access has not progressed quickly. According to the OA2020 Initiative, "Even though open access is now a shared vision of the world's academic communities, research councils, and funding bodies, nearly 85% of the world's scholarly outputs are still locked behind paywalls". Various initiatives have attempted to increase uptake of open access, such as Plan S,

publishing funds, and more. In this work, we argue that there is another fundamental impediment to the uptake of open access: bias towards the status quo.

With this work, we initially set out to identify how definitions of open access varied across different publishers, libraries, and advocacy groups by analyzing the language on Web sites that defined and described open access. In doing this work, we discovered that Web sites for these groups went beyond variation in the definitions of open access, and included considerable variation in the choices and options available to those who wanted to pursue open access publishing. In our initial analysis, we discovered a persistent framing of open access as a "choice" for authors, sometimes being described as an "alternative" to "traditional publishing." This "choice" framing sets up subscription publishing as the status quo, and the use of complex language surrounding open access, we argue, increases the cognitive load on authors who are attempting to make the final decisions about the publication of their research work, and therefore reifies the status quo of subscription publishing [1].

There has been some recent recognition by leaders in scholarly publishing of the need to move beyond the status quo in scholarly publishing. For example, the Open Access Scholarly Publishers Association (OASPA, at https://oaspa.org) — a trade association of which nearly half of the source texts for this study are members — noted that "The nature of publishing is changing as, indeed, is its definition and the function of publishers. We agree there is no good argument for maintaining the status quo" (Redhead, 2019). We hope that this research may inform discussion about the power of such statements as well as address the implicit challenges therein.

This research report proceeds first with a review of the literature on the status quo bias and the compounding effects that lead to that bias: cognitive load and complexity. We then proceed to investigate the ways that the Web sites of scholarly publishers, academic libraries, and advocacy organizations use complicated language and provide choices that could potentially trigger a bias towards subscription publishing. We first set out our sources and methods, then our analyses. Finally, we end with a summary, advice for writing about scholarly publishing, and potential future directions for research in this area.

Literature review

While a growing majority agree that it is desirable to make sharing the new norm for scholarly literature (McKiernan, 2014), what is it that prevents authors from actually engaging in strategies to share their research? The status quo bias is well documented in literature as a widespread source of reluctance in taking up new practices and ideas, but the ways that this mechanism might apply to changing practices in scholarly publishing have not been adequately explored. In this section we examine the concept of the status quo bias as well as its compounding mechanisms: complex texts, and cognitive load that results from the introduction of multiple or alternative choices in the decision-making process. Our review relies largely on literature from the fields of medicine and psychology as these are the areas in which these concepts have been adequately explored, and acknowledge that a lack of literature on these concepts as they relate to scholarly communication exposes a need for further inquiry.

The status quo bias describes peoples tendency to "favor existing and longstanding states of the world" [2]. In other words: "What is, is good." There are many contributors to the existence and strength of the status quo bias in human psychology, some of which are rational and some of which are "non-rational." Rational motivations for an aversion to change include transactional costs, such as time and effort, the cognitive costs of making a novel decision even if it is objectively superior, and a preference for what satisfices over what may be uncertain. "Non-rational" justifications for the status quo include loss aversion or regret avoidance, because regret increases when a decision is changed rather than maintained. Mere exposure to something repeatedly over time and its existence and longevity are rationalized in human psychology as favorable for that thing's continued existence and persistence (Eidelman and Crandall, 2012).

In other words, there is a general human tendency to be averse to new, novel, or what can be perceived as risky decisions. The literature in this area points to various effects all in a similar theme. In Kay and Zanna's (2009) review of key social and psychological literature, this effect is called "system justification" which serves as a "coping mechanism" through which individuals justify existing systems, institutions, and processes even if they are problematic, as any change to that system is perceived as an undesirable threat. System justification theory may also be described as our innate tendency for "'turning lemons into lemonade'" [3]. The status quo bias is central to understanding the advancement of open access publishing because not only does subscription publishing have an innate advantage as the historically established practice in

scholarly publishing, but additionally the multiplication of choices and alternatives may intensify the effects of the status quo bias. The loss aversion and regret avoidance that are emblematic of the status quo bias are rooted in two additional concepts related to the existence of choices and alternatives; these concepts are cognitive load and complexity.

While the status quo bias is a natural bias present at the specter of any choice, the strength of the bias is also compounded by choices that introduce additional intricacies and uncertainty that may strain the cognitive function of the decision-maker. This strain is referred to as cognitive load, defined as "the effort and mental activity imposed on a person's ability to process information. This cognitive load can therefore be seen as an important aspect of bounded rationality" [4]. Allred and colleagues' (2016) study of cognitive load through manipulated memory tasks observes a related effect known as "central tendency bias" Under the central tendency bias, the effect of cognitive load causes one's judgements of present conditions to be skewed by prior memory or experience. In a similar vein to system justification and central tendency bias, the status quo bias makes one highly susceptible to the influence of established experiences and structures. This is of concern because, as we will demonstrate, open access is often presented within a complex matrix of choice. Every additional choice has the effect of compounding the strength of the status quo bias. Eidelman and Crandall argue that "decision-makers are more likely to postpone making a decision as alternatives are added [...], and preference for the status quo increases as a function of the number of options" [5]. In an empirical study on mutual fund market investors, Kempf and Ruenzi conclude that "Our results indicate that fund investors are subject to a SQB [Status Quo Bias]" and that on top of that "it becomes stronger the larger the number of alternatives" [6]. Subscription publishing models do not present any comparable series of choices in order to publish a work, other than the choice of the publication venue itself. If presented through the lens of choice, some models of open access publishing and forms of open access advocacy may be driving authors away from open access rather than encouraging engagement as these authors seek to evade the potential for loss and regret.

Many studies have sought to measure the complexity of a particular text or set of texts through readability level tests. According to Kher, *et al.* [7] "readability is defined through various formulas based on sentence length, word familiarity, syllables, and other factors via scores that identify a grade level needed to attain to comprehend the presented information." Several studies assessing text complexity reviewed patient and consumer health information. Walsh and Volsko (2008) conducted three readability tests on 100 online consumer health information articles in order to determine whether they were written at a comprehensible level for the American public. According to the U.S. Department of Health and Human Services (USDHHS), "material between the 7th and 9th grade levels is viewed as 'average difficulty,' and material above the 9th-grade level is regarded as 'difficult'" [8]. The study found that the mean score of the articles was eight grade levels above the level recommended for comprehension [9]. Similarly, in an assessment of online patient information related to congestive heart failure, Kher, *et al.* applied six readability tests to 70 Google results for "congestive heart failure." Of the 70 Web sites only five were within the USDHHS recommended readability level [10].

In another study seeking to understand the relationship between complexity and consent, Luger, *et al.* applied the SMOG Index ("Simple Measure of Gobbledygook") to Web site terms and conditions. Of the 16 documents selected for the study, all of them fell "beyond the threshold of functional literacy [...] with a mean average readability level of 18.1" [11]. A central concern of complexity in this context is that it may serve to mask crucial information regarding consent and encourage users to "accept" without a full understanding of what they are agreeing to. Complexity, therefore, leads to declining comprehension and applies pressure on onersquo;s ability to provide reasonable consent to new propositions.

Yet with regards to scholarly publishing, is there really a problem with providing a complex vision of open access publishing? It is undoubtedly a complex landscape with no as-of-yet standard set of procedures and practices. Research is conflicted over whether simple or complex Web sites are optimal for user experience, with arguments for the latter citing that complexity provides for a richer experience of the information being presented [12]. However, studies do suggest that when the complexity of a Web site is coupled with a complex task to be completed, a user's cognition suffers greatly. Wang and colleagues describe this "task complexity" as "a function of the amount of task-related information an individual has to process while performing a task. The more information to be processed, the more complex the task is. Complex tasks require more cognitive work, such as psychological comparison" [13]. If there is an overload of information for authors to process on open access, then, the task will appear more complex and thus demand more cognitive energy to achieve success. This increased cognitive energy leads to a higher cognitive load, which — as described above — can lead to a bias towards whatever is presented as the status quo with the least amount of risk and choice involved (Whitney, *et al.*, 2008).

In sum, when decision-makers experience high levels of complexity and cognitive load, there is increased cognitive pressure (in addition to the natural pressure for loss and regret avoidance) to

maintain the status quo. This bias is reinforced with each additional choice introduced, and serves to protect the equilibrium of the system and prevent disruptive changes. To investigate the state of potential status quo bias related to complexity and cognitive load in the scholarly publishing and Open Access models, we analyze source texts from Web sites that define and explain open access and subscription publishing models. In the next section, we describe how we chose the texts and our methods for analysis. We applied a variety of quantitative and qualitative tests to our sample texts in order to assess the various ways in which publishing information provided to authors about may trigger the status quo bias by posing open access as a choice and framing subscription publishing as the status quo, and even further may magnify that bias through complex texts that heighten the cognitive load and uncertainty of authors seeking to make important publication decisions.

Source texts and methods

Because the initial inquiry for this research centered on discrepancies in defining features of the concept of open access, the primary criteria for inclusion within the sample is that all source groups explicitly define "open access" somewhere in the content of the Web site. Other important factors that all source groups included in the sample have in common is that they, at least in part, serve as an "informational" resource about open access and also provide some level of support and advocacy for open access publishing. Twenty-nine source groups were ultimately selected for inclusion within the sample.

Despite these commonalities, the source groups within the sample represent a wide variety of stakeholder groups within the open access movement. We set out to have a sample of 30 organizations, half publishers and half non-publishers. Some clear categories within the sample include non-profit publishers, for-profit publishers, non-profit advocacy groups, informational Web sites, and university libraries. The publishers that were selected for inclusion in the sample are all "large" or "very large" professional publisher members of OASPA (Open Access Scholarly Publishers Association), with the exception of Elsevier which was selected due to its size and prominent role in both subscription and open access publishing. However, the source groups do not include small open access publishers that do not include a paid business model, as these publishers are far less likely to explicitly define open access and rarely serve as a robust information source on open access publishing. Therefor this study favors larger and more prominent source groups since they provide adequate data for our analysis, and this study is not representative of the full spectrum of open access publishers. One source selected at the beginning of the study (Association of Research Libraries) was eventually excluded because of a lack of sufficient information on its Web site for our analyses. The names of sources used in this study appear in Table 1 with the short names used for these sources and the category of that source.

Table 1: Sources used for texts for this study, with short names used in tables throughout this paper, and the categorization of the source.

Source name	Short name	Category
Association of College and Research Libraries	ACRL	Advocacy
Cornell University Libraries	CORNELL U.	Library
Degruyter	DEGRUYTER	Publisher
Securing a Hybrid Environment for Research Preservation and Access	SHERPA	Advocacy
Right to Research Coalition	RIGHT TO RESEARCH	Advocacy
Wikipedia	WIKIPEDIA	Informational

Springer	SPRINGER	Publisher
Open Access Scholarly Information Sourcebook	OASIS	Advocacy
John Wiley & Sons	WILEY	Publisher
Institute of Physics	IOP	Publisher
Nature Publishing Group	NATURE	Publisher
Oxford University Press	OXFORD	Publisher
ProQuest	PROQUEST	Publisher
Hindawi	HINDAWI	Publisher
Creative Commons	CREATIVE COMMONS	Informational
Elsevier	ELSEVIER	Publisher
BMJ Publishing	BMJ	Publisher
Budapest Open Access Initiative	BOAI	Advocacy
Public Library of Science	PLOS	Publisher
Scholarly Publishing and Academic Resources Coalition	SPARC	Advocacy
Massachusetts Institute of Technology Libraries	MIT	Library
University of Kansas Libraries	U. OF KANSAS	Library
Brill Publishers	BRILL	Publisher
Peter Suber's Open Access Overview	PETER SUBER	Advocacy
Cambridge University Press	CAMBRIDGE	Publisher
Harvard University Libraries	HARVARD U.	Library
Taylor and Francis	T & F	Publisher
Sage Publishing	SAGE	Publisher
Open Access Working Group	OAWG	Advocacy

It was prohibitive to this initial scan to conduct a textual analysis of the entire Web site of each group within the sample. We selected three pages from each Web site, which will be referred to throughout this paper as the primary, secondary, and tertiary pages. For each group, the primary page included in analysis is always the page on which the definition of open access was found. The secondary and tertiary pages for each group either link to or are linked from the primary page, and were selected based on the relevance of content [14]. For the qualitative analyses of Choice discussed below, the tertiary page used for coding is either the default tertiary page or, if present, substitutes the default tertiary page with a "Dedicated" tertiary page that addresses the specific topic being coded. For four groups within the sample, fewer than three pages are included in the analysis because there was either only one page with the site's open access content (*e.g.*,

Wikipedia) or the site had fewer than three relevant pages for analysis. A full list of the sample Web pages is available on Github [15].

These pages were converted into text files and analyzed using the Textstat package for Python [16]. Textstat is a package that calculates statistics from a text regarding readability, complexity, and grade level. The measures calculated using Textstat were word count, sentence count, average words per sentence, SMOG index, Flesch Reading Ease, and Combined Grade Level. All texts and analyses, including a Jupyter Notebook with Python commands and results, are available on Github [17].

In addition to the analyses run with Python, two additional qualitative analyses were completed on the text. The first qualitative analysis focused on the definitions of "open access" gleaned from the primary page of each group. The authors selected 10 "factors" that are common concepts to be included in a definition of open access. The authors then coded each definition for whether it did include, maybe included, or did not include that factor. A definition was coded as "Maybe" including a factor, for example, if the language was vague but could be interpreted as being inclusive of that concept/factor.

The second qualitative analysis coded for the presence of choice within the text, and this coding contained two distinct stages. In the first stage, the authors manually searched the sample Web pages for the words "choice," "option," and their derivatives. Web sites that had a specific page addressing hybrid OA substituted the tertiary page with a "dedicated" page, and all others used the default tertiary page. In the second stage of analysis, the authors gathered text manually from all of the sample Web pages about open access "costs." Again, Web sites that had a specific page addressing OA costs (such as fees and article processing charges, also known as APCs) substituted a "dedicated" page for the tertiary page, and all others used the default tertiary page. The authors separately coded the selected text for whether it indicates if the author "always," "usually," "sometimes," or "never" pays for open access, as well as what cost recovery models (other than "author pays") are suggested or promoted by the text.

For both qualitative analyses, the authors coded their responses independently, then collaboratively normalized these codes with discussion of where the coding differed. In rare cases where the authors disagreed on a code after the collaborative discussion, the item was coded as the median of each category ("maybe" or "sometimes"). Vague language sometimes led each author to a different interpretation; we will discuss this vague language in the sections below. Overall, both authors ultimately agreed on the final codes, and there were no disagreements between the authors that could not be traced to what we will later describe as "squishy language."

Results

In this section, we will unfold a story through the lens of the three concepts previously introduced that proposes an explanation for the slow uptake of open access in many fields despite its obvious appeal and advantages. The story starts with the concept of complexity and how it is manifested in the context of information on open access that is transmitted through Web sites. We find in this section that readability scores of these texts categorize the Web sites as very difficult to read, well beyond the recommended levels of readability outlined in other studies of text complexity. Following this, to show the potential effect of these texts on cognitive load, we present the findings of our analysis of the definitions of open access in the texts; we find in this section that definitions vary widely and sometimes include "squishy language" that does not allow for a firm interpretation of meaning. Finally, we address the use of "choice" language on these Web sites and how it may contribute to bias towards the status quo.

Complexity

The Web sites analyzed for this study share some characteristics with each of the studies discussed above. Similar to the samples of online health Web sites, the Web pages assessed for this study seek to be informational. However, the Web pages on open access are also comparable to the Luger, *et al.* study on terms and conditions because of their role in providing policy and publication guidance, which raises similar questions regarding users' knowledge of and consent to espoused publishing routes. On the other hand, a point of distinction between this study and others that have been conducted is that the targeted audience for our sample Web pages will presumably read at a much higher level than the eighth to ninth grade average recommended by the USDHHS.

While we expect a high reading level from those publishing or interested in learning about the publication of scholarly works, the results of the three readability tests indicate that the

complexity of these Web pages is nonetheless exceedingly high. Slightly more than half (52 percent) of the source groups scored at a graduate level or beyond across all three readability scores. Approximately one fifth (21 percent) of the source groups ranked at a college reading level across all three scores, and Public Library of Science (PLoS) was the sole source group that scored at a level recommended for the American public (and achieved this in only one out of three tests). Mean readability level across all source groups is 18.3 for the SMOG Index, 4.6 for the Flesch Reading Ease Score, and 17.6 for the Combined Grade Level. Together, these scores indicate that one must read at a graduate level on average to have a basic comprehension of these sample texts.

Table 2: Three reading level scores by source group.

Key to color code	Beyond post-graduate	Post-graduate	Graduate
	College	High school	
Source	**SMOG Index**	**Flesch Reading Ease Score**	**Combined Grade Level**
ACRL	24.6	-115.6	25
CORNELL U.	24	-11.6	24
DEGRUYTER	Invalid	-212	22
SHERPA	18.9	-22.6	27
RTRC	25.4	-22.2	17
WIKIPEDIA	21.9	2.3	24
SPRINGER	20.1	14.1	23
OASIS	19.5	-14.2	20
WILEY	Invalid	-14.7	20
IOP	18.6	25.3	21
NATURE	18.2	22.6	18
OXFORD	17.1	10.1	18
PROQUEST	Invalid	3.36	20
HINDAWI	Invalid	24.9	17
CC	Invalid	23.9	27
ELSEVIER	20.7	-5.3	13
BMJ	18.2	21.6	13
BOAI	17.1	36.1	17
PLOS	18.4	26.1	9
SPARC	16	-3.21	16
MIT	17.2	37.9	16
U. OF	16.4	17.4	14

KANSAS			
BRILL	14.1	22.1	16
PETER SUBER	16.2	39.9	16
CAMBRIDGE	14.2	44.2	14
HARVARD U.	14.8	44.7	14
T & F	14.4	45.4	14
SAGE	14.2	47.2	13
OAWG	14.5	47.6	13

One of our initial hypotheses was that there would a discernible difference in complexity between source group type, such as between publishers and non-publishers. The results of these tests, however, suggest that readability is a consistent problem for all of those providing information about open access publishing. In fact, of the top ten groups with the "most complex" Web pages, just four are publishers.

Given the strong link between complexity and cognitive load that we previously described, it was pertinent to explore further how the language used in the content of the sample Web pages might increase task complexity by presenting conflicting outcomes or introducing multiple paths to task completion. Qualitative analysis lends itself better to a deeper investigation of language distributed about open access publishing and helps to describe the relationship between complexity and cognitive load. In the following section, we explore how definitions of open access offer conflicting outcomes for open access publications which can increase the perceptual strain and cognitive load of authors seeking an open access route to publication.

Table 3: Link counts, word counts, and sentence counts of sample three Web pages by source group.

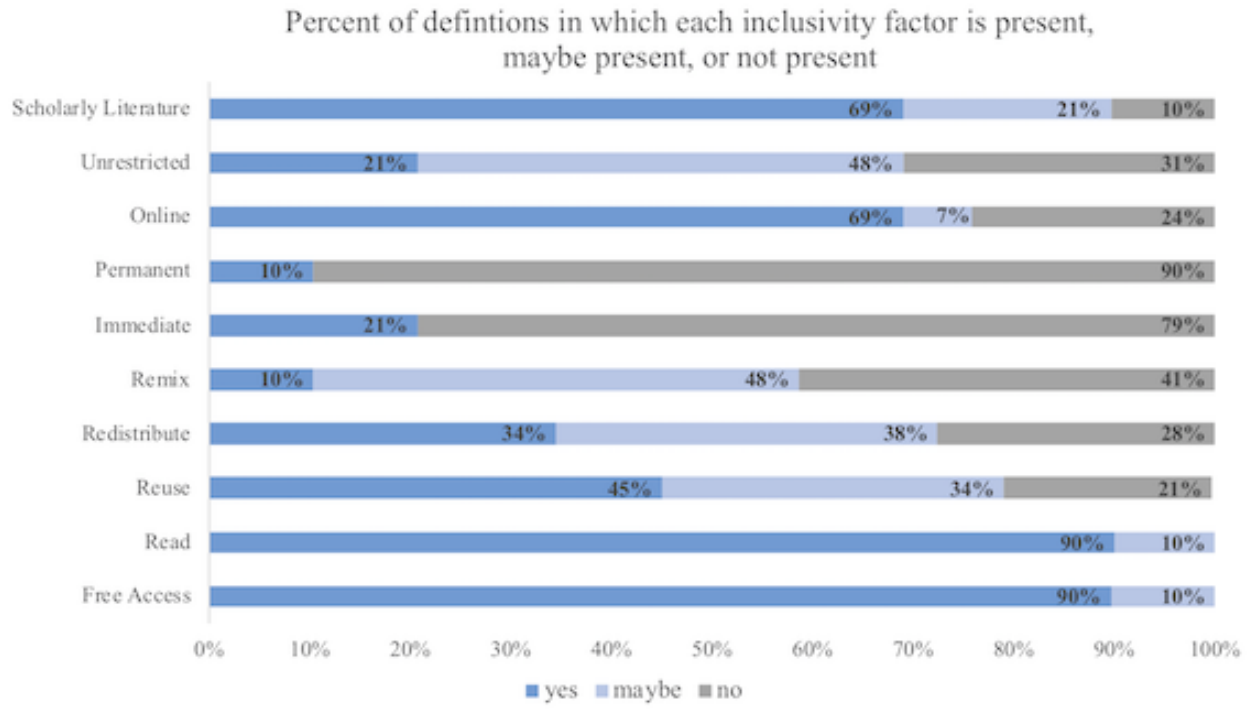
Source	Link count	Source	Word count	Source	Sentence count
PROQUEST	2	CC	544	PROQUEST	15
HINDAWI	7	PROQUEST	631	DEGRUYTER	16
CC	10	DEGRUYTER	734	CC	18
SPARC	24	HINDAWI	818	WILEY	20
WILEY	26	WILEY	1031	HINDAWI	28
BMJ	27	CAMBRIDGE	1586	CORNELL U.	39
OAWG	38	PLOS	1788	PLOS	40
U. OF KANSAS	42	BMJ	1794	BMJ	44
IOP	44	U. OF KANSAS	1828	IOP	49
CAMBRIDGE	49	SPARC	1945	U. OF KANSAS	50
PLOS	50	OAWG	1953	ACRL	52
BRILL	51	CORNELL U.	2216	NATURE	56

SAGE	56	BRILL	2219	CAMBRIDGE	59
SPRINGER	64	IOP	2228	SPARC	61
NATURE	69	NATURE	2229	OASIS	63
SHERPA	74	OASIS	2690	OXFORD	81
OASIS	76	OXFORD	2872	OAWG	83
ELSEVIER	91	SAGE	2921	SPRINGER	86
ACRL	93	ACRL	3085	SHERPA	90
CORNELL U.	93	SHERPA	3841	BRILL	94
DEGRUYTER	122	SPRINGER	4147	RTRC	94
OXFORD	148	WIKIPEDIA	5205	WIKIPEDIA	101
SUBER	150	MIT	5761	ELSEVIER	111
MIT	153	SUBER	6029	SAGE	122
RTRC	188	ELSEVIER	6547	MIT	174
BOAI	199	RTRC	7110	SUBER	194
T & F	211	HARVARD U.	8202	BOAI	255
HARVARD U.	331	BOAI	8889	HARVARD U.	311
WIKIPEDIA	416	T & F	9641	T & F	375

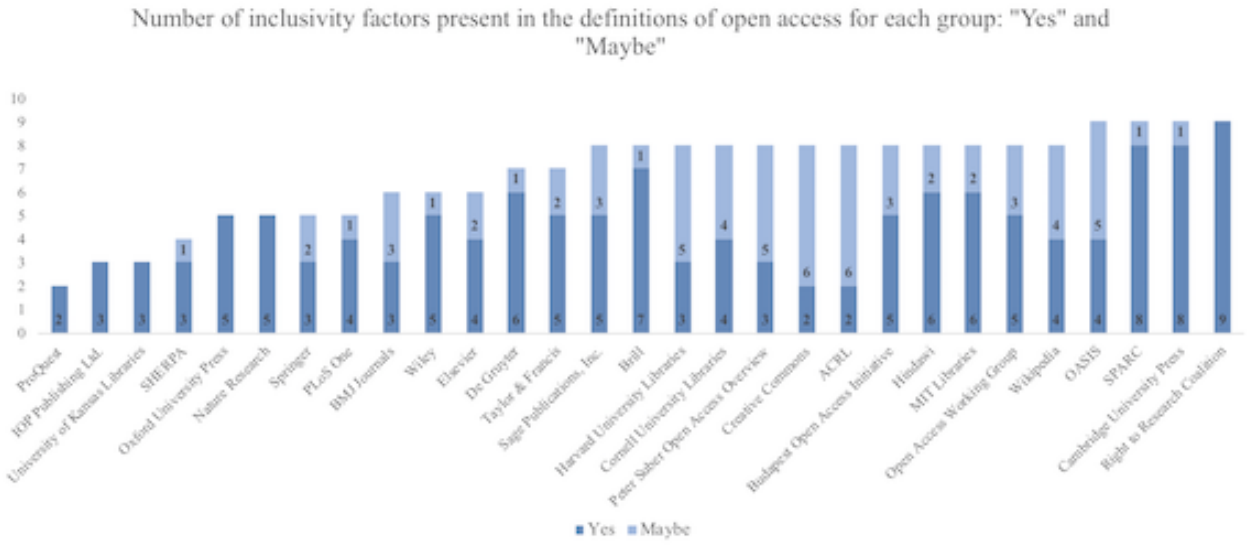
Cognitive load

While all of the texts examined for this study are somewhat-to-extremely complex based on various quantitative readability measures covered in the previous section, the readability of the text may in fact come second to overall comprehension of textual content. Successful task completion relies in part on content being consistent and not presenting outcomes that conflict with one another. In order to assess the consistency of outcomes with open access publishing in the simplest of terms, we conducted a qualitative analysis of the definitions of “open access” provided by each source group. As stated in the Methods section, it was a condition for selection into the study sample that each source group provide a clearly stated definition of open access on their respective Web sites. Definitions are an important component of these texts because they signal to the reader the conditions and results of open access publication.

Definitions were coded according to ten selected “factors” — Free Access, Read, Reuse, Redistribute, Remix, Immediate, Permanent, Online, Unrestricted, and Scholarly Literature. If a definition definitely included a factor, it was coded as “Yes” for that factor. If the definition could be interpreted as including that factor, but did not specifically call out that factor, it was coded as “Maybe.” If the definition definitely did not include that factor, it was coded as “No.”



Note: Larger version available here.



Note: Larger version available here.

The results of this coding exhibit wide disagreement across source groups for the implications of open access publications. Nearly all of the definitions (90 percent) were in strong agreement that open access literature was free to access and to read, and more than two-thirds (69 percent) also agreed that this literature was online and scholarly in nature (Figure 1). There was no strong agreement across definitions regarding the other six inclusivity factors, however. These remaining factors are primarily those that dictate how others can reuse or adapt the works as well as the timing and duration of open access. This lack of clarity about what constitutes an open access publication leaves many in the dark about whether, for example, an article with an embargo can

count as an open access publication, or how they and others might be able to share or modify these works.

There is also a significant variability in the count of factors present in each source group's definition (Figure 2). More than half (*n*=16, 55 percent) of the source groups definitely or maybe include at least eight of the factors. Of these 16 groups, only four (25 percent) are publishers. Of the 13 remaining groups that definitely or maybe contained seven or fewer factors, most of these (*n*=11, 85 percent) are publishers. A particularly notable characteristic of these data are those that were coded as "Maybe" including some factors of a definition. Non-publisher groups were more than twice as often coded as "Maybe" containing a factor, rather than with a definitive "Yes" or "No." When coding for only "Yes", or what definitions definitely contained the factors, there is no clear pattern in which groups' definitions are more or less inclusive, as both publishers and non-publishers equally present restrictive or inclusive definitions of open access. More importantly, however, when including factors coded as "Maybe" into the overall count of factors, it is then that non-publisher definitions become much more inclusive and publishers appear as much more restrictive. However, that inclusivity is problematic because of the conditional nature of the definitions.

We posit that non-publishers may appear to have more inclusive definitions due to the overwhelming presence of what we term "squishy language." For example, non-publisher definitions often include phrases governing re-use by describing open access as "without most permission barriers" or "without severe restrictions on use." While these may be accurate statements, it is up to reader interpretation to determine what would not count as a "severe restriction" or what is not part of "most permission barriers." While publishers may have a higher obligation to set concrete terms, non-publishers have no such obligations and may use this "squishy language" in order to leave broader room for interpretation and expansion of the possibilities under open access. One cannot fault publishers nor non-publishers for these tendencies, but they undoubtedly reveal inconsistent and even conflicting outcomes for readers and authors. It is entirely possible that an author hoping to publish open access could receive a definition from a library or advocacy group that they only find out later is in conflict with the stated terms of their preferred publisher.

In addition to the presentation of conflicting information, cognition is also strained by the existence of multiple paths of achieving a specific task [18]. In an effort to further explore how many paths contribute to ineffective texts, we conducted two additional qualitative analyses on language about "choice" and alternative cost recovery models described in Web site text about open access publishing. The following section explores how this language creates complex tiers of choice that may have the undesirable effect of steering authors away from perceived "new" or "novel" choices about publishing and back into the arms of the status quo.

Status quo bias

Options increase a tendency towards the status quo and the presentation of multiple paths or alternatives also reduce task completion through cognitive load. Our next question, then, was how are choice and alternatives presented to those seeking information about open access publishing? In order to explore this question we identified many ways in which options and alternatives might be presented through information sharing about open access publishing. Some of these include: various models of open access itself (green, gold, hybrid, etc.), options and policies governing the timing and venue of self-archiving, selection of Creative Commons license, methods of distribution and dissemination, pay versus no-pay models, and for models using article processing charges (APCs) the various alternatives to "author pays." In advance of all of these options, however, is the choice of open access itself, upon which all of the other choices follow. Open access publishing, then, consists of complex tiers of decision-making throughout which numerous pathways unfold from the beginning to the very end of the lifecycle of a publication.

Within such a complex landscape of choice, we decided to conduct qualitative analyses on just three of the cases listed above: the presence of choice language itself as related to open access publishing, pay models, and alternative cost recovery models. A full accounting of all of the tiers of choice involved in open access publishing was out of scope for this project, which leaves ample room for future inquiry.

In the first analysis, the three Web page samples from each source group were coded for the presence of the words "choice" and "option" as well as the derivatives of those words. Occasionally these words were used in contexts other than what we were coding for, and these instances were not counted towards our analysis. In other words, we counted text that explicitly advanced the concept of open access as an author choice. For example, consider this excerpt from Cambridge University Press: "The majority of our journals offer the optional open access option." On the other hand, consider this excerpt from SPARC: "Understanding the effect of fully exercising the rights you have as an author can help you make educated choices about the publishing outlets

you choose to submit work to." The more general reference to overall publishing choices by SPARC was not counted as promoting open access through the lens of choice. One source group, the University of Kansas Libraries, used choice language but was unequivocally negative, so these instances were not counted (*e.g.* "Some traditional closed-access publishers have open access (for a fee) options. We call this "hybrid open access" and generally don't recommend it"). It is worth noting that two publishers include the language of choice in the name of the hybrid program itself — Springer Open Choice and SAGE Choice — which accentuates these programs as "different" or "other."

For the second and third analysis, language relating to the costs of open access publishing was pulled from the three sample Web pages, and that smaller sample of language was coded for mention of whether there is "always," "usually," "sometimes," or "never" payment expected for open access as well as what payment models other than "author pays" are mentioned.

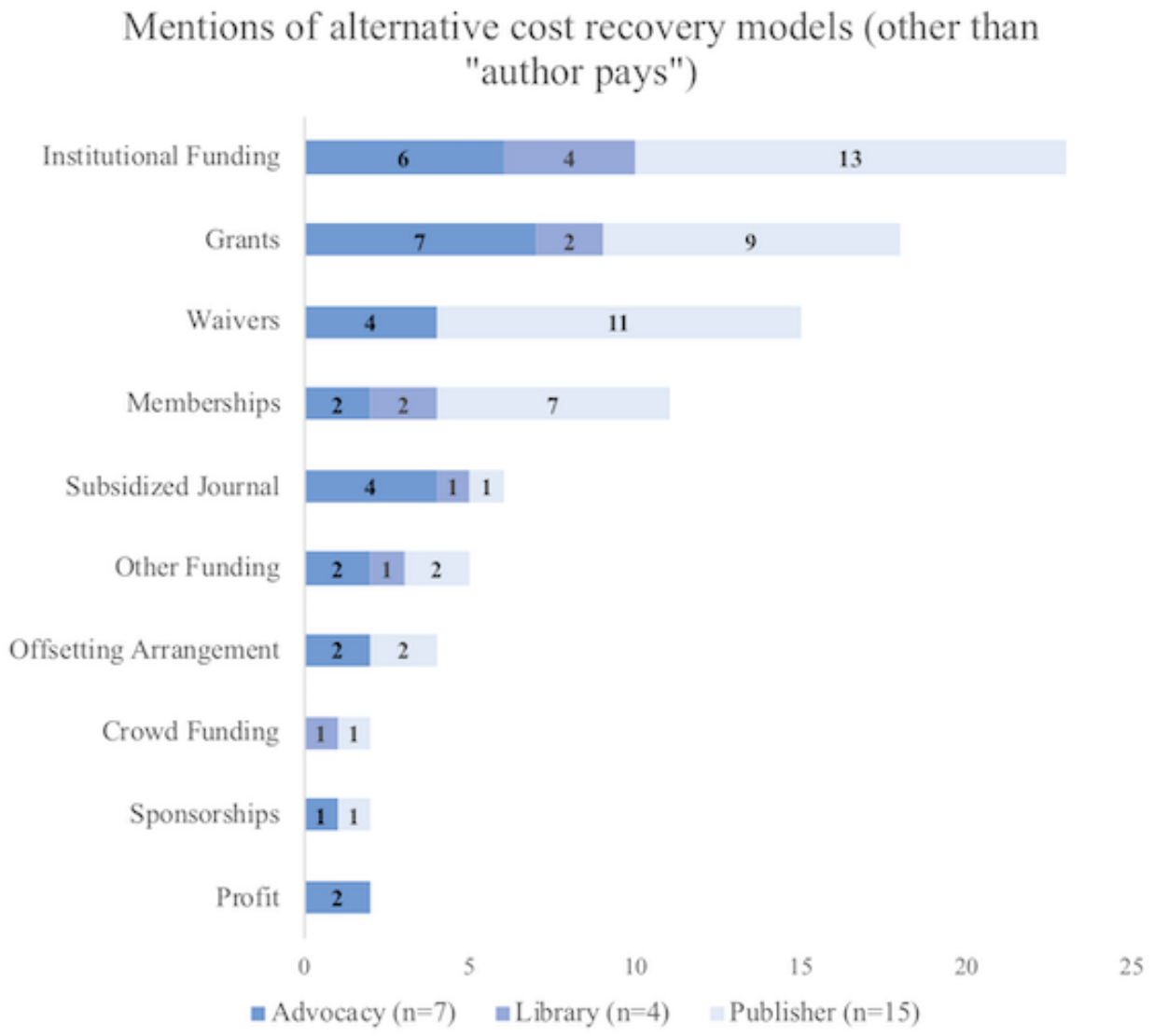
Over two-thirds (69 percent) of the source groups for this study explicitly discussed open access publishing in the sample Web pages as a "choice" or "option." Choice language was largely present on publisher Web pages, but there are several instances of other source groups employing this open-ended language. Notably, the only two publishers that do not employ this language (Hindawi and PLoS) are already open access by default, meaning that they do not require their authors to make a choice about publishing model after choosing the journal, much like subscription publishers that do not offer open access choices.

Open access as a choice was also often compounded with additional choices. One of the most common compounded choices across all source groups was how to pay for open access publishing services. The interaction of these two layers of choice is shown in Table 4. When looking at cost recovery models for open access payments, we coded 10 different methods that may be used to cover the cost of payment other than the author paying "out-of-pocket" (Figure 3). The most common cost recovery model mentioned was institutional funding, which can include funding from university libraries or departmental funding; this method was mentioned by nearly four-fifths (79 percent) of all source groups. Source groups averaged specifying 3.4 payment models other than "author pays" as means of fulfilling fees for open access publication.

Table 4: A compounded choice [19].

	Author always pays	**Author usually pays**	**Author sometimes pays**
Choice language used	OXFORD NATURE WILEY BMJ SAGE PROQUEST	SPRINGER CAMBRIDGE T & F BRILL CORNELL U.	ELSEVIER IOP DEGRUYTER RTRC MIT OASIS WIKIPEDIA HARVARD U. BOAI
Choice language not used	HINDAWI SHERPA PLOS	ACRL	PETER SUBER U. OF KANSAS

Intricate and often contradictory information about the presence, frequency, and source of payments accumulates layers of choice that steer decision-makers towards maintaining the status quo. Furthermore, by presenting open access through multiple lenses of choice, subscription publishing is often coded as the default option, or in other words, the status quo. Eidelman and Crandall (2012) observe that "By simply labeling a stimulus or a process as the status quo, people may assume, quite rationally, that the stimulus or process has passed the inspection of other people. Social norms are often a good guide to correct judgment and conduct; it is reasonable to use the opinions of others as the basis of our own" [20]. Rather than encouraging authors to be more open to new publication models, then, the complex framing of choice around open access may serve to justify, normalize, and further reinforce the predominance of subscription publishing.



Note: Larger version available here.

Discussion

In our view, the analysis in this paper presents a particular challenge to so-called hybrid open access publishing, in which subscription journals present open access as a "choice" from the outset. From that initial choice, if an author makes the decision to publish open access, many further choices are needed — what license to choose, and how to pay for a publication fee (which, in itself, may require more choices and more information gathering) — therefore contributing exponentially to a high cognitive load. While someone unfamiliar with open access may initially pursue it because it is a good way to make their work more accessible, when faced with the many extra tasks and requirements associated that author may turn back to the status quo. This phenomenon may account for the reported low uptake of open access in hybrid journals (28 percent according to Kocher and Kelly, 2016). However, gold and green open access may similarly hinder uptake by framing these routes as "alternatives" or "additions" to "traditional" subscription publishing, since loss aversion and regret avoidance at the outset discourages the pursuit of paths characterized as introducing change or requiring additional effort (Eidelman and Crandall, 2012).

Viewing our analysis through a lens of information literacy, we believe that our findings are consistent with the observed issues surrounding the shift of the information landscape from a scarcity of information to overabundance (Head, 2013), and that information seekers may be "lost in a thicket of information overload" (Head, 2017). Especially as sources with an overabundance of information have experienced a shift from truth to consensus (Garfinkel, 2008), when there is no consensus to draw from as has been shown in this study, readers and users can become lost in a web of conflicting information.

We propose that the most effective way to increase the uptake of open access is to make it the default publishing option. Due to the significant effects of complexity and cognitive load on entrenching the status quo bias, it is essential that information about publishing and publishing routes themselves become simpler rather than more complicated. In the study "Do defaults save lives?" Johnson and Goldstein (2003) found that despite organ donation being favored by the majority in the United States, opt-in conditions made donation rates very low. However, when the default status of organ donation was tested in an online experiment, the opt-out donation agreements were double those of the opt-in. While organ donation is certainly a more high-risk proposition than how to publish an academic paper, the impact of the default in this potentially life-altering decision implies that less-risky situations (like publishing) may also see similar results. Defaults, then, can serve as a powerful mechanism for not only normalizing established practices but encouraging participation by justifying the existence of that default and reducing extraneous barriers. In fact, we have seen this method be effective when open access policies mandate open availability of the research publications from institutions, funders, or other organizations — and the effectiveness in increasing open access sharing is correlated to the strength of the mandate (Gargouri, *et al.*, 2012).

It is true that the open access movement has always been in pursuit of the default in order to make open access publishing basic and ubiquitous. What we propose then is of course easier said than done. Nonetheless, we suggest that much of the current framing of open access by publishers, advocates, and others is not conducive to a shift to open access as the default because it discourages participants from advancing away from the status quo. If we want a more open future, it's imperative to turn the mirror onto our discourse and practices surrounding open access and consider the most productive ways of mitigating the effects of complexity, cognitive load, and especially the status quo bias.

We do not intend to suggest that open access should be pursued blindly at the expense of other scholarly and social values, however. What we would like to bring attention to is the fact that business model agnosticism in open access publishing, whether it be practiced by publishers or accepted by advocacy groups and libraries, will cause uncertainty and strain on authors that has the potential to trigger the status quo bias and obscure the merits of one model over another. Business agnosticism favors a preponderance of choice and ambiguous gestures toward values that cannot possibly be activated by all of the business models being accommodated. Equitable and sustainable models of scholarly publishing make an active decision to build in and prioritize those values (*i.e.*, make them the default) rather than to keep a finger in every slice of the open pie. In the guise of offering author choice, agnostic business models (and those that support them at large) risk stifling the aim of the open access movement to transform the values of scholarly communication.

Next steps and conclusions

A first step towards being more attentive to these effects is to pay close attention to both the complexity and the framing of language. This means creating simple pathways to open access and explaining them in plain language. While it may be counterintuitive, a proliferation of choices does not encourage uptake of newer options and in fact has a stifling effect as participants find themselves confounded and overwhelmed. It may also be productive to re-think words such as "option" and "alternative" used in association with open access and words like "traditional" and "standard" used in association with subscription publishing, because the latter framing connotes something that is well-established and thus superior.

A second step that could dramatically reduce the status quo bias is to present open access as the default form of publication, with a simple track that has minimal choice in order to reduce the burden of further action typically imposed on the author. Subscription publishing should serve as the alternative choice in this system. Open access publication is the only option in major publications such as the Public Library of Science (PLOS), Open Library of Humanities and their associated journals, and many other open access publishers. However, in "hybrid" journal situations where authors must choose between subscription publishing and open access

publishing, it is still by and large subscription publishing that is presented as the default and easiest choice for authors. What would happen if open access publishing were presented as the default in hybrid publishing environments? This would create a shift towards an opt-out rather than an opt-in model of open access publishing; with this framing, would uptake of open access publishing in hybrid journals increase? Based on the literature regarding cognitive load and status quo bias, we suggest that uptake of open access publishing in hybrid journals would increase in these situations, potentially allowing for a gradual phasing out of subscription publishing. However, in the absence of hybrid publishers providing open access publishing as the default, universities and research institutions could apply pressure themselves on the status quo in favor of change by building in open access as an expectation for promotion and tenure. Either of these routes towards open access as the default would reduce the burden of the status quo bias on individual researchers by providing legitimacy for open access at the structural or system level.

This research was conceived of as a first and undoubtedly small-scale step into a wider area of inquiry on the discourse about scholarly communication. One obvious future step for this research is the execution of an empirical study that could test on subjects whether or not language complexity and framing influences decision-making about publication. While related studies mentioned in this paper have found textual characteristics to strongly influence comprehension and cognitive load of online information, only an empirical study of human interaction with texts on open access in particular can draw definitive conclusions about the theoretical findings that are presented here.

There are also abundant areas to extend analysis of the texts themselves, some of which were considered for this study but did not fit within the scope of the present research. Simply expanding the sample used for analysis, whether by including a more diverse set of groups or by analyzing more than three Web pages per source group, would increase the reliability and accuracy of these findings. Another potential avenue of inquiry might include an examination of how structural characteristics of the Web sites, such as layout and navigation, either enable or inhibit use. Future research could also dig deeper into the problem of choice by documenting the tiers of decision-making that make up the "open access decision tree."

There is no shortage of avenues for future inquiry related to this work. We hope that this snapshot of an issue that has been thus far under-examined and potentially undervalued sheds light on the importance of framing to future directions of open access publishing. We cannot assume that the advantages of open access will be self-evident or irresistible, nor can we assume that all variants of open access are created equally. A default, values-driven open access future is not one in which open access is depicted as a revolutionary or even a novel act, but one in which it is recognized as a fundamental and enduring framework for scholarly publishing. FM

About the authors

Melissa H. Cantrell is the Scholarly Communication Librarian at the University of Colorado Boulder.
Direct comments to: melissa [dot] cantrell [at] colorado [dot] edu

Lauren B. Collister is the Director of the Office of Scholarly Communications & Publishing at the University of Pittsburgh.
E-mail: lbcollister [at] pitt [dot] edu

Acknowledgments

The authors are grateful to Chealsye Bowley for recognizing a common area of interest and connecting the co-authors of this paper. Thanks also to Na-Rae Han of the University of Pittsburgh for her assistance with Python, Textstat, and Jupyter Notebooks. Melissa H. Cantrell would like to thank Katie Mika, Nickoal Eichmann-Kalwara, and Andrew Johnson for their input and guidance, particularly in the early stages of this project. Lauren B. Collister sends her gratitude to Melinda Ciccocioppo for conversations about cognitive load and Allison Head for her comments and advice near the end of the project.

Notes

1. We made every effort in this article to re-frame our own word choices in reference to open access and subscription publishing. A document detailing common words/phrases and suggested

ways to re-frame them can be found here: https://docs.google.com/document/d/1ef-of37jAQABAWWfLV-r8qtKo9tqdUqqWop8J7au-uA/edit?usp=sharing.

2. Eidelman and Crandall, 2012, p. 270.

3. Kay and Zanna, 2009, p. 158.

4. Hagedoorn and Hesen, 2009, p. 826.

5. Eidelman and Crandall, 2012, p. 271.

6. Kempf and Ruenzi, 2006, p. 208.

7. Kher, *et al.*, 2017, p. 1.

8. U.S. Department of Health and Human Services (USDHHS), 2008, p. 1,311.

9. U.S. Department of Health and Human Services (USDHHS), 2008, p. 1,314.

10. Kher, *et al.*, 2017, p. 2.

11. Luger, *et al.*, 2013, p. 2,691.

12. Wang, *et al.*, 2014, p. 1.

13. Wang, *et al.*, 2014, pp. 1–2.

14. Links may not be stable. Prior to the completion of data analysis, URLs selected for both Taylor & Francis and OASIS changed, and the Wayback Machine was used to recover text originally selected in early 2018. Other URLs and/or text may alter at any time.

15. https://github.com/parnopaeus/oalanguage/blob/master/datasources_oalanguage.csv.

16. https://pypi.org/project/textstat/.

17. https://github.com/parnopaeus/oalanguage.

18. Wang, *et al.*, 2014, p. 2.

19. Two of the source groups — Creative Commons and the Open Access Working Group — neither use choice language nor mention paying fees for open access publishing. One source group — SPARC — was the only source group in which choice language was not used and "never" paying was explicitly mentioned ("virtually no marginal costs").

20. Eidelman and Crandall, 2012, p. 275.

References

Sarah R. Allred, L. Elizabeth Crawford, Sean Duffy, and John Smith, 2016. "Working memory and spatial judgments: Cognitive load increases the central tendency bias," *Psychonomic Bulletin & Review*, volume 23, number 6, pp. 1,825–1,831.
doi: https://doi.org/10.3758/s13423-016-1039-0, accessed 9 November 2018.

Scott Eidelman and Christian S. Crandall, 2012. "Bias in favor of the status quo," *Social and Personality Psychology Compass*, volume 6, number 3, pp. 270–281.
doi: https://doi.org/10.1111/j.1751-9004.2012.00427.x, accessed 9 November 2018.

Simson Garfinkel, 2008. "Wikipedia and the meaning of truth," *MIT Technology Review* (20 October), at https://www.technologyreview.com/s/411041/wikipedia-and-the-meaning-of-truth/, accessed 17 April 2019.

Yassine Gargouri, Vincent Lariviere, Yves Gingras, Tim Brody, Les Carr, and Stevan Harnad, 2012. "Testing the Finch Hypothesis on green OA mandate ineffectiveness," *arXiv* (2 November), at http://arxiv.org/abs/1210.8174, accessed 22 May 2019.

John Hagedoorn and Geerte Hesen, 2009. "Contractual complexity and the cognitive load of R&D alliance contracts," *Journal of Empirical Legal Studies*, volume 6, number 4, pp. 818–847.
doi: https://doi.org/10.1111/j.1740-1461.2009.01161.x, accessed 9 November 2018.

Alison Head, 2017. "How today's students conduct research" (14 March), at https://www.projectinfolit.org/university-of-nebraska-libraries.html, accessed 17 April 2019.

Alison J. Head, 2013. "Learning the ropes: How freshmen conduct course research once they enter college," *Project Information Literacy* (5 December), at https://www.projectinfolit.org/uploads/2/7/5/4/27541717/pil_2013_freshmenstudy_fullreportv2.pdf, accessed 21 March 2019.

Eric J. Johnson and Daniel Goldstein, 2003. "Do defaults save lives?" *Science*, volume 302, number 5649 (21 November), pp. 1,338–1,339.
doi: https://doi.org/10.1126/science.1091721, accessed 9 November 2018.

Aaron C. Kay and Mark P. Zanna, 2009. "A contextual analysis of the system justification motive and its societal consequences," In: John T. Jost, Aaron C. Kay, and Hulda Thorisdottir (editors). *Social and psychological bases of ideology and system justification*. New York: Oxford University Press, pp. 158–171.
doi: http://dx.doi.org/10.1093/acprof:oso/9780195320916.003.007, accessed 9 November 2018.

Alexander Kempf and Stefan Ruenzi, 2006. "Status quo bias and the number of alternatives: An empirical illustration from the mutual fund industry," *Journal of Behavioral Finance*, volume 7, number 4, pp. 204–213.
doi: https://doi.org/10.1207/s15427579jpfm0704_3, accessed 9 November 2018.

Akhil Kher, Sandra Johnson, and Robert Griffith, 2017. "Readability assessment of online patient education material on congestive heart failure," *Advances in Preventive Medicine*, volume 2017, article ID 9780317.
doi: https://doi.org/10.1155/2017/9780317, accessed 26 February 2019.

Megan Kocher and Julie Kelly, 2016. "Use of the paid open access option in hybrid open access journals in agriculture: A mixed-methods study," *Issues in Science & Technology Librarianship*, at http://www.istl.org/16-fall/refereed2.html, accessed 22 March 2019.
doi: https://doi.org/10.5062/F47P8WDB, accessed 26 February 2019.

Ewa Luger, Stuart Moran, and Tom Rodden, 2013. "Consent for all: Revealing the hidden complexity of terms and conditions," *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2,687–2,696.
doi: https://doi.org/10.1145/2470654.2481371, accessed 28 November 2018.

Erin C. McKiernan, 2014. "Culture change in academia: Making sharing the new norm," at https://core.ac.uk/download/pdf/39962563.pdf, accessed 24 April 2019.

Claire Redhead, 2019. "OASPA response to European Commission Expert Group report," *Open Access Scholarly Publishers Association* (20 May), at https://oaspa.org/oaspa-response-european-commission-expert-group-report/, accessed 31 May 2019.

Jonathan P. Tennant, François Waldner, Damien C. Jacques, Paola Masuzzo, Lauren B. Collister, and Chris. H.J. Hartgerink, 2016. "The academic, economic and societal impacts of open access: An evidence-based review," *F1000Research*, volume 5, at https://f1000research.com/articles/5-632, accessed 22 May 2019.

Tiffany M. Walsh and Teresa A. Volsko, 2008. "Readability assessment of Internet-based consumer health information," *Respiratory Care*, volume 53, number 10, pp. 1,310–1,315, and at http://rc.rcjournal.com/content/53/10/1310, accessed 26 February 2019.

Qiuzhen Wang, Sa Yang, Manlu Liu, Zike Cao, and Qingguo Ma, 2014. "An eye-tracking study of website complexity from cognitive load perspective," *Decision Support Systems*, volume 62, pp. 1–10.
doi: https://doi.org/10.1016/j.dss.2014.02.007, accessed 10 November 2018.

Paul Whitney, Christa A. Rinehart, and John M. Hinson, 2008. "Framing effects under cognitive load: The role of working memory in risky decisions," *Psychonomic Bulletin & Review*, volume 15, number 6, pp. 1,179–1,184.
doi: https://doi.org/10.3758/PBR.15.6.1179, accessed 9 November 2018.

Editorial history

Received 3 May 2019; revised 3 June 2019; accepted 21 June 2019.

The status quo bias and the uptake of open access
by Melissa H. Cantrell and Lauren B. Collister.
First Monday, Volume 24, Number 7 - 1 July 2019
https://firstmonday.org/ojs/index.php/fm/rt/printerFriendly/10089/8053
doi: http://dx.doi.org/10.5210/fm.v24i7.10089