JLSC

ISSN 2162-3309 | JLSC is published by the Pacific University Libraries | http://jlsc-pub.org

Volume 6, General Issue (2018)

Confused about Copyright? Assessing Researchers' Comprehension of Copyright Transfer Agreements

Alexandra Kohn & Jessica Lange

Kohn, A. & Lange, J. (2018). Confused about Copyright? Assessing Researchers' Comprehension of Copyright Transfer Agreements *Journal of Librarianship and Scholarly Communication,* 6(General Issue), eP2253. https://doi.org/10.7710/2162-3309.2253

This article underwent fully-anonymous peer review in accordance with JLSC's peer review policy.

ISSN 2162-3309 10.7710/2162-3309.2253

RESEARCH

Confused about Copyright? Assessing Researchers' Comprehension of Copyright Transfer Agreements

Alexandra Kohn
Head, Office of Copyright Compliance, McGill University
Jessica Lange
Scholarly Communications Librarian, McGill University

INTRODUCTION Academic authors' confusion about copyright and publisher policy is often cited as a challenge to their effective sharing of their own published research, from having a chilling effect on self-archiving in institutional and subject repositories, to leading to the posting of versions of articles on social networking sites in contravention of publisher policy and beyond. This study seeks to determine the extent to which authors understand the terms of these policies as expressed in publishers' copyright transfer agreements (CTAs), taking into account such factors as the authors' disciplines and publishing experience, as well as the wording and structure of these agreements. **METHODS** We distributed an online survey experiment to corresponding authors of academic research articles indexed in the Scopus database. Participants were randomly assigned to read one of two copyright transfer agreements and were subsequently asked to answer a series of questions about these agreements to determine their level of comprehension. The survey was sent to 3,154 participants, with 122 responding, representing a 4% response rate. Basic demographic information as well as information about participants' previous publishing experience was also collected. We analyzed the survey data using Ordinary Least Squared (OLS) regressions and probit regressions. **RESULTS AND DISCUSSION** Participants demonstrated a low rate of understanding of the terms of the CTAs they were asked to read. Participants averaged a score of 33% on the survey, indicating a low comprehension level of author rights. This figure did not vary significantly, regardless of the respondents' discipline, time in academia, level of experience with publishing, or whether or not they had published previously with the publisher whose CTA they were administered. Results also indicated that participants did equally poorly on the survey regardless of which of the two CTAs they received. However, academic authors do appear to have a greater chance of understanding a CTA when a specific activity is explicitly outlined in the text of the agreement.

Received: 04/06/2018 Accepted: 10/15/2018

Correspondence: Alexandra Kohn, McLennan Library Building, 3459 rue McTavish, Montreal, Quebec, H3A 0C9, alexandra.kohn@mcgill.ca

IMPLICATIONS FOR PRACTICE

1. Given the findings of a low level of author comprehension of CTAs, librarians who are involved in author rights education and advocacy should focus on providing academic authors with the tools to parse the language and phrasing of CTAs and to isolate and understand the conditions and limitations around the rights and uses that are most important to the authors. Furthermore, librarians should partner with other relevant departments and offices at their institutions (e.g., research offices, legal offices, faculty councils, promotion and tenure committees, graduate student associations, etc.) to provide relevant information and coaching on understanding author agreements.
2. Academic authors experience difficulty understanding the terms of their CTA regardless of demographic factors such as time in academia, seniority, experience publishing, and discipline. As such, author rights and copyright education and advocacy should be designed for wide dissemination, without focus on a particular demographic within the academy.
3. Results indicate that authors are more likely to understand the rights and obligations stipulated in CTAs when permissible/non-permissible activities are mentioned explicitly. Librarians should advocate for this type of explicit language in CTAs wherever they have influence on the construction of CTAs, be that as members of editorial boards, providers of education and infrastructure through library publishing initiatives, or through scholarly communication education and outreach.

INTRODUCTION

Publishers' conditions and restrictions on academic authors' own published work have a significant effect on academic activities pivotal to teaching, research, publication, and scholarly exchange with colleagues and collaborators. For example, funder and institutional mandates increasingly require researchers to make their journal articles open access within a given time frame from acceptance or publication (ROARMAP, n.d.; SHERPA Juliet, n.d.; Suber, 2012); depositing in institutional and subject repositories (green open access) is often the easiest and most affordable way to achieve this end. The complexity of the restrictions imposed by publishers means that it is no longer sufficient for authors to know if they retain copyright or even if the publisher allows self-archiving. Authors must now be aware of the minute details of the CTA, for example, which version of the paper may be self-archived and where, how the publisher should be acknowledged in any reproduction, under which conditions authors may make the article available to their students, and which methods of sharing the published version with colleagues are acceptable.

The growth of the open access movement and rise in institutional and funder open access mandates discussed above have resulted in more publishers explicitly permitting some variant of self-archiving (Archambault et al., 2014; Gadd & Troll Covey, 2016; Laakso, 2014). However, in tandem with this rise, the number of conditions under which self-archiving is allowed have proliferated at a staggering rate. According to one study, between 2004 and 2015, "the volume of restrictions around *how, where* and *when* self-archiving may take place has increased 119%, 190% and 1000% respectively" (Gadd & Troll Covey, 2016, p. 1).

By many accounts, academic authors experience a great deal of difficulty navigating these complexities. The literature consistently points to authors' self-reported confusion and anxiety around copyright, publisher policies, and CTAs as a major barrier to self-archiving (Cullen & Chawner, 2011; Frass, Cross, & Gardner, 2014; Kim, 2011). In addition, recent studies suggest that academic authors are active and frequent participants in the posting and distributing of their articles on academic social networking sites in contravention of CTAs they have signed. An analysis of a random sample of papers uploaded to ResearchGate found that just over half the non–open access papers analyzed were posted to the site in violation of publisher policy (Jamali, 2017). In the vast majority of these cases of infringement, the authors would have been permitted to deposit a postprint but had instead deposited the publisher's PDF (which was not permissible according to the publisher's policy). A recent survey of University of Rhode Island faculty corroborates this assumption, with over 60% of respondents unsure of the legality or compliance with publisher policy of posting their articles in ResearchGate (Lovett, Rathemacher, Boukari, & Lang, 2017). This practice, sometimes dubbed black open access (Björk, 2017), has increasingly attracted publishers' notice, resulting in their issuing takedown notices and filing lawsuits in an attempt to combat this practice (Chawla, 2017; Solon, 2013; Van Noorden, 2017).

This study seeks to move beyond academics' self-reported confusion and trepidation around CTAs to empirically test author comprehension of these agreements. We hypothesized that scholars with more publishing experience would demonstrate greater comprehension of CTAs. Furthermore, we hypothesized that authors who received a CTA from a publisher with whom they had previously published would similarly demonstrate greater comprehension than those who not worked with that publisher previously. We also expected that shorter agreements (calculated by word count) would have a positive effect on author comprehension, as this would improve the likelihood that the participants would read the agreement in full and circumscribe the amount of text that respondents would have to search to determine whether a condition of use is stated in the CTA.

By investigating the level and determinants of author comprehension, we hope to empiri-

cally address the extent to which CTAs are barriers to self-archiving and other author rights behaviors. This article will also discuss proactive behaviors librarians and publishers can take to improve author comprehension and engagement with their publishing agreements.

LITERATURE REVIEW

Surveys to Date

To the best of our knowledge, no published surveys have been administered focusing exclusively on author understanding of CTAs. However, there is a sizeable corpus of surveys that pose questions related to copyright assignment and licensing in the context of open access or toward self-archiving practices in particular. Such surveys have been conducted at institutional, disciplinary, and national levels, typically focusing on authors' attitudes, awareness, and self-reported actions concerning CTAs. Questions posed that are germane to the current study included whether respondents read and/or modified their CTAs and whether a funder's open access policy had made them examine the terms of those agreements more closely (Charbonneau & McGlone, 2013); the extent to which publishers' copyright terms affected participants' selection of publication venue; whether authors engaged in or were willing to engage in negotiation of CTAs; and whether participants felt that CTAs limited their ability to self-archive or disseminate their articles for teaching, research, or publication purposes (Moore, 2011; Odell, Dill, & Palmer, 2014; University of California Office of Scholarly Communication, California Digital Library eScholarship Program, & Greenhouse Associates, Inc., 2007). A number of surveys posed questions about the effect of CTAs and copyright concerns on participants' willingness to self-archive (Austin, Heffernan, & David, 2008; Cullen & Chawner, 2011; Kim, 2011). In addition, a 2006 survey of academics in the health sciences (Smith et al., 2006) used previous publishing experience as a possible gauge of copyright knowledge about educational and publishing uses of third-party copyright material, but did not touch upon the academics' rights regarding their own published materials.

Copyright Concerns and Lack of Understanding as a Barrier to Deposit in Institutional Repositories

As mentioned previously, although CTAs potentially restrict a range of author activities, a significant body of literature focuses on the effects of CTAs on self-archiving in institutional repositories (IRs).

Where surveys have asked about comprehension, findings indicate that the majority of researchers experience difficulty in this area. Several studies point to copyright issues as

a major barrier to self-archiving in IRs (Cullen & Chawner, 2011; Frass et al., 2014; Kim, 2011). Creaser et al. (2010) report that scholarly authors' top two concerns when it comes to self-archiving are fear of committing copyright infringement and lack of certainty regarding embargo periods. Furthermore, 63% of respondents indicated a lack of understanding of publisher policy as an important or very important factor in failing to deposit an article in an IR (Frass et al., 2014, p. 14). Respondents in another international survey (Spezi, Fry, Creaser, Probets, & White, 2013) reported that determining their rights regarding self-archiving was the most challenging part of the process, with two-thirds reporting some degree of difficulty in doing so. Confusion concerning publisher policies is cited as a main contributory factor in studies on the massive gap between the proportion of scholarly journal articles that could be self-archived and those that are (Covey, 2009; Hansen, 2012; Laakso, 2014). For example, over half of respondents reported being unsure as to which version of an article, if any, they were permitted by their most recent publishing agreement to deposit in a repository or post online (Austin, Heffernan, & David, 2008, p.34).

Authors' Perceptions of Rights Retained vs. Terms of CTAs

In addition to self-reported uncertainty about rights retained in CTAs, many studies find that there is a significant mismatch between authors' perceived and actual rights retained under the CTAs they sign. According to Spezi et al. (2013), authors believe they have fewer rights in general than they do in practice. A 2009 study that compared authors' perceived rights with publishers' policies found more specifically that authors overestimate their rights to self-archive the publishers' PDF, while underestimating the scope of rights publisher policies permit for the submitted and the accepted versions of their articles (Morris, 2009). Creaser et al.'s (2010) findings echo this point.

While much of the literature focuses on the self-archiving rights of authors, there are a variety of ways that authors may wish to disseminate their articles as dictated by the terms of their CTAs. This may include sharing the article with colleagues, incorporating it into other published works, or distributing it for teaching purposes. Morris's study indicated that authors consistently underestimate their rights in this regard (Morris, 2009).

The self-reported anxiety and confusion over copyright constraints paired with the discrepancy between perceived and actual rights accorded by CTAs raises the question of whether authors are unsure about copyright because they do not read their publishing agreements or because they do not understand them. Although there are no studies on this topic in the library or publishing literature, there is a broad consensus in contract law scholarship that very few consumers read standard form contracts (like CTAs) before accepting them (see

Bakos, Marotta-Wurgler, & Trossen, 2014). Additionally, some publishers have incorporated the author's acceptance of the CTA into an automated manuscript submission system. This means that instead of receiving a hard copy or PDF of the document that the author can examine at leisure, authors are presented with an online click-through agreement with a checkbox to indicate consent. Recent legal and consumer behavior studies indicate that a small minority of users actually read contracts that are delivered in this manner (known as click-wrap agreements), often agreeing to onerous or one-sided terms and conditions (Helberger, Loos, Guibault, Mak, & Pessers, 2013). While the legal literature suggests that authors simply sign CTAs without reading them, when asked, the majority (with a range of 68 to 78%) of academic authors claim that they do indeed read and/or examine these agreements prior to signing them (Austin et al., 2008; Lawson & Lawson, 2011; Moore, 2011; University of California Office of Scholarly Communication et al., 2007).

As discussed above, previous studies have indicated that authors consistently state that they read and examine CTAs and that confusion and anxiety around copyright has a negative impact on green open access self-archiving. The literature also reveals a discordance between authors' perceptions of their rights and the actual scope of rights retained in CTAs. The question arises as to whether understanding the agreements themselves is a major barrier to self-archiving or if authors are not reporting their actual behavior. Instead of attempting to interrogate the veracity or accuracy of academics' self-reported behavior, our study aims to test how well academics understand the rights retained and relinquished in CTAs when they *do* read them.

METHODS

In order to determine the level of comprehension of CTAs by researchers we constructed a survey experiment including several testing questions (see Appendix A) to compare results from two different copyright transfer agreements. Specifically, the test questions were designed to gauge respondents' understanding of scholarly sharing in these agreements (e.g., self-archiving, presenting at a conference). We received research ethics board approval from our institution to conduct this survey.

The survey included basic demographic questions (i.e., academic discipline, academic publishing experience, how long ago participants published their first article, past publisher experience as corresponding author with a variety of publishers) alongside the anonymized text of one of two existing publisher copyright transfer agreements. In order to replicate natural conditions as much as possible, the CTAs were kept in their original formatting; however, any identifying information about the publisher (e.g., name, address) was removed.

In order to control for the effect of language complexity and readability on the scores, the SMOG readability formula (as detailed in Hedman, 2008) was applied to the agreements of four major publishers: Elsevier, Wiley, Springer, and Taylor & Francis. To include CTA length as a control variable, we selected one longer and one shorter CTA with proximate readability scores to be used in the experiment.

Participant Population

Participants were drawn from the article abstract and citation database Scopus (Scopus, 2018). Articles in Scopus typically contain the email address of the corresponding author, and this information was used to generate an email list for the sample population. To narrow our population in Scopus, we limited the search to English-language articles published in 2016. The following is the search string used: (DOCTYPE (ar) AND LANGUAGE (english) AND PUBDATETXT (2016) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (PUBYEAR , 2016)). The initial search provided a list of 1,656,414 articles. Due to Scopus's export limit of 2,000 items, in order to be able to extract randomly sorted data to Excel, the results were sorted newest-oldest then oldest-newest, then using the Scopus relevance ranking, in each case exporting the first 2,000 results to Excel (Elsevier, 2017). Using the resulting 6,000 items, articles with no e-mail address provided were filtered out and then duplicates removed. This provided us with a sample set of 3,245 articles from which to contact the corresponding authors as participants.

Participants were randomly assigned to one of two groups. The survey distributed to the participants was identical except for which CTA was included. One group received CTA 1, while the other received CTA 2. Participants were initially e-mailed on May 30, 2017. A reminder e-mail was then sent on June 13, 2017. Of the 3,245 e-mails sent, 91 emails (45 and 46 for CTA 1 and 2, respectively) were undeliverable, leaving a set of 3,154 participants. In total, 122 full responses were received. This represented a 4% response rate overall.

To determine participants' comprehension of the agreements as presented via the survey, we co-created an answer key and then individually scored each respondent's answers. A review of the demographics of each population group (CTA 1 and CTA 2) demonstrated that our randomization process had been successful and that there were similar demographic makeups in both groups. Indeed, difference in means is never statistically significant between the two groups (see Appendix B for balance of controls and Table 1 for the demographic makeup of the two sample populations).

Variables Used in Analysis	N (CTA 1)	N (CTA 2)	% (CTA 1)	% (CTA 2)
Discipline				
Science (includes life sciences, engineering, medicine & health sciences, physical sciences, & mathematics)	50	48	79	80
Social Sciences & Humanities (architecture & planning, arts & humanities, business, education, law, behavioral & social sciences)	11	12	17	20
Other or not indicated	2	0	3	0
Number of Articles for which Respondent Has Been the Corresponding Author in the Past 5 Years				
0–3	9	7	14	12
4–7	22	15	35	25
8+	31	38	49	63
not indicated	1	0	2	0
Number of Years since First Academic Journal Publication				
0–3	10	6	16	10
4–7	8	14	13	23
8+	44	40	70	67
not indicated	1	0	2	0

Table 1. Descriptive Statistics for Variables Used in Analysis

Statistical Analysis

To capture relationships between dependent variables and independent variables, we ran two main statistical analyses: one to capture the effect of the agreement on the overall score and another to capture how well respondents scored on individual questions depending on the agreement they received.

Pooled Analysis: Effect of Agreement on Overall Score

We investigated whether the overall score for each participant's survey was affected by which copyright transfer agreement they received (i.e., CTA 1 or CTA 2). We controlled for potential mitigating factors on the participants' scores by asking a series of demographic questions regarding academic discipline, academic publishing experience, how

long ago they published their first article (called "Seniority" in our analysis), as well as past publisher experience as corresponding author with the agreement they received (called "Publisher 1 Before" and "Publisher 2 Before" in our analysis).

For academic discipline, we divided the respondents into two categories: those whose main academic focus was in the hard sciences (i.e., life sciences, engineering, medicine & health sciences, physical sciences & mathematics) and those in the humanities and social sciences (i.e., architecture & planning, arts & humanities, business, education, law, behavioral & social sciences). This division was made as the sample size for many disciplinary categories (e.g., law, humanities, etc.) was too small to complete analysis on its own.

To determine the effect of the independent variables (academic discipline, academic publishing experience as corresponding author, seniority, and past publisher experience) on our dependent variable (score), we estimated our models using OLS regressions. This regression was chosen because our dependent variable (score) is a ratio variable.

Single-Question Analysis: Effect of Agreement on Individual Questions

In addition to determining if one of the demographic variables had an effect on the participants' overall score, we conducted an analysis of each of the individual question's scores using probit regressions and controlling for the same independent variables as listed previously. We chose a probit model, as the dependent variables in this instance were nominal (i.e., a score of one for a correct answer, a score of zero for an incorrect answer). This analysis was done to ascertain if participants were more likely to answer certain questions correctly than others and whether or not participants had higher scores on individual questions according to which CTA they had received.

RESULTS

Effect of Agreement

Our analysis shows that those who received CTA 1 answered questions correctly overall 34% of the time, while those who received CTA 2 answered the questions correctly 33% of the time. The results between the two agreements are not significantly different (see Table 2), and thus the results indicate that which agreement the respondents received had no discernible effect on participants' overall scores.

Demographic Variables

Our analysis additionally showed that the independent variables (i.e. Seniority, Discipline, etc.) had no statistically significant effect on participants' scores (see Table 2) ($p<.05$). For example, a participant who had published more than seven articles was not more likely to provide correct answers than someone who had published fewer articles.

	Share of correct answers				
VARIABLES	(1)	(2)	(3)	(4)	(5)
CTA 1	**0.340***	**0.313***	**0.327***	**0.330***	**0.331***
	(0.029)	**(0.049)**	**(0.070)**	**(0.083)**	**(0.084)**
CTA 2	**0.334***	**0.310***	**0.331***	**0.337***	**0.340***
	(0.029)	**(0.048)**	**(0.071)**	**(0.086)**	**(0.087)**
Science		0.031	0.024	0.020	0.026
		(0.050)	(0.050)	(0.051)	(0.051)
No. Articles 3–7			0.029	0.017	0.029
			(0.068)	(0.072)	(0.073)
No. Articles 8+			-0.036	-0.058	-0.035
			(0.063)	(0.076)	(0.079)
Seniority 4–7 years				-0.009	-0.006
				(0.075)	(0.076)
Seniority 8+ years				0.025	0.029
				(0.072)	(0.073)
Publisher 1 Before					-0.035
					(0.045)
Publisher 2 Before					-0.032
					(0.045)
Wald test: CTA 1=CTA2					
F statistics	0.02	0.01	0.01	0.03	0.05
Prob.>F	0.88	0.93	0.92	0.87	0.83
Observations	124	122	122	122	122
R-squared	0.688	0.696	0.701	0.702	0.705

Table 2. Share of Correct Answers Controlling for Demographic Variables: OLS Regressions with Standard Errors in Parentheses *p<0.05.

Individual Survey Questions

While overall respondents obtained the same scores regardless of which agreement they received, there were significant differences in the correct answers for *specific* questions between the two agreements. Table 3 outlines the probability that a respondent was more likely to answer a question correctly in the CTA 1 treatment than in CTA 2. For example, in Question 1 ("According to the agreement you just viewed you can post the article in a course management system"), participants who had received the CTA 1 treatment were almost 27% more likely to answer this question correctly. Conversely, participants who received the CTA 1 treatment were almost 44% *less* likely than those who received the CTA 2 treatment to correctly answer Question 7 ("According to the agreement you just viewed, you can post which version of the article to which of the following sites without seeking further approval or information: Personal website"). As previously, all demographic control variables were included in the model but had no statistically significant effects. Questions for which there were statistically significant responses are marked with an asterisk below and their corresponding p values. [1]

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
CTA 1	0.268*	0.212*	0.016	0.037	0.237*	0.133	-0.435*	-0.387*	-0.027
	(0.084)	(0.082)	(0.086)	(0.094)	(0.070)	(0.071)	(0.084)	(0.088)	(0.053)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	122	122	122	122	122	122	122	122	122
Pseudo R-squared	0.0811	0.0908	0.0372	0.0412	0.173	0.0578	0.148	0.147	0.0544

Table 3. Coefficient of CTA1 estimates the probability of respondents answering each question correctly when receiving CTA1 compared to when they receive CTA 2. Probit regression with marginal effects. Standard errors in parentheses *p<0.05.

[1] For instance: a positive and significant coefficient for Q1 implies that respondents receiving CTA 1 are 26.8% more likely to answer Q1 correctly compared to respondents treated with CTA 2 (the reverse would also be true, e.g. for Q1 respondents treated would CTA 2 would 26.8% less likely to answer correctly than respondents receiving CTA 1).

In our analysis respondents completing the CTA 1 survey were more likely to answer these three questions correctly:

- Q1 "According to the agreement you just viewed you can post the article in a course management system." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)
- Q3 "According to the agreement you just viewed you can republish the article in an edited book." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)
- Q5 "According to the agreement you just viewed you can: Reuse tables/graphs from the article in a commissioned textbook." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)

Conversely, participants who had received CTA 2 were more likely to answer correctly the following questions:

- Q7 "According to the agreement you just viewed, you can post which version of the article to which of the following sites without seeking further approval or information: Personal website." (Possible responses: Publisher version; Accepted version; Submitted/Prior Version; Not permitted for any version)
- Q8 "According to the agreement you just viewed, you can post which version of the article to which of the following sites without seeking further approval or information: Institutional repository." (Possible responses: Publisher version; Accepted version; Submitted/Prior Version; Not permitted for any version)

DISCUSSION

Contrary to what we expected, demographic variables had no effect on participants' overall scores on the survey. We anticipated that participants who had more publishing experiencing as a corresponding author would have more exposure to copyright transfer agreements and thus would have greater comprehension than those who had less experience; this turned out to not be the case. Similarly, we also anticipated that participants who had experience as a corresponding author with the publisher whose CTA they were administered would have higher scores than those who did not. For instance, participants who had had past experience publishing with Publisher 1 AND received CTA 1 had similar scores to those who had not published previously with Publisher 1 and received CTA 1. This lends credence to the perception that even experienced researchers lack understanding of copyright transfer agreements and that no particular population of researchers we identified in our

survey (e.g. senior/junior, certain disciplines, past publishing experience, etc.) is more adept at reading and comprehending these agreements. Further, our results belie the assumption of a learning effect; while it might be expected that authors who had published previously with the publisher whose CTA they received would be more familiar with that agreement and therefore more likely to understand it, there was no evidence to support this hypothesis.

Explicit Wording

Although the overall number of correct answers was 33% for both surveys (or close to what might be expected if respondents had guessed at random) this does not appear to be due to random selection of answers by participants. The finding that between the agreements there was a significant difference on *which* questions they got right or wrong overall suggests that the respondents were indeed reading the agreement and attempting to answer correctly. In reviewing the answers for the five questions for which there was a significant difference between the two agreements, we hypothesize that *explicitness* was a contributing factor to the respondents' success rate on those questions. We define explicitness as the degree to which the question asked in the survey has an answer that is explicitly found in the text, with a prohibited or permitted activity and any attendant conditions detailed, as opposed to being inferred or not discussed at all.

For example, below are the three questions for which respondents were more likely to answer correctly if they received the CTA 1 treatment. Included under each question is the paragraph from the agreement that directly corresponds.

- Q1 "According to the agreement you just viewed you can post the article in a course management system." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)
    - "Electronic posting of the Final Published Version in connection with teaching/training at the Contributor's company/institution is permitted subject to the implementation of reasonable access control mechanisms, such as user name and password."
- Q3 "According to the agreement you just viewed you can republish the article in an edited book." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)
    - "The right to re-use the Final Published Version or parts thereof for any publication authored or edited by the Contributor (excluding journal articles) where such re-used material constitutes less than half of the total material in such publication."

- Q5 "According to the agreement you just viewed you can: Reuse tables/graphs from the article in a commissioned textbook." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)
    - "Contributors may re-use figures, tables, artwork, and selected text up to 250 words from their Contributions, provided the following conditions are met:…"

In addition, respondents were somewhat more likely to answer the following question correctly for CTA 1 (though the effect is only weakly significant, i.e., $p<0.01$):

- Q6 "According to the agreement you just viewed you can: present the article at a conference." (Possible responses: Yes, absolutely; Yes, with conditions; No; Unclear)
    - "The right to include the Final Published Version in teaching or training duties at the Contributor's institution/place of employment including in course packs, e-reserves, presentation at professional conferences.."

For CTA 2, conversely:

- Q7 "According to the agreement you just viewed, you can post which version of the article to which of the following sites without seeking further approval or information: Personal website." (Possible responses: Publisher version; Accepted version; Submitted/Prior Version; Not permitted for any version)
    - "Authors may self-archive the Author's accepted manuscript of their articles on their own websites."
- Q8 "According to the agreement you just viewed, you can post which version of the article to which of the following sites without seeking further approval or information: Institutional repository." (Possible responses: Publisher version; Accepted version; Submitted/Prior Version; Not permitted for any version)
    - "Authors may also deposit this version of the article in any repository, provided it is only made publicly available 12 months after official publication or later."

Although intuitively explicitness seems to aid in comprehension, and our results bear this intuition out to some extent, it does not appear to be a panacea. Both surveys included several explicit question/answer pairs that did not demonstrate statistically significant results. For example, respondents on the CTA 2 survey did very well in determining which version

of the article was permitted for various uses, whereas respondents for the CTA 1 survey did poorly in spite of the answers being clearly stated in the text. While there is no readily apparent reason for this discrepancy, we speculate that this is one area in which CTA length was a factor; since CTA 2 was shorter and outlined fewer possible permitted uses, it may have been easier for respondents to isolate and identify those uses that were made explicit. In addition, the statement of permitted uses in CTA 1 was organized by what was allowed for each version of the article, rather than by use, which may have confounded participants' efforts to identify the permitted activity.

Assuming Permission If Wording Is Unclear

Further analysis of the data revealed that in instances where the answer to the question was "unclear" (i.e., not explicitly discussed in the text nor implicit), typically 60–75% of respondents chose either "Yes, with conditions" or "Yes, absolutely." Given that both CTAs included as treatments in the survey constituted copyright transfers, authors should assume that they do not retain any rights that are not specifically licensed back to them in the text of the agreement. As such, this is either wishful thinking on the part of the respondents or a tendency to believe something is permitted unless it is stated otherwise.

Negative Response to Republishing in an Edited Book

Another interesting finding was that for Question 3 ("According to the agreement you just viewed you can republish the article in an edited book"), a majority of respondents for both surveys (CTA 1 and CTA 2) selected "No" as their response. This is notable as it was the only question for which more than 50% of respondents selected a negative answer. What is even more intriguing is that, in the case of CTA 1, this usage was explicitly allowed with listed conditions, and in the case of CTA 2, such permission could be inferred. It is unclear why respondents had such a strongly negative (and incorrect) response to this question; however, this finding is in keeping with Morris's (2009) study, which found that authors significantly underestimated the extent to which publisher policies permitted them to incorporate their published articles into subsequent publications. This misperception may be informed by scholarly custom, according to which republishing a piece of scholarly work would necessitate explicit permissions to be sought (Gadd, 2017).

Future Studies and Applications

There are several immediate applications of this research. Firstly, author understanding of these agreements is minimal. Respondents averaged a correct response rate of 33% demonstrating the importance of improving the comprehension of these agreements. Bearing this

in mind, librarians who are involved in author rights education and advocacy should focus on providing academic authors with the tools to parse the language and phrasing of CTAs, particularly on how to isolate the conditions that are most important to the authors, such as the ability to self-archive to fulfill open access mandates or otherwise. Librarians should additionally focus on reaching out to related departments on their campuses such as research offices, promotion and tenure committees, faculty councils, graduate student associations, and legal offices to partner in coaching faculty on reading these agreements. Understanding legal documents related to publishing should be a campus-wide effort to engage and inform authors for which the library can take a leading role.

Second, we found that demographic variables and experience had no effect on the results. This finding overrides any assumptions librarians and publishers might have about CTA knowledge and understanding among faculty. For example, while one might assume that those who publish frequently or are senior faculty may be more adept at reading these agreements, this is not the case. Armed with this knowledge, librarians should cast a wide net in their author rights educational and advocacy efforts in order to reach all demographics. This could once again be aided by developing campus partnerships outside the library.

Third, explicitness has a potential effect to improve the readability of copyright transfer agreements. As one respondent wrote, "The wording [of the agreement] is very subjective and is not clear." Improving the clarity of wording in CTAs would have a twofold benefit: it would ensure that what is permitted/not permitted is clearly laid out, potentially mitigating the chilling effect of copyright confusion on author' self-archiving or other scholarly reuse behavior, and greater explicitness would avoid the "unclear" effect, whereby authors assume they have greater permission in the absence of a clear indication otherwise.

It would be interesting to test the explicitness hypothesis further, as a positive finding could be immediately utilized by publishers to their benefit. For example, there have been many conflicts recently between authors and publishers as publishers increasingly take steps to enforce their copyright policies by issuing takedown notices to social networking and university sites and, in some cases, take legal action (Chawla, 2017; Mika, 2017). Our research suggests that in many cases, the authors were not aware that they were breaching copyright by posting the publisher's final version to these sites. It has also been suggested that the language many publishers use on their websites portray their scholarly sharing policies as being more permissive than they actually are, which may lead to further confusion (Gadd, 2017). In either case, a clearer statement in the CTA forbidding these activities may aid in mitigating such transgressions. Although many authors may still choose to ignore their agreements, it is also likely that some authors would adjust their behavior accordingly.

In addition, increasing explicitness in copyright transfers agreements would improve literacy, and librarians may find more researchers engaging with their author rights as they become more aware of what they are signing. With this in mind, librarians can advocate for more unambiguous language in CTAs wherever they have the influence to do so, be that as members of editorial boards, as providers of education and infrastructure through library publishing initiatives, or through scholarly communication education and outreach and engaging with faculty members and colleagues who sit on editorial boards.

Limitations

Our response rate, though sufficient for running our analyses, was a small percentage of the population that received our study invitation. Further studies could improve the response rate by offering incentives to respondents. Additionally, in attempting to recreate real-world settings by using the actual text and layout of existing CTAs, we acknowledge that even controlling for readability, there may have been elements of the layout or typography that affected respondents and for which we did not control.

Finally, the scope of our study was limited to testing author understanding of two CTAs. Subsequent research could examine the content, length, and syntax of a larger corpus of CTAs. Where the current study was limited to CTAs that effected a copyright transfer, an expanded study could examine agreements which take the form of exclusive and nonexclusive licenses, including the Creative Commons license variants.

CONCLUSION

Our results indicate that academic authors have poor comprehension of the CTAs that govern the terms under which they may reuse and make available their own published journal articles. This lack of understanding does not vary significantly with seniority, academic discipline, experience publishing, or previous publishing experience with a given publisher. Nor does the length of the agreement seem to increase or decrease levels of understanding. The one factor that appears, in some cases, to have a positive effect on authors' levels of understanding is when permitted or prohibited activities are mentioned explicitly in the text of the CTA.

ACKNOWLEDGMENTS

We wish to thank Leonardo Baccini for his valuable assistance and suggestions regarding the statistical analysis and Dawn McKinnon for comments on earlier drafts of this paper. Many thanks also to Francesco Amodio, Robin Desmeules, Krzysztof Pelc, Jane Secker, and Nikki Tummon for advice on the design of the survey instrument.

REFERENCES

Archambault, É., Amyot, D., Deschamps, P., Nicol, A., Provencher, F., Rebout, L., & Roberge, G. (2014). Proportion of open access papers published in peer-reviewed journals at the European and world levels—1996–2013. *Copyright, Fair Use, Scholarly Communication, Etc*. Retrieved from http://digitalcommons.unl.edu/scholcom/8

Austin, A. C., Heffernan, M. E., & David, N. (2008, May). Academic authorship, publishing agreements and open access: Survey results [Other]. Retrieved from http://eprints.qut.edu.au/13623/

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. R. (2014). Does anyone read the fine print? Consumer attention to standard-form contracts. *Journal of Legal Studies, 43*(1), 1–35. https://doi.org/10.1086/674424

Björk, B.-C. (2017). Gold, green, and black open access. *Learned Publishing, 30*(2), 173–175. https://doi.org/10.1002/leap.1096

Charbonneau, D. H., & McGlone, J. (2013). Faculty experiences with the National Institutes of Health (NIH) public access policy, compliance issues, and copyright practices. *Journal of the Medical Library Association 101*(1), 21–25. https://doi.org/10.3163/1536-5050.101.1.004

Chawla, D. S. (2017, October 6). Publishers take ResearchGate to court, alleging massive copyright infringement. *Science | AAAS*. Retrieved from http://www.sciencemag.org/news/2017/10/publishers-take-researchgate-court-alleging-massive-copyright-infringement

Covey, D. T. (2009). Self-archiving journal articles: A case study of faculty practice and missed opportunity. *Portal: Libraries and the Academy, 9*(2), 223–251. https://doi.org/10.1353/pla.0.0042

Creaser, C., Fry, J., Greenwood, H., Oppenheim, C., Probets, S., Spezi, V., & White, S. (2010). Authors' awareness and attitudes toward open access repositories. *New Review of Academic Librarianship, 16*(sup 1), 145–161. https://doi.org/10.1080/13614533.2010.518851

Cullen, R., & Chawner, B. (2011). Institutional repositories, open access, and scholarly communication: A study of conflicting paradigms. *Journal of Academic Librarianship, 37*(6), 460–470. https://doi.org/10.1016/j.acalib.2011.07.002

Elsevier. (2017, January 8). What does "Relevance" mean in Scopus? Retrieved from https://service.elsevier.com/app/answers/detail/a_id/14182/kw/relevance/supporthub/scopus/

Frass, W., Cross, J., & Gardner, V. (2014). *Taylor & Francis open access survey.* Taylor & Francis/Routledge. Retrieved from http://www.tandf.co.uk/journals/explore/open-access-survey-june2014.pdf

Gadd, E. (2017, October 3). Academics and copyright ownership: Ignorant, confused, or misled? Retrieved from https://scholarlykitchen.sspnet.org/2017/10/31/guest-post-academics-copyright-ownership-ignorant-confused-misled/

Gadd, E., & Troll Covey, D. (2016). What does "green" open access mean? Tracking twelve years of changes to journal publisher self-archiving policies. *Journal of Librarianship and Information Science*. https://doi.org/10.1177/0961000616657406

Hansen, D. (2012). Understanding and making use of academic authors' open access rights. *Journal of Librarianship and Scholarly Communication, 1*(2). https://doi.org/10.7710/2162-3309.1050

Hedman, A. S. (2008). Using the SMOG Formula to Revise a Health-Related Document. *American Journal of Health Education, 39*(1), 61–64. https://doi.org/10.1080/19325037.2008.10599016

Helberger, N., Loos, M. B. M., Guibault, L., Mak, C., & Pessers, L. (2013). Digital content contracts for consumers. *Journal of Consumer Policy, 36*(1), 37–57. https://doi.org/10.1007/s10603-012-9201-1

Jamali, H. R. (2017). Copyright compliance and infringement in ResearchGate full-text journal articles. *Scientometrics, 112*(1), 241–254. https://doi.org/10.1007/s11192-017-2291-4

Kim, J. (2011). Motivations of faculty self-archiving in institutional repositories. *Journal of Academic Librarianship, 37*(3), 246–254. https://doi.org/10.1016/j.acalib.2011.02.017

Laakso, M. (2014). Green open access policies of scholarly journal publishers: A study of what, when, and where self-archiving is allowed. *Scientometrics, 99*(2), 475–494. https://doi.org/10.1007/s11192-013-1205-3

Lawson, A., & Lawson, A. (2011, August). Report on the UWE research repository survey, June 2011. Retrieved from http://eprints.uwe.ac.uk/15720/

Lovett, J., Rathemacher, A., Boukari, D., & Lang, C. (2017). Institutional repositories and academic social networks: Competition or complement? A study of open access policy compliance vs. ResearchGate participation. *Journal of Librarianship and Scholarly Communication, 5*(1). https://doi.org/10.7710/2162-3309.2183

Mika, A. (2017, June 15). Authors peeved by APA's article takedown pilot. Scientist. Retrieved from https://www.the-scientist.com/?articles.view/articleNo/49670/title/Authors-Peeved-by-APA-s-Article-Takedown-Pilot/

Moore, G. (2011). Survey of University of Toronto faculty awareness, attitudes, and practices regarding scholarly communication: A preliminary report. Retrieved from https://tspace.library.utoronto.ca/handle/1807/26446

Morris, S. (2009). Journal authors' rights: Perception and reality (PRC Summary Paper No. 5). London: Publishing Research Consortium. Retrieved from http://publishingresearchconsortium.com/index.php/110-prc-projects/summary-papers/journal-authors-rights-summary-paper/140-journal-authors-rights-perception-and-reality

Odell, J. D., Dill, E., & Palmer, K. L. (2014). Authors' rights to share scholarship: A survey of IUPUI faculty attitudes. Retrieved from https://scholarworks.iupui.edu/handle/1805/4246

ROARMAP. (n.d.) Browse by policymaker type - ROARMAP. Retrieved from http://roarmap.eprints.org/view/policymaker_type/

Scopus. (2018). Retrieved March 22, 2018, from https://www.scopus.com

SHERPA Juliet. (n.d.) Juliet statistics - SHERPA services. Retrieved from http://v2.sherpa.ac.uk/view/funder_visualisations/1.html

Smith, K. H., Tobia, R. C., Plutchak, T. S., Howell, L. M., Pfeiffer, S. J., & Fitts, M. S. (2006). Copyright knowledge of faculty at two academic health science campuses: Results of a survey. *Serials Review, 32*(2), 59–67. https://doi.org/10.1080/00987913.2006.10765031

Solon, O. (2013, December 17). Elsevier clamps down on academics posting their own papers online. *Wired*. Retrieved from http://www.wired.co.uk/article/elsevier-versus-open-access

Spezi, V., Fry, J., Creaser, C., Probets, S., & White, S. (2013). Researchers' green open access practice: A cross-disciplinary analysis. *Journal of Documentation, 69*(3), 334–359. https://doi.org/10.1108/JD-01-2012-0008

Suber, P. (2012). OA Policies at funding agencies and universities. In *Open access* (pp. 78–95). Cambridge, MA: MIT Press. Retrieved from http://archive.org/details/9780262517638OpenAccess

University of California Office of Scholarly Communication, California Digital Library eScholarship Program, & Greenhouse Associates, Inc. (2007). *Report on faculty attitudes and behaviors regarding scholarly communication*. Retrieved from http://osc.universityofcalifornia.edu/2007/08/report-on-faculty-attitudes-and-behaviors-regarding-scholarly-communication/

Van Noorden, R. (2017). Publishers threaten to remove millions of papers from ResearchGate. *Nature News*. https://doi.org/10.1038/nature.2017.22793

APPENDIX A

Survey Instrument

Background and purpose:

The aim of this survey is to investigate researchers' understanding of publisher copyright transfer agreements and the effect that word choice, syntax, and document length have on comprehension.

Study Procedures:

The survey should take approximately 5–10 minutes to complete. You may begin the survey by clicking on "Next" at the end of this information letter. Participants will be asked to read a sample agreement and then answer a series of questions.

Benefits and Risks:

Participating may improve your awareness and/or comprehension of copyright transfer agreements. There are no perceived risks with completing this survey.

Voluntary participation:

Your participation in this survey is completely voluntary. You may choose not to answer any questions, and you may withdraw from the survey at any time before submission of your results by closing your browser or selecting "Exit and clear." Incomplete surveys will not be saved. As the information being gathered is anonymous, it will not be possible to have answers withdrawn once the complete survey is submitted.

Confidentiality and Anonymity:

Your responses are completely anonymous and will not be traced back to you. Data collected via this survey may also be used in research articles or professional presentations. Any shared results will be in aggregate form only. Any free text responses that are shared will have all possible identifying information removed. This survey has been approved by McGill University's Research Ethics Board. If you have questions about this survey, please contact either of the researchers of this project: Jessica Lange, McGill University Library & Archives or Alexandra Kohn, McGill University Library & Archives. If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the McGill University Ethics Manager.

By clicking Next, you consent that you are willing to answer the questions in this survey. Please save or print a copy of this document to keep for your own reference.

1. Please select your area/discipline.
   a. Architecture & Planning
   b. Arts & Humanities
   c. Behavioral & Social Sciences
   d. Business
   e. Education
   f. Engineering
   g. Law
   h. Life Sciences
   i. Medicine & Health Sciences
   j. Physical Sciences & Mathematics
   k. Other (please indicate)

2. On how many articles have you been the corresponding author[2] in the past 5 years?
   a. 0
   b. 1-3
   c. 4-7
   d. 8+
   e. No answer

3. With which publishers have you published an article as corresponding author? (Check all that apply.)
   a. American Chemical Society
   b. Cambridge University Press
   c. Elsevier
   d. Emerald
   e. IEEE

[2] Corresponding author is defined as the researcher who looks after the manuscript submission process and is the primary contact for the article.

f. Oxford University Press
g. Sage
h. Springer Nature
i. Taylor & Francis
j. Wiley
k. Other

4. How many years ago was your first academic journal publication? (Choose one of the following answers.)

a. 0-3
b. 4-7
c. 8+
d. No answer

Below is a sample publisher agreement. Please read this agreement. You will be asked questions on the content of this text afterwards.

[TEXT OF CTA 1 OR CTA 2 HERE]

5. According to the agreement you just viewed you can:

	Yes, absolutely	Yes, with conditions	No	Unclear	No answer
Post the article in a course management system.					
Use the article in a course pack / course reader / distribute print copies.					
Republish the article in an edited book.					
Include the article in a dissertation or thesis.					
Reuse tables/graphs from the article in a commissioned textbook.					
Present the article at a conference.					

6. According to the agreement you just viewed, you can post which *version of the article* to which of the following sites without seeking further approval or information:

	Publisher version	Accepted version	Submitted/ Prior Version	Unclear which version	Not permitted for any version	No answer
Personal website						
Institutional Repository						
Social networking sites (e.g. ResearchGate, Academia.edu, etc.)						

7. Please add any additional comments related to publisher copyright transfer agreements.

APPENDIX B

	Mean		
	CTA 1	CTA 2	Same Mean
Science	0.81	0.77	Yes
No. Articles	2.36	2.52	Yes
Seniority	2.53	2.57	Yes
Publisher 1 Before	0.42	0.48	Yes
Publisher 2 Before	0.34	0.38	Yes

Balance of Controls; Mean of demographic groups in each treatment (CTA 1 and CTA 2)