

# Who's Talking about Scholarly Communication? An Examination of Gender and Behavior on the SCHOLCOMM Listserv 

Clayton Hayes, Heidi Elaine Kelly

Hayes, C., \& Kelly, H.E. (2017). Who's Talking about Scholarly Communication? An Examination of Gender and Behavior on the SCHOLCOMM Listserv. Journal of Librarianship and Scholarly Communication, 5(General Issue), eP2017. https://doi.org/10.7710/2162-3309.2017

## External Data or Supplements:

Hayes, C., Kelly, H.E. (2017). Data from: Who's Talking about Scholarly Communication? An Examination of Gender and Behavior on the SCHOLCOMM Listserv. https://doi.org/10.7910/DVN/MABOGR, Harvard Dataverse.

## Erratum:

Added citation for external data 4/20/2017


# Who's Talking about Scholarly Communication? An Examination of Gender and Behavior on the SCHOLCOMM Listserv 

Clayton Hayes<br>STEM and Digital Publishing Librarian, Wayne State University Libraries, Wayne State University<br>Heidi Elaine Kelly<br>Digital Preservation Librarian, Herman B Wells Library, Indiana Unviersity Bloomington


#### Abstract

INTRODUCTION This study analyzes the gender dynamics of the American Library Association's SCHOLCOMM listserv in order to determine the accuracy of concerns expressed by participants in early 2016 regarding an overrepresentation of male voices on the listserv. METHODS Utilizing the SCHOLCOMM listserv archive, openly available online, the authors analyzed metadata related to individual messages in order to create a comprehensive list of participants, which was then analyzed to determine gender identity. The authors utilized this information to correlate the frequency of new messages and replies sent to the list with the gender identity of participants. RESULTS While men represented 35\% of the SCHOLCOMM list's participants, they contributed over half of the messages sent to the listserv and two-thirds of those sent as replies on existing message threads. DISCUSSION The opinion of several SCHOLCOMM participants that male voices were overrepresented in listserv discussions proved to be true. The gender identity breakdown of those most active on the list may also influence the perceptions and/or behaviors of other listserv participants, however, and should be investigated further. CONCLUSION While this study substantiates the opinion of several listserv participants that male SCHOLCOMM participants account for a disproportionately large amount of listserv discussion, we argue that the dynamics of the listserv can and should be changed in order to better represent the participant population.


Received: 12/08/2016 Accepted: 03/07/2017

Correspondence: Clayton Hayes, Room 134, Purdy/Kresge Library, Wayne State University, Detroit, MI, 48202, as6348@wayne.edu

## IMPLICATIONS FOR PRACTICE

1. As the dynamics on the SCHOLCOMM listserv reflect the larger community of scholarly communication professionals, this study will inform further inquiries into gender dynamics of online communities in this field.
2. This study serves to support the claim that there is a need for more group and selfawareness in terms of gender dynamics in online scholarly communities of librarians and scholarly communication professionals.
3. This study provides several concrete steps that will enable scholarly communities to develop awareness and initiate behavioral changes to provide more support for inclusivity.

## INTRODUCTION

Scholarly communication is "the system through which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community, and preserved for future use" (Association of College \& Research Libraries [ACRL], 2003); stakeholders in scholarly communication include a wide array of academics and information professionals, notably scholarly communications librarians. The American Library Association (ALA) is the largest library association in North America and, as such, its listservs are uniquely situated to provide participants with the opportunity to communicate with colleagues throughout the United States and, as is often the case, the world. Its Scholarly Communication discussion list, commonly referred to as SCHOLCOMM, is provided through the ALA's higher education division, the ACRL, and should ideally serve as an environment in which participants feel comfortable expressing opinions and contributing to conversations. In March of 2016, however, as part of an unrelated conversation on the list, a number of participants posted to the list expressing the opinion that male voices were dominating discussions and discouraging non-male contributions. This was quickly met with dissent; replies were posted to the list which asserted that such opinions were baseless and that statistical proof was required in order for them to be taken seriously. Kevin Smith, at that time the director of the Office of Copyright and Scholarly Communication at Duke University, performed an ad-hoc analysis of several recent list discussions with more than ten replies. He found that $40 \%$ of the discussions came from four male individuals, but did not undertake further analysis at the time (Smith, 2016).

The current study was undertaken as a systematic extension of Smith's work, with the hopes that it may provide a foundation of data and analysis on which an informed discussion of gender dynamics on the SCHOLCOMM list can be based and on which future research can build. This paper examines basic metadata harvested from the SCHOLCOMM archive
provided by the ALA on every post to the list from its inception in February of 2003 to the end of December in 2015. By identifying each participant uniquely and correlating their gender identity with the frequency with which they posted and replied to the list, it seeks to determine what difference, if any, exists in how frequently men and women participate in discussions on the SCHOLCOMM list.

## LITERATURE REVIEW

## Gender Distribution in LIS and Scholarly Communication

For the purposes of this study, it is important to begin to parse the distribution of gender identities within the field of library and information science (LIS) broadly, and more specifically within scholarly communications librarianship. According to the Oxford University Press Librarian Census, which surveyed the field of librarianship from 1880-2009, women made up $83 \%$ of the field at the end of the study, leveling down from $92 \%$ in 1930 (Beveridge, Weber, \& Beveridge, 2011). Likewise, a 2014 survey of ALA members found that $81 \%$ of respondents were women (American Library Association, 2014). This suggests that there is still a strong bias towards women in the profession, though this varies based on type of library and position. Within academic libraries, where the field of scholarly communications librarianship is largely situated, gender distribution is not as clearly delineated. Based on the most recent openly accessible Association of Research Libraries survey of member institutions (Kyrillidou \& Young, 2006), women made up $63 \%$ of all professional staff within academic libraries, but only $39 \%$ of academic librarians. Within scholarly publishing, a field with close ties to academic and scholarly communications librarianship, a 2014 analysis of the Society of Scholarly Publishers (SSP) members found that $58 \%$ were women (Kane \& Meadows, 2014). As an analysis by West, Jacquet, King, Correll, and Bergstrom (2013) found, however, even in academic fields where genders were relatively evenly distributed, gender distribution within subfields can vary widely. It is therefore unclear which trends, if any, are applicable to scholarly communications librarianship, which has yet to be singled out for analysis.

## Online Communication in Scholarly Communications Librarianship

Established in the mid-1970s, the subfield of scholarly communications librarianship has a robust community of practitioners and, as with other academic research areas, various mechanisms for online interaction across the greater community. A study by Sugimoto, et al. (2012) of librarians' information dissemination and consumption practices found that, while the main modes for sharing information are the more traditional conference presentations and research articles, many librarians are now engaging with the professional commu-
nity via social media, listservs, and blogs (p. 151). Within scholarly communications librarianship, several online resources are deeply embedded in the community's practices around sharing knowledge, including organizational blogs such as the SSP's Scholarly Kitchen, ${ }^{1}$ individual university blogs such as that of Duke University ${ }^{2}$ or Indiana University, ${ }^{3}$ independent blogs such as In the Open, ${ }^{4}$ and the SCHOLCOMM listserv. ${ }^{5}$ There is likely heavy crossover within these online communities in terms of users and levels of participation; the SCHOLCOMM listserv was considered for this study because it is open and available to any individual who would like to participate.

## Gender in Online Scholarly Communities

The issue of gender dynamics in online communities has been widely studied across a variety of fields and disciplines (e.g., Herring, 1992; van Doorn \& van Zoonen, 2009). Many of these studies are unique to their realms of investigation and may not necessarily apply directly to scholarly communications librarianship; they do provide, however, a clearer view of the gender dynamics at play in online scholarly communities, especially communities of librarians. According to Herring, while earlier research presumed that the mediation of the online world would eliminate gender differences, it has been shown through various studies that gender continues to have a deep impact on the ways in which users interact with one another (as cited in Sierpe, 2001, p. 340). This impact varies based on study, however, which is perhaps due to the differences in topics of conversation and user contexts; these can play a role in defining the general style and frequency of communication (Baym, as cited in van Doorn \& van Zoonen, 2009, p. 262). While McGee and Briscoe (2003) found that women were more active on a general faculty listserv (p. 139), Sierpe (2001), using a methodology similar to that of the current study, found that men contributed nearly $59 \%$ of all messages on an LIS forum, though they comprised only $40 \%$ of those subscribed to the forum (p. 345). They also found that male top contributors were more likely than female top contributors to be active in many discussions, and that men were more likely to contribute multiple times to the same discussion (p. 346). In terms of more traditional publishing in scholarly communications librarianship, the gender gap is clearer. Gul, Shah, Hamade, Mushtaq, and Koul (2014) found that only about one-quarter of the authors published by the Electronic

[^0]Library Journal were women, or teams of women (p. 496).

## METHODS

The ALA website provides a public-facing archive of all emails distributed via the various ALA discussion lists; SCHOLCOMM is no exception. These archives are formatted as a list wherein each line represents an individual message and consists of the subject, author, and timestamp separated by commas. Each line links to the content of that particular email, though only the metadata (subject, author, and timestamp) were collected for this study. This formatting is very convenient in that these lists can be easily copied and pasted into a text editor (such as SublimeText, which was used in this study) and saved as a Comma-Separated Values (CSV) file. After a small amount of clean-up, these CSV files can be imported into Microsoft Excel or a statistical analysis tool, such as the R statistical software, in order to evaluate the data contained therein.

Initially, one CSV file was created for each month in the SCHOLCOMM archives' existence; these files were combined into thirteen Excel workbooks, each covering an entire year. From these workbooks, a master list of SCHOLCOMM participants was established. This was essential, as a particular author's name did not always remain uniform throughout the archive; one author, for example, had at least five distinct versions of their name appear in the data. As a result, a fair amount of sleuthing needed to be done in order to establish a single identity for each author.

Once each participant was uniquely identified, pivot tables were used to count the number of emails sent to the SCHOLCOMM list by each author for each year. This was done in order to gauge overall participation in the list. It was important, however, to attempt to address how frequently male and female participants interacted with others on the list; many of the emails sent to the list are of a closed nature, such as job postings, calls for papers, or conference announcements. In order to get a rough idea of interactions via the list, the number of "reply" emails sent by each participant, i.e. emails sent as a response to a posting on the list, were also counted. This was established using an Excel formula which noted whether or not "re:" appeared in the subject field of each message. Some manual checking was required at this point but, thanks to the fact that the list's archive is organized into conversation threads by default, this was not as onerous as it could have been. These yearly data were organized into a single Excel workbook, and pivot tables were again used to combine the yearly data into a single spreadsheet covering the entire history of the SCHOLCOMM list.

It is important to note that not all SCHOLCOMM interactions appear in the archive. Replies can and do happen off-list, and there is no viable method for collecting off-list mes-
sages for inclusion in the data. Such messages fall outside of the scope of this study, however, as public participation in the SCHOLCOMM list is of primary interest. It should also be noted that, on occasion, a line reading "Message not available" appeared in the archive for the listserv. There is no way of determining why a particular message may not appear in the archive, according to the list's Server Administrator, but possible explanations include deletion of messages by administrators due to privacy concerns and a variety of technical issues (personal communication, December 1, 2016). No such "Message not available" instances exist before 2011, and all "Message not available" items appear to be replies sent to the list. In total, 132 of these errors appear in the SCHOLCOMM archive for the years 2003 to 2015. This would constitute approximately $2.6 \%$ of total messages, and they have not been included in the final data set analyzed as part of this study. After establishing yearly counts for both total number of messages sent to the list and number of replies sent to the list, the gender identity of each participant was coded for analysis.

Solely for the purposes of this study, four gender categories were established: M , signifying male-identifying participants; F, signifying female-identifying participants; U, signifying participants who either do not identify as either male or female or whose gender identity could not be established; and N/A for non-human participants such as organizational and SPAM emails. Positively identifying each participant's gender identity with absolute certainty would require asking them directly; with 650 list participants, over a third of which had only ever sent a single email to the SCHOLCOMM list, this was not perceived to be a viable option. Instead, each participant was placed into one of the above categories as follows: Names which the authors considered to be regularly associated with either the male (e.g. John or Robert) or female (e.g. Christina or Maria) gender, based largely on common first names in the United States, were categorized accordingly; this accounted for the majority of participants. For any name that could conceivably be considered agendered, or for any name that was unfamiliar to the researchers, a title or place of employment was established based on messages sent to the list if possible. Searches were then performed through online search engines or the participant's institution or employer in order to find any biographical information; the gender identity of the participant was then based on exclusively any gendered pronouns contained in said biographical information. A total of 375 female and 228 male participants were identified. In situations where it was not possible to locate such information (e.g., participants did not include a full name, a title, or a place of employment in their messages; there was no biographical information available; or biographical information did not include any gendered pronouns), the participant was placed in the "U" category as described above. Though the authors had intended the "U" category to include any gender nonconforming individuals, no participants were positively identified as such. A total of 21 participants were assigned to the "U" category. Finally, institutional or organizational participants were easily identified by the name given in the data as a post's author, e.g. NISO
or ARL Communications. SPAM emails were not as immediately obvious as they were often sent from accounts with human names but were generally evident from the message's subject line. These were all grouped into the "N/A" category, which totaled 26 participants.

Though the authors have determined that the method outlined here is adequate for the purposes of the current study, it is fully acknowledged that this is an oversimplification of gender identity and sex. In particular, despite the biological connotations that may accompany the terms "male" and "female", they are used to signify gender identity in the present study. Furthermore, as discussed above, research for this study was undertaken without contacting the listserv for permission or for further information, such as preferred gender identities, in order to avoid unnecessary time constraints or skewing of the data due to low participation rate. The section below discussing possible avenues for future research outlines steps which may be taken to build on these findings while also supporting a more nuanced understanding of gender dynamics on the listserv.

## RESULTS

The number of unique male and female individuals who posted to SCHOLCOMM along with the total number of posts from male and female participants each year is presented in Table 1, with percentage representations of those data given in Table 2. Overall, male participants comprised $35.1 \%$ of individuals posting to the SCHOLCOMM list but contributed $51.2 \%$ of the messages sent to SCHOLCOMM from 2003 to 2015. In comparison, female participants comprised $57.7 \%$ of the individuals posting to the list and accounted for $45.8 \%$ of the messages posted to the list.

This gender disparity becomes more pronounced when considering only those messages that were sent as replies to other posts on the list. Table 3 displays the number of unique male and female individuals who posted replies to the list in a given year along with the total number of replies posted by male and female participants in that year, with percentage given in Table 4. Female participants accounted for $53.8 \%$ of individuals replying to the list, but only accounted for $32.9 \%$ of reply emails. Male participants, however, accounted for $66.7 \%$ of reply emails while only comprising $43.8 \%$ of individuals replying to the list.

Some simple statistical treatments, performed with the R statistical software, were used in order to examine the correlation between gender and interaction with the SCHOLCOMM list. Two measures of list interaction were used, both based on replies posted to the list: a) whether or not an individual had ever replied to a list message and b) the number of replies and the number of non-replies sent to the list by an individual. As above, only male and female list participants were considered for these statistical treatments.
Table 2．Unique male and female SCHOLCOMM participants total post count，percentages

| \％でしs | \％697 | \％1•9s | \％1－99 | \％ガOS | \％l｀¢s | \％909 | \％9「ワワ | \％ャワ¢ | \％ 1 －0¢ | \％し「ヤの | \％6－99 | \％が18 | \％l｀Sb | W |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| \％8＇¢ワ | \％S＇$\angle \square$ | \％ع＇6を | \％9｀ャ | \％6＊$\angle \square$ | \％ぐらワ | \％6¢¢ | \％と＇てS | \％ャ゙て9 | \％8．69 | \％ع．$¢ \varepsilon$ | \％し｀ャを | \％でャレ | \％6『๑ | 」 |
| ｜0¢0］ | sioz | ャレOZ | £ı0乙 | てıOZ | lıOZ | OLOZ | 600Z | 800Z | LOOZ | 900て | s00z | ャ00て | \＆00乙 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| \％l｀$¢ \varepsilon$ | \％ャ゙とદ | \％ガてを | \％$\angle$－8غ | \％0・ワワ | \％8『ャ | \％6゙で | \％s＇97 | \％9＇で | \％ $1 \cdot 8 \varepsilon$ | \％0087 | \％ぐ99 | \％S＇$\angle 8$ | \％＜＇\＆ऽ | W |
| \％$\angle \subset \angle S$ | \％どと9 | \％6．8S | \％1｀09 | \％と＇てS | \％009 | \％0067 | \％s＇97 | \％688 | \％6．19 | \％0＾8 ${ }^{\text {r }}$ | \％ع＇દદ | \％8．97 | \％ど9ヤ | 」 |
| $\|10+0\|$ | ¢ıOZ | ャレOZ | $\varepsilon ⿺ 𠃊 乙$ | てıOZ | lioz | Oloz | 6002 | 800Z | LOOZ | 9002 | ¢00Z | ャ00て | ع00乙 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |


| てZ¢Z | LSS | と97 | $88 \varepsilon$ | 99 ¢ | 8てて | 89 | 98 | $\triangleright 9$ | $8 \varepsilon$ | 86 | 011 | Z6 | $\varepsilon \angle$ | W |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ¢¢てZ | 8¢¢ | ャてع | て0ع | £¢乙 | 961 | $\bigcirc 01$ | 101 | $\varepsilon!1$ | 88 | ャ | $\angle S$ | 91 | 68 | $\pm$ |
| $10+1$ | sIOZ | ャレOZ | $\varepsilon$ ıoz | てIOZ | 110Z | OLOZ | 6002 | 800乙 | LOOZ | 9002 | s00Z | ャ00て | ع00Z |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8てて | EOL | 1／ | $\varepsilon 9$ | 87 | $\bigcirc \varepsilon$ | I乙 | OZ | OZ | 8 | てし | OZ | 6 | こて | W |
| ¢ $\angle \varepsilon$ | S61 | 6て1 | 86 | $\angle G$ | Ot | ャ乙 | OZ | £乙 | $\varepsilon 1$ | てし | 01 | 11 | 61 | 」 |
| $10+1$ | sıOZ | $\checkmark$ ロOZ | $\varepsilon 10 乙$ | てIOZ | lioz | 010Z | 6002 | 800Z | LOOZ | 9002 | sooz | ャ00乙 | \＆00Z |  |



| \％L゚19 | \％く゚19 | \％ $8.1 /$ | \％s＇89 | \％8・ャ9 | \％S＇LL | \％ガレL | \％でてく | \％6＂ss | \％000s | \％6でて | \％9＇\＆ร | \％ع．$¢ 8$ | \％${ }^{\circ} 69$ | W |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| \％8＇てを | \％6＊$\angle \varepsilon$ | \％9＊$\angle 乙$ | \％s＇ıE | \％8・ャ | \％s＇てて | \％9＇8乙 | \％8＊$\angle 乙$ | \％でしヤ | \％000s | \％しゃ $\angle$ ¢ | \％ャ゙9ヤ | \％ぐ91 | \％く゚०ヤ | 」 |
| $10+01$ | ¢ıOZ | $\downarrow$－OZ | $\varepsilon ⿺ 𠃊 Z$ | てしOZ | lıOZ | O10Z | 6002 | 800Z | LOOZ | 9002 | sooz | ャ00て | ع00Z |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| \％ $8^{\text {® }}$ ¢ $\downarrow$ | \％9007 | \％0＾を | \％と＇ワワ | \％0085 | \％0＾¢ $\angle$ | \％9｀¢9 | \％0008 | \％1「で | \％000s | \％0．0s | \％6．9L | \％0008 | \％ナ゙¢9 | W |
| \％8＇\＆ऽ | \％$\angle \subset \angle S$ | \％でヤら | \％L｀乌S | \％000s | \％0＾¢ | \％ガ9を | \％000 | \％9＇ZS | \％009 | \％000 | \％l｀をて | \％000 | \％9｀もを | $\pm$ |
| $10+1$ | sioz | ャレOZ | $\varepsilon ⿺ 𠃊 乙$ | てıOZ | lıOZ | OlOZ | 600Z | 800Z | LOOZ | 9002 | sooz | $\downarrow$ ヤOZ | ع00乙 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |







| LS！ | $1 /$ | 97 | ¢ | ャて | $\angle 乙$ | $\angle$ | 8 | 8 | 乙 | 乙 | 01 | $\checkmark$ | $\angle 1$ | W |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 6 11 | 101 | 8G | $\triangleright \nabla$ | sz | 6 | $\checkmark$ | 乙 | 01 | 乙 | て | $\varepsilon$ | 1 | 6 | $\pm$ |
| $10+01$ | sıloz | ヤا0て | عıoz | てıOZ | 1102 | 0102 | 6002 | 8002 | L00Z | 9002 | s00乙 | ャ00て | ع00z |  |
| 人｜uO sə！｜dəy ग！$\downarrow \mathrm{D}_{\mathrm{d}}$ ənb！uก |  |  |  |  |  |  |  |  |  |  |  |  |  |  |




First, a 1-factor $\chi^{2}$ test of independence using the chisq.test() function in R was performed comparing the gender of each participant with whether or not they had ever posted a reply to the list. The resulting $\chi^{2}$ value, 11.476 , indicates a very strong correlation between these two variables, giving a p-value of less than 0.005 . A similar test was run by considering each email sent to the list and comparing the gender of the author with whether or not the email was sent as a reply to another message. The resulting $\chi^{2}$ value, 272.17, indicates an extremely strong correlation, with a p -value less than $2.2 \times 10^{-16}$.

Table 5 shows the contingency tables for each $\chi^{2}$ test, along with the resulting statistical information.

|  | Did Not Reply | Replied | Non-Replies | Replies |
| :--- | :--- | :--- | :--- | :--- |
| $F$ | 196 | 179 | $F$ | 1599 |
| $M$ | 86 | 142 | $M$ | 1193 |
| $\chi^{2}=11.476$, p-value $=0.0007051$ |  |  | $\chi^{2}=272.17, p-$ value $<2.2 \times 10^{-16}$ | 1329 |

Table 5. Contingency tables and statistical information related to list replies
It is important to note that two male participants replied to the list with a great deal more frequency than any other list participants, combining for 464 replies total. No other participant replied to the list more than 65 times. As can be seen in Table 6, even with these two individuals removed, the $\chi^{2}$ test still implies an extremely strong correlation between gender and both the likelihood of an individual to send a reply and the number of replies sent, with p-values of less than 0.005 and less than $2.2 \times 10^{-16}$, respectively.

|  | Did Not Reply | Replied |  | Non-Replies | Replies |
| :--- | :--- | :--- | :--- | :--- | :--- |
| F | 196 | 179 | $F$ | 1599 | 656 |
| $M$ | 86 | 140 | $M$ | 1187 | 865 |
| $\chi^{2}=10.875, p$-value $=0.0009746$ |  |  | $\chi^{2}=79.686, p$-value $<2.2 \times 10^{-16}$ |  |  |

Table 6. Contingency tables and statistical information related to list replies with two most frequent repliers removed.

It seemed a natural step to consider the same statistical tests on the data for each year. This was done but is not included in this study for two reasons: First, a $\chi^{2}$ test of independence generally requires at least five occurrences to exist in each category in order to be considered valid and, as a result, several years would have been left out of the analysis as a result. Second, the resulting $\chi^{2}$ and $p$-values did not add information to the study beyond what can be obtained through simple observation of the yearly raw data as discussed below.

## DISCUSSION

These data show a contrast in how men and women both utilize and interact with the SCHOLCOMM listserv. Overall, the average male participant posted to the list just over eleven times throughout the list's existence, almost twice the average for female participants at just over six posts. This means that, despite comprising only about $35 \%$ of the participants on SCHOLCOMM, male participants accounted for over $50 \%$ of the activity on the listserv.

This gap grows more conspicuous when considering replies to the list, which can be used as a basic measure of interaction. Again, though women accounted for over 50\% of the individuals who interacted with the list via posted replies, over $60 \%$ of the replies posted to the list were from male participants. In fact, the average number of replies sent by male participants was 9.4 as opposed to only 3.7 for female participants, a difference of almost two-thirds.

| Gender | Participants | Posts | Replies |
| :--- | :--- | :--- | :--- |
| M | 16 | 1103 | 965 |
| F | 4 | 408 | 163 |

Table 7. Summary of the 20 most frequent repliers to the SCHOLCOMM list
When ordering the SCHOLCOMM list participants by number of replies, there is a clear distinction between male and female participants. As can be seen in Table 7, of the 20 individuals who replied to the list most frequently, 16 were men. The top six men together had 682 replies to the list, more than all female list participants combined.

These observations, along with the strong statistical correlation between gender and list interaction, help to provide a more complete understanding of the gender dynamics on the SCHOLCOMM listserv and support the opinion expressed by several participants in 2016 that male voices on SCHOLCOMM are overrepresented. However, it is important to bear in mind that the observed correlation does not provide an explanation as to why men are more likely to interact on the listserv. Similarly, this study considered only the number of posts by individuals, which does not provide any information on the content, purpose, or effect of participation; suggestions for further research are outlined in the conclusion below.


Figure 1. Yearly trend data, male and female SCHOLCOMM participants and total post count


Figure 2. Yearly trend data, male and female SCHOLCOMM repliers and reply post count

Participation in SCHOLCOMM has changed significantly since its inception in 2003, however, and it is helpful to consider the data on a year-by-year basis. As can be seen in Figure 1, both the number of participants and the amount of participation on the listserv remained fairly steady from 2003 to 2010. However, an increase in the number of both can be seen beginning in 2011 and continuing through 2015.

That trend is also present when limiting to replies only, as is evident from the trend data in Figure 2. Again, the number of participants and the number of replies remain fairly
constant from 2003 to 2010, but an increase in both can be seen from 2011 onwards. Over 200 replies were posted to the list in 2011; to contrast, SCHOLCOMM received only 165 replies total from 2003 to 2010 . Though less pronounced in the number of posts, both posts and replies saw a sharp increase in 2011.

Though the identification of outliers would likely be beneficial, it would require deeper statistical analysis and would likely include many of the most frequent contributors to the SCHOLCOMM listserv due to the high number of participants with only a single post. We therefore leave this task to future researchers.

## CONCLUSION

Participants on the ALA's SCHOLCOMM listserv voiced concerns in early 2016 over the state of the list; these concerns included a perceived overabundance of male voices on the list, which participants felt discouraged contributions from other groups. This notion was immediately met with dissenting opinions and, specifically, the rebuttal that such opinions could not be taken seriously without proof. Our study substantiates the initial opinions, showing that male participants are both more active in sending out initial messages as well as in replying to threads.

Despite this, the identification of any underlying causes for the overrepresentation of male voices on SCHOLCOMM was outside the scope of this study. Though future research can and should be undertaken to better understand the dynamics of the SCHOLCOMM listserv, initial steps towards more inclusive communication can be immediately implemented by participants. List members can expressly seek contrasting opinions and contributions to contentious topics, challenge themselves to speak up if they are not frequent contributors to the list, or critically evaluate their level of activity in the SCHOLCOMM community before posting to the list. Perhaps most importantly, any list member can amplify the contributions of women, members of underrepresented groups, and fresh voices on the list by repeating their ideas and attributing them to the original author. Building in a practice of listing and requesting preferred gender pronouns (e.g., she/her/hers, they/ them/theirs) would allow participants to further actions like amplification by making underrepresented groups more clearly identifiable. This self-identification does, however, bring it with the possibility of discrimination, and therefore this is only recommended if the community has a written code of conduct prohibiting such behavior, which SCHOLCOMM does indeed have, or a stated mission involving inclusivity. ${ }^{6}$ Gender dynamics in

[^1]online communication are not fixed; such dynamics can and should be altered to better represent the diverse makeup of the community.

## Limitations and Suggestions for Further Research

As discussed above, this study's identification of unique participants as well as their gender identity was basic and likely involved some amount of error. The number of individuals as well as variations in the names of authors posed major challenges in undertaking this initial research. Similarly, because identification of participants' gender identity was undertaken indirectly, it is possible that certain participants' preferred gender identities were misrepresented.

In terms of potential future research, the authors hope that this study can serve as a basis for future analysis of online interactions in the field of scholarly communications librarianship. Expanding the statistical analysis to include position, rank, and years of experience would likely provide a more comprehensive picture of the dynamics of the SCHOLCOMM listserv. While, anecdotally speaking, no clear difference in experience level was observed during the data-gathering process, a more systematic analysis would provide a definitive picture of whether there are larger dynamics informing the findings of this current study. Studying the participation of individuals who are transgender or those who do not conform to the male/female gender binary would also provide a more comprehensive understanding of dynamics on the listserv, as well as more broadly in online communication in the field. Additionally, delving deeper into the issue of gender identity and the intersection of other underrepresented groups would provide a better understanding of the findings of this study.

While the research undertaken for this study only included the metadata associated with messages, a discourse analysis looking more closely at the actual content of the SCHOLCOMM listserv, such as the one described in McGee and Briscoe (2003), would allow for a better understanding of the actual roles and contributions of participants. The comprehensive preservation of data spanning the history of the listserv would also allow future researchers to undertake a more thorough historical analysis than is presented here, though in the years before 2011, SCHOLCOMM did not see a large amount of participation. This makes statistical analysis of many of those years through a $\chi^{2}$ test of independence impossible. The identification of outliers is similarly challenging due to the large number of list participants who posted or replied only once. As such, the simplistic statistical treatments utilized for this study leave space for future researchers to apply more sophisticated tools and to delve further into the content of the SCHOLCOMM listserv archives.

## ACKNOWLEDGMENTS

The authors would like to that Katherine G. Akers and Jamie Casbon for providing guidance at various stages in the research process.

## REFERENCES

American Library Association. (2014). ALA demographics studies. Chicago, IL: ALA Office for Research \& Statistics. Retrieved from http://www.ala.org/research/sites/ala.org.research/files/content/initiatives/ membershipsurveys/September2014ALADemographics.pdf

Association of College \& Research Libraries (2003). Principles and strategies for the reform of scholarly communication 1. Chicago, IL: ACRL Scholarly Communications Committee. Retrieved from http://www. ala.org/acrl/publications/whitepapers/principlesstrategies

Beveridge, A. A., Weber, S., \& Beveridge, S. (2011). Librarians in the United States from 1880-2009 [Web log post]. Retrieved from http://blog.oup.com/2011/06/librarian-census/

Gul, S., Shah, T. A., Hamade, S. N., Mushtaq, R., \& Koul, I. (2014). Effects of gender in library and information science research: A case study of The Electronic Library. The Electronic Library, 34(3), 488-503. https://doi.org/10.1108/EL-08-2014-0126

Herring, S. C. (1992). Gender and participation in computer-mediated linguistic discourse. Paper presented at the Linguistic Society of America, Philadelphia, PA. Retrieved from ERIC database. (ED345552)

Kane, L., \& Meadows, A. (2016). From the podium to the boardroom: Encouraging gender parity in scholarly publishing. Learned Publishing, 29(2), 135-138. https://doi.org/10.1002/leap. 1022

Kyrillidou, M., \& Mark, Y. (Eds.). (2006). ARL annual salary survey 2005-06. Washington, D.C.: Association of Research Libraries. Retrieved from http://www.libqual.org/documents/admin/2012/ARL_ Stats/2005-06ss.pdf

McGee, P., \& Briscoe, F. (2003). Discourse analysis: Power/knowledge on an academic listserv. Journal of Information, Communication \& Ethics in Society, 1(3), 133-148. https://doi. org/10.1108/14779960380000232

Smith, K. (2016, March 27). Content analysis. Message posted to http://lists.ala.org/sympa/arc/ scholcomm/2016-03/msg00200.html

Sugimoto, C. R., Tsou, A., Naslund, S., Hauser, A., Brandon, M., Winter, D.,...Finlay, S. C. (2012). Beyond gatekeepers of knowledge: Scholarly communication practices of academic librarians and archivists at ARL institutions. College \& Research Libraries, 75(2), 145-161. https://doi.org/10.5860/crl12-398
van Doorn, N., \& van Zoonen, L. (2009). Theorizing gender and the internet: Past, present, and future. In A. Chadwick \& P. N. Howard (Eds.), Routledge handbook of Internet politics (pp. 261-274). New York, NY: Taylor \& Francis.

West, J. D., Jacquet, J., King, M. M., Correll, S. J., \& Bergstrom, C. T. (2013). The role of gender in scholarly authorship. PLoS One, 8(7), e66212. https://doi.org/10.1371/journal.pone. 0066212


[^0]:    ${ }^{1}$ https://scholarlykitchen.sspnet.org
    ${ }^{2}$ http://blogs.library.duke.edu/scholcomm/
    ${ }^{3}$ https://blogs.libraries.indiana.edu/scholcomm/
    ${ }^{4}$ http://intheopen.net/
    ${ }^{5}$ http://www.ala.org/acrl/issues/scholcomm/scholcommdiscussion

[^1]:    ${ }^{6}$ http://www.ala.org/acrl/issues/scholcomm/scholcommdiscussion

