Trends, costs, benefits, challenges, and prognoses for supplementary materials

Authors: Shutler, Dave, and Murray, Ashley

Source: The Auk, 133(4) : 733-737

Published By: American Ornithological Society

URL: https://doi.org/10.1642/AUK-16-92.1

The Auk
Ornithological Advances
AmericanOrnithology.org

Volume 133, 2016, pp. 733–737
DOI: 10.1642/AUK-16-92.1

COMMENTARY

Trends, costs, benefits, challenges, and prognoses for supplementary materials

Dave Shutler* and Ashley Murray

Department of Biology, Acadia University, Wolfville, Nova Scotia, Canada
* Corresponding author: dave.shutler@acadiau.ca

Submitted May 14, 2016; Accepted June 20, 2016; Published August 31, 2016

ABSTRACT

Supplementary materials (SM) are a relatively new addition to published research. The volume of SM has been increasing over time for many journals, and in a few cases SM appear to be replacing appendices. We review the costs and benefits of SM to authors, reviewers, production teams, and readers. We conclude that SM can have value but that they place additional demands on all participants. Given the difficulty of recruiting reviewers for journal submissions and the added effort required by all participants, this may lead to incomplete vetting that may undermine the reliability of SM in some instances.

Keywords: appendices, dead links

Tendencias, costos, beneficios, desafíos y prognosis del material suplementario

RESUMEN

Los materiales suplementarios (MS) son un agregado relativamente nuevo a las publicaciones científicas. El volumen de MS ha ido aumentado a lo largo del tiempo en muchas revistas y en algunos pocos casos parece estar reemplazando a los apéndices. Revisamos los costos y beneficios del MS para los autores, los revisores, los equipos de producción y los lectores. Concluimos que el MS puede tener valor, pero que plantea demandas adicionales a todos los participantes. Dada la dificultad de reclutar revisores de las presentaciones a las revistas y el esfuerzo adicional requerido de todos los participantes, esto puede dar lugar a una revisión incompleta que puede vulnerar la confiabilidad del MS en algunos casos.

Palabas clave: apéndices, enlaces muertos

The world of scientific publishing is undergoing exponentially increasing rates of upheaval, as exemplified by sinister, "predatory" journals (Butler 2013). One upheaval in "legitimate" publishing has been the increased relegation of information to supplementary materials (SM). Most of us have encountered SM, although perhaps not in ornithological journals, and some have wondered about these materials. Our intent here is to provide an overview of SM for ornithologists and other interested readers. We begin with an assessment of trends in SM in a subset of ornithological and interdisciplinary journals, and we follow that with a critical overview of the limited literature that considers this new frontier. We distinguish between appendices, which are properly attached to a paper, and SM, which are online.

Trends

We chose a set of journals that many ornithologists read: *The American Naturalist*, *Animal Behaviour*, *The Auk*, *Behavioral Ecology*, *BMC Ecology* (online only), *The Condor*, *Ecology*, *Ibis*, *Journal of Avian Biology*, *Oecologia*, *Oikos*, *PLoS Biology* (online only), and *Proceedings of the National Academy of Sciences* (biology articles only). To capture trends, we chose the years 2000 (but 2001 for *BMC Ecology*, the year it was first published), 2005, 2010, and 2015, recording pages of appendices and pages of SM for each article to evaluate whether SM were replacing appendices. Initially we reviewed every article in *Animal Behaviour* for 2015, but, because of time constraints, we thereafter subsampled only 2 randomly chosen issues within the specified years. Our initial forays revealed that many journals did not begin having SM until after 2000; thus, to avoid wasted effort, we began in 2005 for *Behavioral Ecology*, *Oecologia*, *Oikos*, *Proceedings of the National Academy of Sciences*, and *PLoS Biology*. We reviewed only full papers that reported original data; commentaries, reviews, and short communications were not considered. We evaluated trends with general linear models where the response variable was either the log of the number of pages of appendices + 1 or the log of the

TABLE 1. Comparison of journal use of appendices and supplementary materials evaluated for a subset of articles from 2000, 2005, 2010, and 2015. Statistics are from general linear models where year was a class explanatory variable, the total number of pages we reviewed for a journal in a year was a covariate, and pages of appendices or pages of supplementary materials were the response variables. A dash (–) indicates that there were no instances of the response variable. Letters denote least square means (LS-means) comparisons of years in temporal sequence; if there are only 3 letters, 2000 was not evaluated. A's are associated with the highest means, followed by B, followed by C. Years sharing letters did not differ in numbers of pages devoted to appendices or supplementary materials.

		Appendices				Supplementary materials			
Journal	Articles (*n*)	R^2	*F*	*P*	LS-means comparison	R^2	*F*	*P*	LS-means comparison
The American Naturalist	89	0.34	14.5	<0.0001	A,B,B,B	0.31	1.6	<0.0001	B,A,A,A
Animal Behaviour	414	0.25	10.1	<0.0001	B,B,B,A	0.03	3.6	0.01	B,B,B,A
The Auk	156	0.12	1.6	0.06	AB,B,A,A	0.03	3.8	0.18	AB,B,AB,A
Behavioral Ecology	149	0.08	2.4	0.09	A,B,B	0.25	22.7	<0.0001	B,B,A
BMC Ecology	49	–	–	–	–	0.25	6.5	0.003	B,B,A
The Condor	150	0.15	3.1	0.03	B,B,B,A	–	–	–	–
Ecology	201	0.19	14.7	<0.0001	B,A,A,B	0.67	130.8	<0.0001	B,B,B,A
Oecologia	163	0.05	3.3	0.04	A,B,B	0.31	23.5	<0.0001	C,B,A
Oikos	110	0.11	2.6	0.08	AB,A,B	0.57	66.3	<0.0001	B,B,A
PLoS Biology	77	–	–	–	–	0.44	0.1	0.26	A,A,A
Proceedings of the National Academy of Sciences [a]	91	–	–	–		0.61	61.9	<0.0001	C,B,A

[a] Biology articles only.

number of pages of SM + 1, year was a class explanatory variable, and log of the number of pages in an article was a covariate.

In total we reviewed 1,876 articles, of which 262 (14.0%) had appendices and 373 (20.0%) had SM. Although there were significant temporal trends in numbers of pages of appendices for 7 journals, there was no clear, overall pattern of decreases or increases (Table 1). By contrast, 10 journals had statistically significant positive trends in SM over time, including one exclusively online journal (Table 1). These findings are consistent with previous studies on different sets of journals (Schaffer and Jackson 2004, Evangelou et al. 2005, Schriger et al. 2011, Kenyon and Sprague 2014, Rafferty et al. 2015, Williams 2016); in fact, in some cases, journals now have SM associated with 100% of their articles (Borowski 2011, Pop and Salzberg 2015). SM were less often found in ornithological journals, particularly in North America: 1 of 156 (0.6%) articles in *The Auk*, 0 of 150 articles in *The Condor*, 21 of 116 (18.1%) articles in *Ibis*, and 9 (all in 2015) of 111 (8.1%) articles in the *Journal of Avian Biology*. Finally, for 4 journals, we found evidence that as pages of appendices decreased, pages of SM increased (Table 1).

Another aspect of SM is content; with the exceptions of audio and video files, allocation of types of materials within appendices and SM was broadly similar (Figure 1). Schriger et al. (2011) found that the number of tables was the fastest-growing content in SM; we did not evaluate trends.

Costs

The first cost of SM is to authors. In the past, authors summarized relevant data as succinctly as possible in their papers, and filed data of less relevance. With the option of SM, authors now must judge whether less relevant data are worth sharing with readers (Kenyon and Sprague 2014, Pop and Salzberg 2015). If authors opt for SM, they must then invest in the organization and production of these

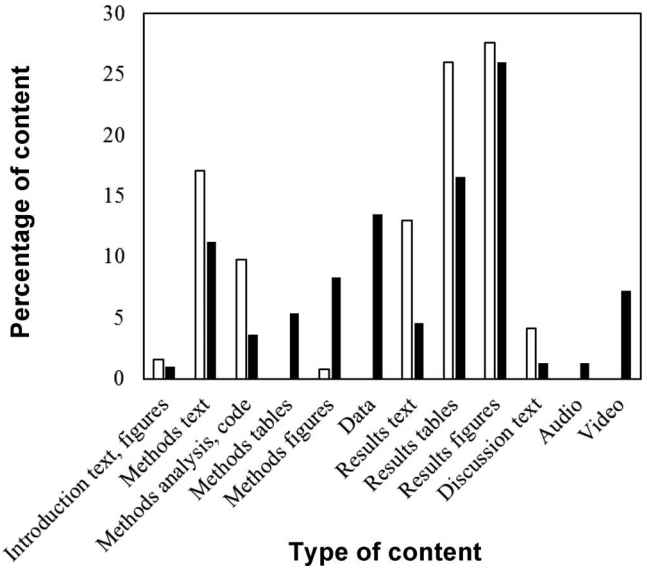

FIGURE 1. For a subset of articles, percentages of each type of content found in 262 articles (white bars) that had appendices and 373 articles (black bars) that had supplementary materials.

materials (Carpenter 2010), investments that may be more than trivial. To reduce these investments, authors may not follow due diligence, which may result in poor-quality SM. Moreover, if authors invest insufficiently in assessing the value of SM, irrelevant and overly detailed information could become entrenched in a burgeoning digital archive; this would be detrimental to science (Pop and Salzberg 2015). As an illustration, in a recent review of publicly mandated data-archiving, Roche et al. (2015) reported that >50% of data may not be usable. Another cost, on the flip side, is that busy readers may deem SM of limited value or too costly, in terms of time, to access (Rosenthal and Reich 2010); thus, authors may invest significant effort for low return. Formatting inconsistencies among journals are an irritation for many of us who publish; reformatting each time one resubmits a paper is painstaking. Now, with SM, a whole new set of guidelines may have to be reviewed; these may include size limits, file formats, and so on (NISA/NSAIS 2013). All of this adds additional time to the submission process. Also, if a reviewer decides to pursue a link that doesn't work (see below), this can delay peer review (Brooks and Markwell 2008). At one extreme, ubiquitous inclusion of SM in some journals implies that SM are required, which places additional burdens on authors, reviewers, production teams, and readers.

The second cost of SM is to reviewers. Some journals request that SM be reviewed along with associated articles (Pop and Salzberg 2015). Reviewers may refuse, arguing against further burdening what is essentially volunteerism. If reviewers can be convinced to evaluate SM, it could increase turnaround times on papers (particularly if a link doesn't work because it was entered improperly), which is neither in authors' nor in journals' best interests (Brooks and Markwell 2008, Borowski 2011). Reduced motivation for authors to invest in editing SM, and reluctance by reviewers to vet, can lead to poor-quality SM of questionable scientific merit. Maunsell (2010) cited these issues in an editorial in which her journal announced the end of SM. By contrast, Marcus (2009) argued that any limit to SM is arbitrary. We suspect that there is a continuum in how researchers view this.

The third and perhaps most significant cost of SM is to production teams (editors and copyeditors). One of the biggest costs to journals can be for copyediting, and there is always pressure to reduce costs. Thus, production teams may decide not to review SM (Kenyon and Sprague 2014). If we layer this atop inadequate attention from authors and reviewers, it presents another threat to the quality of SM. An additional production cost is for curation of SM, but this may not be a dire expense, given that digital storage space continues to come down in price.

As we prepared this commentary, we became most acutely aware of the fourth cost of SM, which is to readers. First, navigating SM can be extremely time consuming (see Anonymous 2015). For example, there is significant variation in how one accesses SM at different journals; one may have to scroll to find links, and multiple links are sometimes required to get to SM. Second, if storage space is not maintained (e.g., because a journal goes bankrupt, leading to link rot) or if uniform resource locators (URLs) are entered incorrectly (see above), dead links may make it impossible to retrieve SM (Anderson et al. 2006, Aronsky et al. 2007, Brooks and Markwell 2008, Nagaraja et al. 2011). Third, the sheer volume of material can be staggering, making it difficult or impossible to extract what is relevant (Pop and Salzberg 2015). It is already a daunting challenge to stay atop even a subset of the literature; this is a truly dark side of cheaply accessible information. In the journals we searched, it was not unusual to find dozens of pages of SM per paper (Figure 2). In an added irony, we found reference to SM within SM! Pop and Salzberg (2015) cite a few particularly amazing extremes: 165 pages of SM for a 6-page paper, and 144 pages of SM for a 5-page paper. The SM in the latter case had multiple sections, with authors listed in different orders. Fourth, SM can be a minefield of storage formats, and in some instances require downloading unfamiliar software with which a reader has to become conversant (Aronsky et al. 2007, Pop and Salzberg 2015). Fifth, one never knows whether SM are worth tracking down, for reasons already given. Sixth, readers who prefer to work from hard copies of papers or from print subscriptions lose that option when they have to return to the Internet to acquire SM. SM becomes unavailable to readers who venture to places where Internet access is unavailable or too expensive. In these circumstances, even if one works from on-screen versions of papers, access to potentially relevant SM is lost. Seventh, it is not uncommon to find that authors refer within the paper to only a fraction of associated SM links (Pop and Salzberg 2015). Do we allow

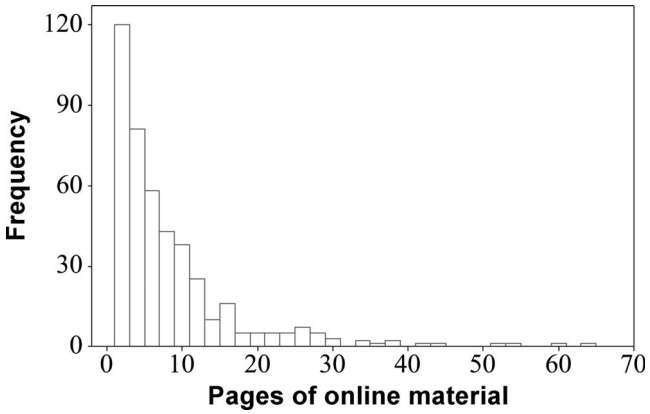

FIGURE 2. Frequency distribution (excluding 1,433 zeros) of the numbers of pages of supplementary materials for 443 articles we evaluated.

our students to leave tables and figures uncited in their lab reports?

A somewhat distinctive cost of SM is that they may contain citations that do not appear in an associated paper; these citations may not be captured by search engines—and, hence, authors may not be credited for their contributions to research (Rafferty et al. 2015). This may affect a researcher's competitiveness in seeking positions or promotions, or in applying for grants. However, this becomes a problem only if there are biases in the kinds of research that are relegated to SM. Nonetheless, this latter trend should continue to be monitored.

Benefits

We live in an era of phenomenal growth in data acquisition. In the past, this would have created a dilemma for authors because large volumes of data can be difficult to summarize. Now, authors have the option of providing many more summaries of information than may be necessary in a paper, but that may be of interest to part of a paper's readership; authors have virtually unlimited capacity to link information they deem relevant. This may include raw data and programs to evaluate the data in the same manner as the authors. Retaining all this information in a paper may impede interpretation for the majority of its audience. Second, journals' adoption of SM has freed authors (in some cases) from storing those materials on their personal websites; this is likely to reduce link rot (Anderson et al. 2006). Third, a variant of the file-drawer problem— wherein only statistically significant results are presented in detail—may be addressed by SM. Authors now have the option of depositing statistically nonsignificant results in SM. Fourth, data can be hard-won, and even if the data have no immediate purpose, it would be wasteful to discard them, assuming they are of reasonable quality (although such data may be deposited in locations other than SM; e.g., Dryad, http://datadryad.org/). Long-term repositories may pay substantial dividends; indeed, many who have had to provide their data to accompany publications have received notice that someone is using the data to support additional research.

A potential benefit to reviewers of SM is that they can access more information than is provided in a terse paper, possibly facilitating their ability to evaluate scientific merit (Borowski 2011). On the other hand, if freed from reviewing SM, a reviewer may need less time to complete a review of a paper that would otherwise have contained volumes of peripheral information.

Production teams may incur substantial benefits from SM if they are not tasked with formatting the materials; this is often the case with some raw data that appear in SM (e.g., csv or mpeg files; Carpenter 2009). Indeed, SM appear in a variety of forms that have limited or inconsistent formatting. An additional benefit, primarily for print journals, is that costs of printing are saved, and these can be substantial for color images or voluminous SM (Evangelou et al. 2005, Carpenter 2009). Finally, the more SM a journal produces, the greater its web presence. However, we are not convinced that this latter perk will have important consequences in the mushrooming cloud of web volume.

One of the biggest benefits to readers may be from having access to SM that are too expensive to produce in print, such as high-quality images, or SM that cannot be produced in conventional print format, such as sound and video files (Figure 1). Of all the SM we encountered, sound and video files had the most immediate impact for the least investment on our part. A second benefit to readers is flexibility in how one reads a paper. Granted, one can always skip sections, but if all SM appeared within the body of the paper, sheer length might discourage readers from even attempting to read it, and readers might become distracted from the main thread. A paper that has only the barest essentials may flow much better (Pop and Salzberg 2015).

Challenges and Prognoses

We believe that there is a place for SM, but it needs to be carefully considered at every step, and this creates dilemmas for everyone involved. If we don't confront these dilemmas, we risk accepting poor-quality SM that will deter scientific advances. As we skimmed SM in >400 papers, we observed the gamut of inaccessible to wonderful SM. An optimistic view is that wonderful SM will become the norm and will follow guidelines provided by NISO/NFAIS (2013), including recommendations for improving links. All SM links should appear in both the print and online versions of a paper; be cited within the paper where the information is relevant, as well as at the end of the paper; and be included in the table of contents (Williams 2016). For readers who disconnect, it might be valuable if the default were for SM to download along with the paper (Marcus 2009, Laue 2010), which would alert readers who want to print only the paper to follow attendant instructions. One reviewer of this commentary suggested that journals indicate whether their SM have been peer reviewed, a view we share.

One can only wonder about the future of scientific publishing and whether this is all moot. Will there always be publications, or will we be carrying around multimedia, multidimensional synopses of research with interactive data interfaces (Carpenter 2009)? Brave new world.

ACKNOWLEDGMENTS

We thank A. Mullie for valuable suggestions that improved the manuscript, M. Mallory and other colleagues with whom we have discussed it, and two fine reviewers.

LITERATURE CITED

Anderson, N. R., P. Tarczy-Hornoch, and R. E. Bumgarner (2006). On the persistence of supplementary resources in biomedical publications. BMC Bioinformatics 7:260.

Anonymous (2015). 50 years ago. Nature 522:427.

Aronsky, D., S. Madani, R. J. Carnevale, S. Duda, and M. T. Feyder (2007). The prevalence and inaccessibility of Internet references in the biomedical literature at the time of publication. Journal of the American Medical Informatics Association 14:232–234.

Borowski, C. (2011). Enough is enough. Journal of Experimental Medicine 208:1337.

Brooks, D. W., and J. Markwell (2008). Evaluating web-based information: Access and accuracy. Journal of Chemical Education 85:458–459.

Butler, D. (2013). Investigating journals: The dark side of publishing. Nature 495:433–435.

Carpenter, T. (2009). Standards column—Journal article supplementary materials: A Pandora's box of issues needing best practices. Against the Grain 21(6):6.

Carpenter, T. (2010). Outside the core: Working towards an industry recommended practice for supplemental journal materials. Serials 23:155–158.

Evangelou, E., T. A. Trikalinos, and J. P. A. Ioannidis (2005). Unavailability of online supplementary scientific information from articles published in major journals. The FASEB Journal 19:1943–1944.

Kenyon, J., and N. R. Sprague (2014). Trends in the use of supplementary materials in environmental science journals. Issues in Science and Technology Librarianship 75.

Laue, A. (2010). Hosting supplementary material: Technical challenges and suggested best practices. Information Standards Quarterly 22(3):10–15.

Marcus, E. (2009). Taming supplemental material. Cell 139:11.

Maunsell, J. (2010). Announcement regarding supplemental material. The Journal of Neuroscience 30:10599–10600.

Nagaraja, A., S. A. Joseph, H. H. Polen, and K. A. Clauson (2011). Disappearing act: Persistence and attrition of uniform resource locators (URLs) in an open access medical journal. Program: Electronic Library and Information Systems 45:98–106.

NISO/NFAIS (2013). Recommended practices for online supplemental journal article materials. National Information Standards Organization, Baltimore, MD, USA. http://www.niso.org/publications/rp/rp-15-2013

Pop, M., and S. L. Salzberg (2015). Use and mis-use of supplementary material in science publications. BMC Bioinformatics 16:237.

Rafferty, A. R., B. B. M. Wong, and D. G. Chapple (2015). An increasing citation black hole in ecology and evolution. Ecology and Evolution 5:196–199.

Roche, D. G., L. E. B. Kruuk, R. Lanfear, and S. A. Binning (2015). Public data archiving in ecology and evolution: How well are we doing? PLOS Biology 13:e1002295.

Rosenthal, D. S. H., and V. A. Reich (2010). Archiving supplemental materials. Information Standards Quarterly 22(3):16–21.

Schaffer, T., and K. M. Jackson (2004). The use of online supplementary material in high-impact scientific journals. Science & Technology Libraries 25(1–2):73–85.

Schriger, D. L., A. C. Chehrazi, R. M. Merchant, and D. G. Altman (2011). Use of the Internet by print medical journals in 2003 to 2009: A longitudinal observational study. Annals of Emergency Medicine 57:153–160.

Williams, S. C. (2016). Practices, policies, and persistence: A study of supplementary materials in crop science journals. Journal of Agricultural & Food Information 17:11–22.