

Biological science  
practices

**Cite this article:** Martínez A, Mammola S.

2021 Specialized terminology reduces the number of citations of scientific papers.

*Proc. R. Soc. B* **288**: 20202581.

<https://doi.org/10.1098/rspb.2020.2581>

Received: 15 October 2020

Accepted: 12 March 2021

**Subject Category:**

Ecology

**Subject Areas:**

environmental science, cognition, evolution

**Keywords:**

glossary, linguistics, scientific communication, scientific writing, specialized vocabulary, terminology

**Author for correspondence:**

Alejandro Martínez

e-mail: [amartinez.ull@gmail.com](mailto:amartinez.ull@gmail.com)

<sup>†</sup>Equal contribution.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5356446>.

## Specialized terminology reduces the number of citations of scientific papers

Alejandro Martínez<sup>1,†</sup> and Stefano Mammola<sup>1,2,†</sup>

<sup>1</sup>Molecular Ecology Group (Dark-MEG), Water Research Institute (IRSA), National Research Council (CNR), Verbania Pallanza, Italy

<sup>2</sup>Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History (LUOMUS), University of Helsinki, Helsinki, Finland

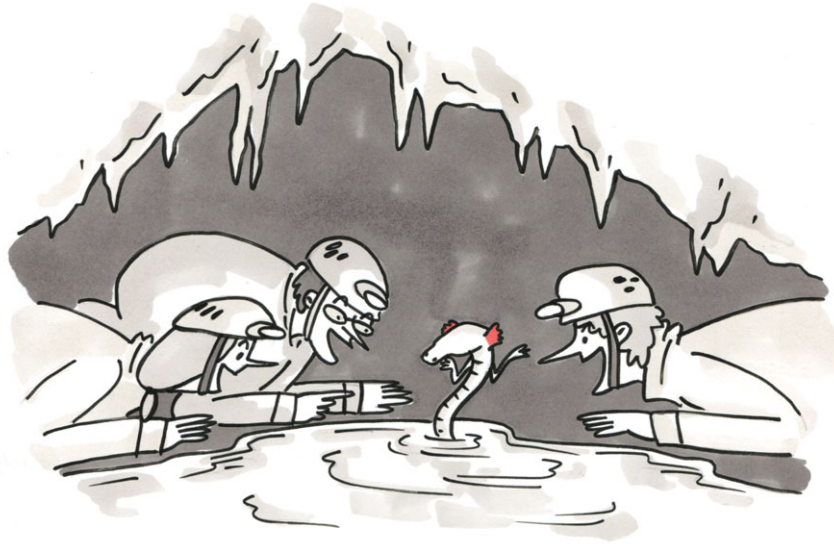
AM, 0000-0003-0073-3688; SM, 0000-0002-4471-9055

Words are the building blocks of communicating science. As our understanding of the world progresses, scientific disciplines naturally enrich their specialized vocabulary (jargon). However, in the era of interdisciplinarity, the use of jargon may hinder effective communication among scientists that do not share a common scientific background. The question of how jargon limits the transmission of scientific knowledge has long been debated but rarely addressed quantitatively. We explored the relationship between the use of jargon and citations, using 21 486 articles focusing on cave research, a multidisciplinary field particularly prone to terminological specialization, and where linguistic disagreement among peers is frequent. We demonstrate a significant negative relationship between the proportion of jargon words in the title and abstract and the number of citations a paper receives. Given that these elements are the hook to readers, we urge scientists to restrict jargon to sections of the paper where its use is unavoidable.

## 1. On jargon and the reach of knowledge

A stumbling toddler babbling ‘mummy’ and a famous scholar writing a 500-page lifetime essay have at least one thing in common: they both navigate reality through words. We all do, as much as we can speak or read. ‘The limits of my language mean the limits of my world’, believed philosopher Ludwig Wittgenstein [1], and not surprisingly, our education is largely devoted to learning new terms and their meanings [2]. Whether it is a zoologist defining a white blind salamander as a *neotenic metazoan with anophthalmia*, or a geologist describing marble as a *metamorphic rock produced by the recrystallization of calcite or dolomite*, the importance of specialized terminology (jargon) is undisputed. Jargon, although difficult at first, condenses years of knowledge into a precise mental image (e.g. [3–5]). *Metazoan* depicts a *multicellular eukaryote*; hence an organism consisting of multiple cells with a *nucleus*; which brings up, if we understand this jargon, images of membranous structures containing the salamander’s genetic information. In this example, as in a Russian doll, each new term enriches the initial message with information, structuring and systematizing concepts into the corpus of Science [6].

However, words are not Science. Physicist Richard Feynman [7] believed that learning the meaning of words only informs about the limit of people’s imagination, but nothing about nature itself. What he meant was that understanding *what is going on* is more important than knowing the name of things or processes—even though he admitted that words are needed to communicate. Likewise, the late Wittgenstein pointed at the futility of using a ‘private language’ that prevents us from connecting with other people [8]. In a similar way, jargon is restrictive because it illustrates complex concepts only in the minds of those sharing a common background while precluding everyone else from understanding [9]. When stepping out of their linguistic comfort zone, readers might not understand the jargon at all, get the message only partially or, after an extra mental effort, figure out its meaning. In other words,



'What a remarkable troglobiomorphic species!'

'In my opinion, it's an obligate hypogeobiont ...'

'Are you kidding me? It's obviously an eu-styophile!'

**Figure 1.** Jargon underground. Another day of terminological debate in the cave-office. Since the nineteenth century, much ink has been spilled to discuss terminological nuances regarding the ecological classification of the subterranean fauna [19–21], even though such classifications are just an attempt to simplify nature's complexity, which is far from being rigidly defined [22]. Illustration by Irene Frigo (<https://www.instagram.com/irene.frigo/>). (Online version in colour.)

upon reading about the *neotenic metazoan* above, one reader might be able to picture an olm, another might imagine some sort of weird animal, and a third might throw up their arms in despair before googling the meaning of 'neotenic' and 'metazoan'. It is within this range of differential understanding where the evil of jargon abuse expresses itself in all its glory: not only does it reinforce the distinction between a geologist and a zoologist, but potentially, it also divides zoologists into an endless number of subgroups. In conclusion, jargon may communicate ideas powerfully, but also, and perhaps more often, artificially define 'insiders' and 'outsiders', reinforcing the isolation of academics within their respective ivory towers.

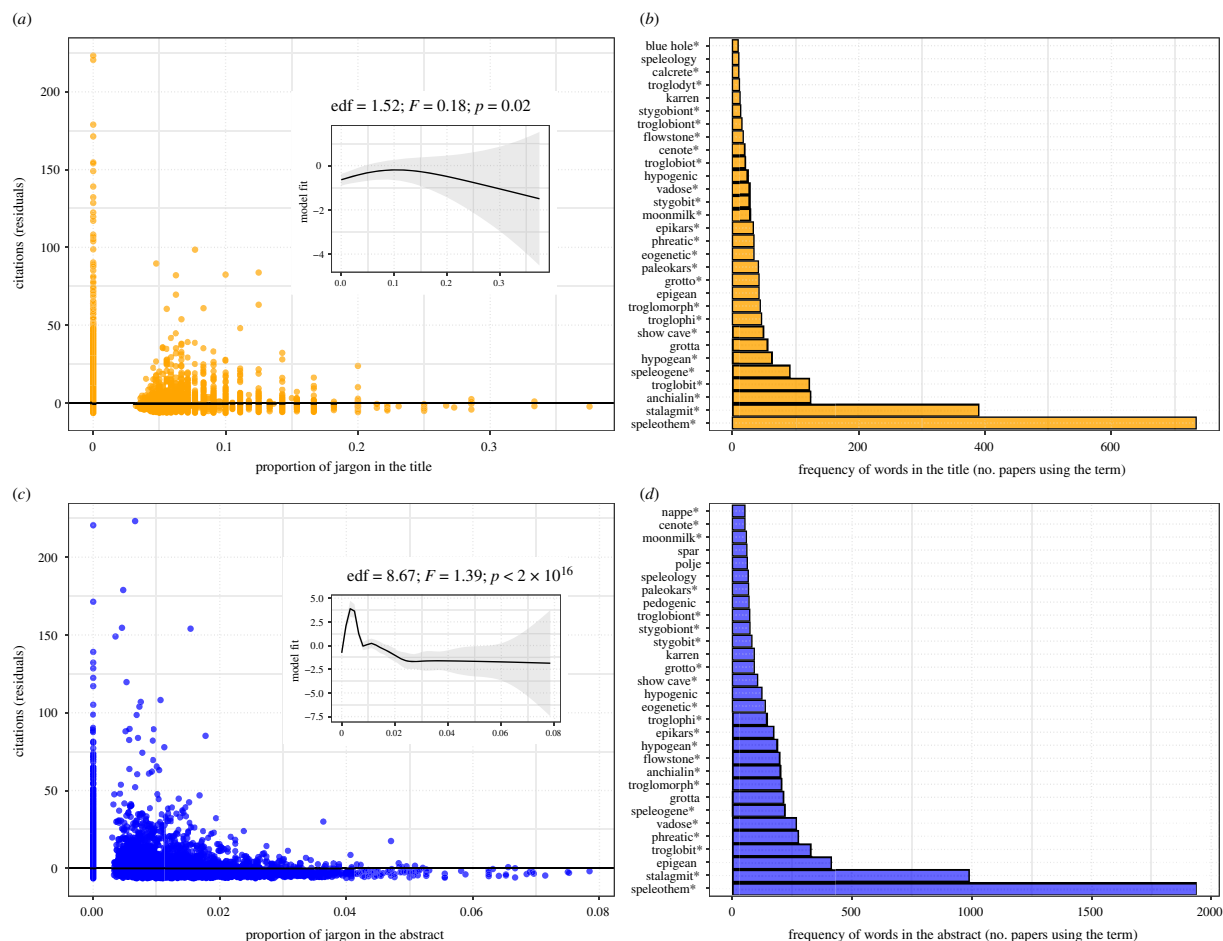
Researchers from disparate disciplines have warned us about the dangers of jargon abuse [9–14]. 'Never use a [...] jargon word if you can think of an everyday English equivalent', George Orwell [15] famously stated. Therefore, it seems important to quantitatively analyse the effects of jargon use in science, and more specifically, how it may limit the reach of scientific knowledge. Recently, Plavén-Sigra *et al.* analysed the abstract of greater than 700 000 articles across 12 sub-disciplines of life and medical sciences, showing that an increase in the use of jargon decreases the readability of texts [16]. This study, however, focused on general scientific jargon such as 'robust', 'therefore' and 'underlying'. All these terms have become an integral part of scientists' writing routine and thus are unlikely to undermine communication among scientists. What remains to be quantified is the role of discipline-specific jargon in driving the impact of a paper across scientists with different backgrounds. Such an analysis could hardly be performed on a broad multidisciplinary database such as the one used by Plavén-Sigra *et al.* [16], because the diversity of specialized terms and the factors affecting their use vary too much across disciplines. These two confounding factors, however, can be alleviated by

looking at the literature within a single multidisciplinary community of scientists. We hypothesize that, in such a context, jargon papers will be less understood, remembered and ultimately cited.

## 2. Effect of jargon on citations: a case study

Since the early twentieth century [17], cave research has been the colliding point of generations of scientists with diverse scientific backgrounds [18]. Geologists, zoologists, anthropologists, ecologists and evolutionary biologists have interacted in the darkness of caves populating 120 years of cave literature with a maze of specialized terms, either borrowed from their scientific backgrounds or just coined *ex novo* using diverse etymological roots. The lack of terminological agreement among cave scientists has preserved most of these words, which are still commonly found in the literature and are the central subject of heated etymological debates (figure 1). We took advantage of the long tradition of multidisciplinary [23,24] and high terminological specialization [25] offered by cave literature to investigate the effect of jargon use on article success—measured as the number of citations.

In the Web of Science (Clarivate analytics), we sourced 21 486 research articles on cave environments published over the last 30 years. By using a curated selection of keywords, we ensured that we cover articles dealing with caves published in both cave-specific and general international journals. Given that older papers had more time to accumulate citations, we normalized the number of citations by article age. We predicted the expected pattern of citations over time using a Poisson generalized additive model, and expressed the number of citations of each individual paper as the Pearson residual from the curve [26] (electronic supplementary material, appendix S1). To define discipline-



**Figure 2.** Effect of jargon use on citations and most frequently used jargon. (a,c) Relationship between the proportion of jargon and citations. The number of citations for each article is normalized by its age, expressing it as Pearson's residuals from the regression curve (electronic supplementary material, figure S1 in appendix S2) representing the predicted number of citations over time [26]. Dots below the horizontal black line are articles under-cited for their age, and vice versa. Insets show the predicted trend based on generalized additive mixed models, with a random structure to account for the similarity of jargon between articles published in the same subject area and the variation of jargon through time (details in electronic supplementary material, appendix S2). (b,d) Thirty most frequently used jargon terms in titles and abstracts. Using the regular expression notation, the asterisk (\*) at the end of the words is a metacharacter for zero or more instances of the preceding characters (e.g. 'speleogenom\*' matches 'speleogenome', 'speleogenomics', etc.). (Online version in colour.)

specific jargon, we manually assembled a comprehensive list of roughly 1500 words using glossaries of books focused on caves, reviews and other sources (full list in electronic supplementary material, appendix S2). We calculated the proportion of jargon in the title and abstract of each article relative to the total number of words. We focused on titles and abstracts given that these elements are the hook to readers [27] and reflect the overall writing style of entire articles [16]. A detailed description of the methods and analyses is in electronic supplementary material, appendix S1.

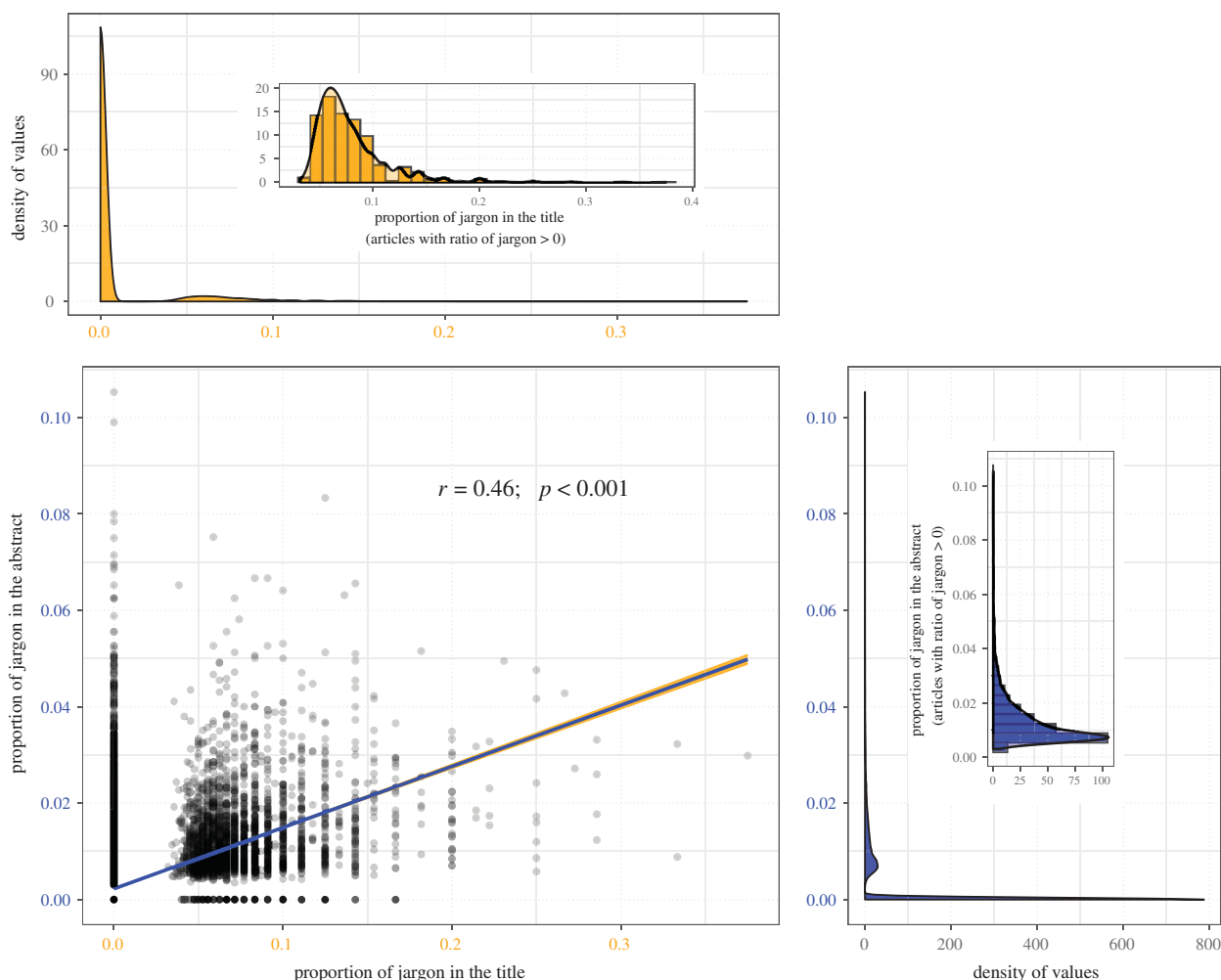
We observed a negative and nonlinear relationship between the number of jargon words and the normalized number of citations, which significantly decreased as the proportion of jargon in the title (figure 2a) and abstract (figure 2c) increased. This trend was particularly evident in abstracts, with a sudden drop in citations when the proportion of jargon was above 1% (figure 2c, inset). Interestingly, none of the highly cited papers (Pearson's residuals of citations greater than 100, corresponding to greater than 450 citations) used jargon in the title, and almost all highly cited papers had a proportion of jargon in the abstract below 1%. This trend was recovered even when we restricted the analysis to either specialist cave journals

or multidisciplinary ones (electronic supplementary material, figure S2).

All in all, the type of specialized words used in the abstract and title were similar (cf. figure 2b,d). Although we did not analyse the effect of individual words on citations, we noted that the most widely used terms (e.g. stalagmite, epigean) were seemingly less detrimental for the citation success of a paper than terms occurring sporadically. This may be because these frequently used specialized terms (figure 2b,d) are being integrated into the vocabulary of most scientists. We also found a positive correlation between the use of jargon in the title and abstract, with only about one-third of articles using jargon in the abstract also including jargon in the title (figure 3).

### 3. Concluding remarks

While our analysis does not inform about the epistemological basis driving the choice for one word or another, it clearly emphasizes the negative effect of jargon on the success of a paper. With an estimated annual rise in the volume of scientific articles between 4% and 9% per year [28,29], and over 5 million papers published in 2019 alone (*Dimensions* database,



**Figure 3.** Correlation between the proportion of jargon in the abstract and title. Correlation is based on Pearson's  $r$ . Density plots show the distribution of the proportion of jargon values, obtained by computing a kernel density estimate. Insets are the distribution of the proportion of jargon values only for those articles with abstract or titles including jargon (i.e. proportion greater than 0). (Online version in colour.)

accessed on 30 December 2020), there is increasing pressure to publish papers that stand out amid so many others [30,31]. A global estimate pointed out that scientists skim an average of over 1100 titles and 200 abstracts a year, but they go on to read 97 full texts [27]. This suggests that the stylistic features of titles and abstracts act as important filters [32–34]: if overuse of jargon prevents a reader from understanding the message of a paper, this paper is unlikely to end up being among the 97 chosen few. Given that the title and abstract bait readers' interest, scientists might want to restrict jargon use to sections of the paper where its use is unavoidable.

As a corollary, it must be noted that in the Internet era, thanks to modern search tools such as online databases, bibliographic portals and hyperlinks, there have been significant shifts in the way scientists retrieve papers [35,36]. A high frequency of specialized terms in the abstract and title of a given paper may further reduce its retrievability, insofar as it will less frequently appear in online searches. This is problematic because the better visibility a paper has in engine optimization search results, the more likely it is to attract readers and garner citations. This, in turn, makes the case for favouring accessible and widely searched words in the abstract, title and keywords.

Not without irony, we would like to conclude by introducing a new jargon ourselves: the 'Wittgensteinian shortfall'. Since words are tools to communicate ideas, let this obscure

combination of terms be used to raise awareness of the problems previously discussed by associating the philosophical ideas of the late Wittgenstein [8] with the shortfall metaphor frequently used in ecology [37]. We define the Wittgensteinian shortfall as the inability to successfully communicate specific ideas across different scientific communities. These different communities could thus be seen as characterized by different language games. By introducing this definition, we hope to create a framework—a 'secret language' [38]—that we want to share with our *exclusive* readers to promote future discussions on this topic. After all, each community needs a password to identify its members and allow them to engage in the same *word games* [8]. Except that in this case, we believe even outsiders will appreciate the irony in this.

**Data accessibility.** All the data and scripts used in this manuscript are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.9w0vt4bdp> [39].

**Authors' contribution.** A.M. sourced articles in Web of Science, curated the list of jargon words and prepared the glossary. S.M. performed analyses and prepared figures. Both authors equally contributed to the writing.

**Competing interests.** We declare we have no competing interests.

**Funding.** S.M. acknowledges support from the European Commission through Horizon 2020 Marie Skłodowska-Curie Actions (MSCA) individual fellowship program (Grant no. 882221).

**Acknowledgements.** This paper started as a discussion with our friend Martina Pavlek in the Simphon Music Café in Pallanza. By reading



with hindsight some of our early manuscripts, we realized how rich in jargon—and, in parallel, poorly cited—they were: this further stimulated us in exploring this subject. We thank Irene Frigo for preparing the illustration. We are grateful to Ester M. Eckert, Ricardo Martínez-García and Ziga Fišer for their useful comments. We also

want to thank the associated editor Stephanie Meirmans, as well as E. J. Sayer and one anonymous referee, for several constructive suggestions that greatly improved the quality and clarity of this manuscript. Irene and Maria, thanks for helping us to avoid jargon in our everyday lives.

## References

- Wittgenstein L. 1933 *Tractatus logico-philosophicus*. London, UK: Routledge & Kegan Paul.
- Yager RE. 1983 The importance of terminology in teaching K-12 science. *J. Res. Sci. Teach.* **20**, 577–588. (doi:10.1002/tea.3660200610)
- Torres JM. 1995 The importance of accurate terminology in the field of human gene transfer. *Hum. Gene Ther.* **6**, 133–135. (doi:10.1089/hum.1995.6.2-133)
- Köhler C, Baum R, Grobe P, Meid S, Quast B, Vogt L. 2018 Using semantics for morphological descriptions in morph D base. *Biodiver. Inform. Sci. Stand.* **2**, e25535. (doi:10.3897/biss.2.25535)
- Keiter DA, Mayer JJ, Beasley JC. 2016 What is in a 'common' name? A call for consistent terminology for nonnative *Sus scrofa*. *Wildlife Soc. Bull.* **40**, 384–387. (doi:10.1002/wsb.649)
- Hoyningen-Huene P. 2013 *Systematicity: the nature of science*. New York, NY: Oxford University Press.
- Feynman R. 1969 What is science? *The Physics Teacher* **7**, 313–320.
- Wittgenstein L. 1953 *Philosophical Investigations*. Oxford, UK: Blackwell.
- Montgomery SL. 1989 The cult of jargon: reflections on language in science. *Sci. Cult.* **1**, 42–77. (doi:10.1080/09505438909526248)
- Adams DC, Bitetti MSD, Janson CH, Slobodkin LB, Valenzuela N. 1997 An 'Audience effect' for ecological terminology: use and misuse of Jargon. *Oikos* **80**, 632–636. (doi:10.2307/3546641)
- Hirst R. 2003 Scientific jargon, good and bad. *J. Tech. Wri. Commun.* **33**, 201–229. (doi:10.2190/J8JJ-4YD0-4R00-G5N0)
- Rakedzon T, Segev E, Chapnik N, Yosef R, Baram-Tsabari A. 2017 Automatic jargon identifier for scientists engaging with the public and science communication educators. *PLoS ONE* **12**, e0181742. (doi:10.1371/journal.pone.0181742)
- Barnett A, Doubleday Z. 2020 The growth of acronyms in the scientific literature. *Elife* **9**, e60080.
- Tan H-T, Wang EY, Yoo G-S. 2019 Who likes jargon? The joint effect of jargon type and industry knowledge on investors' judgments. *J. Account. Econ.* **67**, 416–437. (doi:10.1016/j.jacceco.2019.03.001)
- Orwell G. 1968 Politics and English Language. In *The collected essays, journalism and letters of George Orwell* (eds S Orwell, I Angos), pp. 127–140. New York, NY.
- Plavén-Sigra P, Matheson GJ, Schiffler BC, Thompson WH. 2017 The readability of scientific texts is decreasing over time. *Elife* **6**, e27725. (doi:10.7554/eLife.27725)
- Racovitza ÉG. 1907 Biospéologia: Essai sur les problèmes biospéologiques. *Archs Zool. Exp. Gen.* **4**, 371–488.
- Mammola S, Martínez A. 2020 Let research on subterranean habitats resonate! *Subterranean Biol.* **36**, 63–71. (doi:10.3897/subtblol.36.59960)
- Sket B. 2008 Can we agree on an ecological classification of subterranean animals? *J. Nat. Hist.* **42**, 1549–1563. (doi:10.1080/00222930801995762)
- Giachino PM, Vailati D. 2017 Considerations on biological and terminological aspects of the subterranean and endogean environments. Diversity, correlations and faunistic interchange. *Atti Accad. Naz. Ital. di Entomol.* **65**, 157–166.
- Trajano E, de Carvalho MR. 2017 Towards a biologically meaningful classification of subterranean organisms: a critical analysis of the Schiner-Racovitza system from a historical perspective, difficulties of its application and implications for conservation. *Subterranean Biol.* **22**, 1–26. (doi:10.3897/subtblol.22.9759)
- Mammola S. 2019 Finding answers in the dark: caves as models in ecology fifty years after Poulson and White. *Ecography* **42**, 1331–1351. (doi:10.1111/ecog.03905)
- Poulson TL, White WB. 1969 The cave environment. *Science* **165**, 971–981. (doi:10.1126/science.165.3897.971)
- Mammola S *et al.* 2020 Fundamental research questions in subterranean biology. *Biol. Rev.* **95**, 1855–1872. (doi:10.1111/brv.12642)
- Culver DC, Pipan T. 2019 Ecological and evolutionary classifications of subterranean organisms. In *Encyclopedia of caves* (ed. W White, D Culver, T Pipan), pp. 376–379. New York, NY: Academic Press.
- Mammola S, Fontaneto D, Martínez A, Chichorro F. 2020 Impact of the reference list features on the number of citations. *Scientometrics* **126**, 785–799. (doi:10.1007/s11192-020-03759-0)
- Mabe MA, Amin M. 2002 Dr Jekyll and Dr Hyde: author-reader asymmetries in scholarly publishing. *Aslib Proceedings* **54**, 149–157. (doi:10.1108/00012530210441692)
- Bornmann L, Mutz R. 2015 Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references. *J. Assoc. Inform. Sci. Technol.* **66**, 2215–2222. (doi:10.1002/asi.23329)
- Johnson R, Watkinson A, Mabe M. 2018 *The STM report: an overview of scientific and scholarly publishing*. The Hague, The Netherlands: International Association of Scientific, Technical and Medical Publishers.
- Pinker S. 2014 Why academics stink at writing. *The Chronicle of Higher Education*, 1 October.
- Ioannidis JP, Klavans R, Boyack KW. 2018 Thousands of scientist publish a paper every five days. *Nature* **561**, 167–169. (doi:10.1038/d41586-018-06185-8)
- França TFA, Montserrat JM. 2019 Writing papers to be memorable, even when they are not really read. *Bioessays* **41**, 1900035. (doi:10.1002/bies.201900035)
- Bowman D, Kinnan S. 2018 Creating effective titles for your scientific publications. *VideoGIE* **3**, 260. (doi:10.1016/j.vgie.2018.07.009)
- Freeling B, Doubleday ZA, Connell SD. 2019 Opinion: how can we boost the impact of publications? Try better writing. *Proc. Natl Acad. Sci. USA* **116**, 341–343. (doi:10.1073/pnas.1819937116)
- Evans JA. 2008 Electronic publication and the narrowing of science and scholarship. *Science* **321**, 395–399. (doi:10.1126/science.1150473)
- Gingras Y, Larivière V, Archambault É. 2009 Literature citations in the internet era. *Science* **323**, 36. (doi:10.1126/science.323.5910.36a)
- Hortal J, de Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ. 2015 Seven shortfalls that beset large-scale knowledge of biodiversity. *Annu. Rev. Ecol. Syst.* **46**, 523–549. (doi:10.1146/annurev-ecolsys-112414-054400)
- Sand-Jensen K. 2007 How to write consistently boring scientific literature. *Oikos* **116**, 723–727. (doi:10.1111/j.0030-1299.2007.15674.x)
- Martínez A, Mammola S. 2021 Specialized terminology reduces the number of citations of scientific papers. Dryad Digital Repository. (doi:10.5061/dryad.9w0vt4bdp)