

ARTIST: ARNICA KHATON ¹

¹Bachelor of Integrated Science (Honours), Class of 2026, McMaster University



RESEARCH INSIGHT

**SHARING SCIENCE MADE SIMPLE:
EXPLORING THE QUALITY AND READABILITY
OF PUBLISHED LAY SUMMARIES**



ABSTRACT

Lay summaries exist to bridge the gap that separates the scientific community from the general public. To foster improved science communication, this study examined the overall quality and readability of published lay summaries across peer-reviewed journals. We obtained 200 lay summaries published in four science journals: eLife, PLOS Medicine, Proceedings of the National Academy of Science (PNAS), and the Journal of Hepatology. Over 900 students across three semesters participated as raters of each summary using a rubric developed to assess the overall quality, accuracy, and accessibility of lay summaries across these journals. The Flesch Reading Ease formula was used to determine the readability of the highest and lowest scoring summaries from each journal. eLife and the Journal of Hepatology had the highest and lowest mean scores for overall quality of 15.6 and 11.7 out of 20, respectively. There were statistically significant differences in accuracy and accessibility found across all journals ($p < 0.0001$). eLife had the highest scoring lay summary for readability. The differences in and lack of consistent scoring across journals with the rubric indicate that deficits exist in the overall quality and readability of published lay summaries. These findings may support the development of guidelines that incorporate elements of the rubric used to write effective lay summaries.

INTRODUCTION

Scientific journals are the predominant medium for communicating research findings. Scientists have the necessary training to read and contextualize scientific articles, however, the general public may lack these skills.¹ This is a significant issue as non-specialist audiences are key stakeholders in scientific research through their roles as patients, research participants, and members of society, and need to be informed of important advancements.^{2,3} Unlike technical abstracts, lay summaries are designed to communicate research to the general public in an accessible format.⁴ These texts are typically 250 to 300 words of relevant summary content that clarify complex ideas, and are free of jargon.^{5,6} Further, scientific journals use the passive voice while lay summaries use the active voice to foster greater reader engagement.⁷

The Importance of Lay Summaries

Lay summaries benefit both the general public and the scientific community. Greater accessibility facilitates higher visibility of research, correlating with increases in the number of citations and future scientific collaborations.⁸ Additionally, lay summaries extend the reach of research articles on social media and news platforms, removing the barrier for general audiences to understand and engage with relevant science.⁹ This translation of knowledge can increase public interest in science, leading to greater public participation in research.¹⁰ Lay summaries can help to combat misinformation, particularly in the health sciences.¹⁰ Accuracy and accessibility are essential to a powerful lay summary. Accuracy refers to making information easier to understand without compromising the clear communication of facts. Accessibility refers to the removal of barriers to public

understanding, such as jargon. Many researchers find it challenging to write lay summaries that accurately and accessibly communicate findings to the public.³ This is troubling as lay summaries are the strongest defense against a reporting practice known as spin –the biased and inaccurate translation of scientific information, leading to misinterpretations.¹¹ Disparities in these aspects may be due to a lack of consensus and training on instructions to write lay summaries across peer-reviewed journals.¹² For example, PLOS Medicine requires authors to include a bullet point-structured summary of their research that is accurate and accessible to non-scientists.¹³ However, apart from omitting technical terminology and acronyms, the terms accurate and accessible are not further defined.¹³ According to a sentiment analysis published in the Cambridge University Press, the readability of scientific writing has declined, while the use of scientific jargon has increased over the past three decades.¹⁴ The Flesch Reading Ease (FRE) score is a readability measure that assesses how easily a piece of text can be understood. Applying the FRE to scientific abstracts has revealed that the average number of syllables per word and the average sentence length in scientific abstracts have steadily increased since 1960.¹⁴ More than 20% of abstracts currently have an FRE score at the post-secondary level, well-above the target of 8th and 9th grade students.¹⁵ These findings indicate that current science communication efforts may not effectively convey research findings to general audiences.

Knowledge Gap and Research Question

This study builds on data collected from three semesters of undergraduate students in the LIFESCI 2AA3 course at McMaster University. The aim of this ongoing research is to help establish a gold standard for the quality of lay summaries across peer-reviewed journals. This study compares the quality of selected lay summaries across journals using a rubric to assess overall quality, including accuracy and accessibility, and the FRE score to assess readability. This investigation also explores the potential of this rubric in helping to establish guidelines for writing lay summaries.

RESEARCH DESIGN

This study analyzed 200 lay summaries from four reputable (impact factor of 3 or higher) scientific journals: eLife, PLOS Medicine, Proceedings of the National Academy of Science (PNAS), and Journal of Hepatology. These journals were selected as they consistently publish lay summaries alongside articles. Fifty lay summaries were assessed from each journal using two methods: a subjective, specially developed rubric to determine accessibility and accuracy, and an objective FRE formula score to determine readability. Over 900 undergraduate students in the LIFESCI 2AA3 participated as raters. Each lay summary was graded by six to nine students using a rubric (Appendix Figure A) created by Dr. Katie Moisse, Assistant Professor and Associate Director of the School of Interdisciplinary Science at McMaster University. The rubric had four categories, each worth five points for a total score for overall quality out of 20. Line 1 asked whether the study's methods, results, and conclusions were accurately summarized, while Line 2 asked the same of the rationale, implications, and limitations. Together, these lines evaluated accuracy. Line 3 assessed the clarity and organization of the writing, and Line 4

assessed whether it was tailored towards non-expert audiences, together evaluating accessibility. The statistical significance of the data was determined using a one-way ANOVA and Tukey Test. The FRE formula was used to calculate readability of the highest and lowest scoring lay summaries on the rubric from each journal, using the total words per sentence and total syllables per word.¹⁶ The numerical scores and the corresponding United States school grade reading level necessary to understand the article were determined using the FRE table (Appendix Table A).¹⁶

RESULTS

eLife, PLOS, PNAS, and Hepatology had mean total rubric scores of 15.6, 14.7, 13.2, and 11.7, respectively. Total scores were significantly different across the four journals ($p < 0.0001$), and between each journal ($p < 0.0001$) (Figure 1).

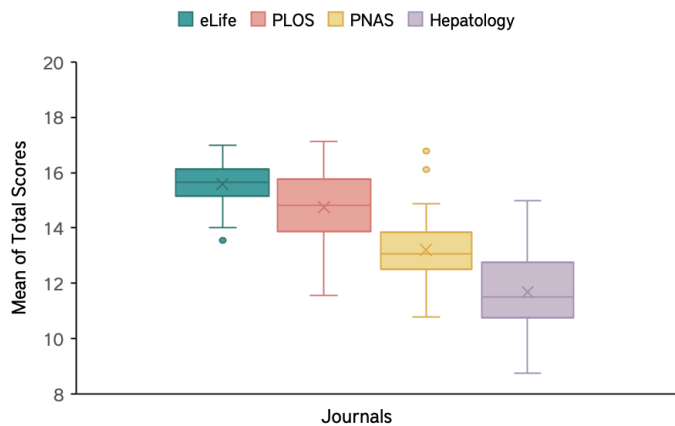


Figure 1. Mean of total scores for each journal using the LIFESCI 2AA3 rubric. Each of the four categories were scored out of five, 1 being strongly disagree and 5 being strongly agree. Each journal received a total score out of 20, which was used to assess overall quality (accuracy and accessibility). Statistically significant differences in quality were found across all journals ($p < 0.0001$).

Figure 2A highlights the outlier in the accuracy trend mentioned above; eLife vs. PLOS show no significant differences in accuracy ($p = 0.6134$). Of the two, PLOS scored the highest in accuracy with mean scores of 3.9 and 3.6 out of 5 for Lines 1 and 2, respectively. Additionally, Figure 2B highlights the outlier in the accessibility trend mentioned above; PNAS vs. Hepatology show no significant differences in accessibility ($p = 0.8752$). Of the two, eLife had the highest accessibility scores of 4.3 and 4.0 out of 5 for Lines 3 and 4, respectively.

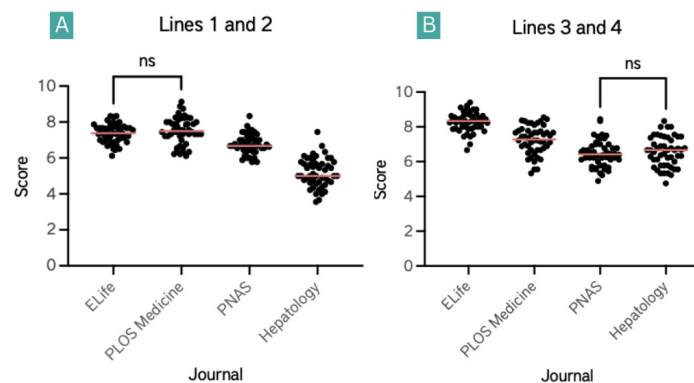


Figure 2. Scatterplot of combined total scores. (A) Lines 1-2 represent accuracy. (B) Lines 3-4 represent accessibility. Each dot represents an individual lay summary. The median is represented by the red lines. No statistical significance is indicated by “ns”.

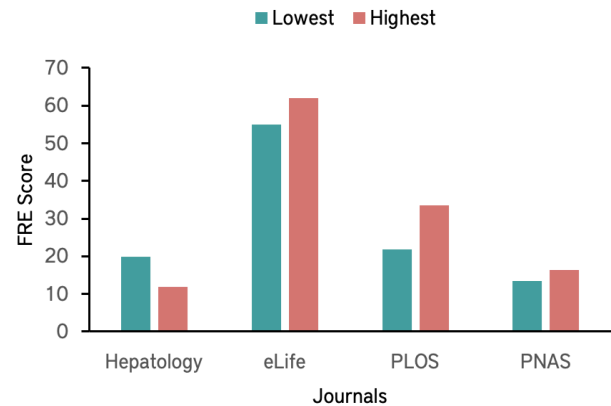


Figure 3. Bar graph with FRE scores for the highest and lowest scoring lay summaries from each journal. The highest scoring lay summary was from eLife, which also had the highest FRE score of 62.0, corresponding to the reading level of an 8th to 9th grade student (Figure 3). Lay summaries from PLOS, PNAS, and Hepatology scored between 10.0 and 30.0 on average, corresponding to university level or higher (Appendix Table 1).

DISCUSSION

Overall Quality

In order to promote effective communication of scientific research, the quality of lay summaries must be improved, as demonstrated by the results of this study. eLife had the highest overall quality, with a mean rubric score that is 1.4 times that of Hepatology, the lowest scoring journal (Figure 1). As seen in Figure 1, eLife also had the smallest range of 3.0 for overall quality, suggesting a greater consistency in scoring across raters of the three semesters for the journal and its published lay summaries. Hepatology scored the lowest for accuracy and overall quality, indicating greater variability in grading. The differences discovered between journals and eLife’s high rubric score may be a reflection of the guidelines used to publish their lay summaries, called *eLife digests*. eLife collaborates with freelance writers and science communicators to provide clear and engaging insights into research.¹⁷ Unlike the other journals examined, eLife has invested resources into researching different approaches to translate science for general audiences.¹⁷ *eLife digests* are typically 200-400 words, and are considered to have an optimal length by 90% of non-scientist readers, as found in a 2016 *eLife digest* reader survey.¹⁸ *eLife digests*’ high quality scores on the rubric demonstrate the value of a lay summary writing process with clear guidelines to produce accurate and accessible lay summaries.

Readability

As seen in Figure 3, the lowest and highest quality summaries from eLife had the highest FRE score of 52.0 and 62.0, respectively. This, along with the fact that eLife had the highest score for accessibility and overall quality in this study, suggest that improved readability may contribute to a higher score on the rubric. The average assessed lay summary scored at the undergraduate and graduate level of readability. Thus, many lay summaries may be difficult to understand for individuals without a higher level of education. Studies of lay summaries from a similar level of journals

(impact factor higher than 3) have also demonstrated lower FRE scores. Wen et al. found that while lay summaries published in Autism Research were significantly more readable than their corresponding abstracts, the mean FRE score for lay summaries was 34.39.¹⁹ Shiely and Daly assessed a random sample of lay summaries from randomised controlled trials from the National Institute for Health and Care Research and found a mean FRE score of 42.77.²⁰ These findings align with the undergraduate literacy level requirement found in this investigation. The restrictive readability illustrates the need for researchers and editors to adopt more accessible writing techniques. As shown in Figure 3, the highest scoring summary from Hepatology on the rubric had a lower FRE score than the lowest scoring summary. This finding suggests that readability scores may not always function as an accurate measure of an individual's comprehension. Studies have found that some general audiences prefer plain-language summaries of medium complexity, written for a reading level between 14-17 years.²¹ Hence, a higher FRE score may not necessarily correspond to a higher quality lay summary on the rubric. Additionally, oversimplified lay summaries may not only miscommunicate findings but may also lead to negative public perception.²¹

Implications, Limitations and Next Steps

The findings of this study should be used to inform guidelines for the communication of scientific findings to the public. The scores obtained from the combined dataset of the three semesters of students did not differ significantly from that of each semester alone. Thus, the consistency in gradings for each journal may support the use of the specially developed rubric as a standardized guideline for writing lay summaries. The depth of raters' evaluations in the study was limited by the rubric's design, which featured only four sections to assess accuracy and accessibility. Free-text feedback and thematic coding could be used with the rubric to identify nuanced factors and perceptual differences that may be missed using the current rubric. Readability measures such as FRE are also limited in the scope of their evaluation, as the formulas cannot assess other features that affect a reader's understanding such as headers, visuals, and line spacing. Future research should assess how a lay summary that is written in accordance with the rubric in this study alters the overall comprehension of the text by non-technical audiences.

CONCLUSION

This study explored the overall quality of published lay summaries across peer-reviewed journals. We found significant differences in the overall quality between journals. Additionally, we found differences in readability scores that support the need to use more inclusive writing techniques in science. This research can help improve science communication by supporting guidelines for writing lay summaries across journals.

ACKNOWLEDGEMENTS

We would like to extend a huge thank you to Dr. Katie Moisse for her guidance and support throughout this project and creation of the rubric. Thank you to each student in the LIFESCI 2AA3 course who participated in the study. This research was conducted as part of a summer research study opportunity which entailed paid research under the Office of Undergraduate Research. There are no conflicts of interest to declare.

APPENDIX

Figure A. LIFESCI 2AA3 lay summary grading rubric.

Does the lay summary accurately summarize the study methods, results and conclusions?	Does the lay summary accurately summarize the study rationale, implications and limitations?	Is the lay summary's writing clean, clear and logically organized?	Is the lay summary's writing tailored to its audience and purpose?	Total

Table 1. Flesch Reading Ease (FRE) Scores.¹⁴

Reading Ease Score	Reading Level/Grade	Description
90 – 100	5 th Grade	Very Easy
80 – 90	6 th Grade	Easy
70 – 80	7 th Grade	Fairly Easy
60 – 70	8 th and 9 th Grade	Standard
50 – 60	10 th to 12 th Grade	Fairly Difficult
30 – 50	College/University	Difficult
0 – 30	Graduate	Very Difficult

REVIEWED BY: DR. MAYA NISHIMURA (PHD)

Dr. Mayu Nishimura is an Assistant Professor of Psychology, Neuroscience & Behaviour at McMaster University and a research associate at SickKids Hospital in Toronto. She conducted her PhD at McMaster University. She is passionate about the presentation of scientific information in a way that is interesting and understandable to the general public.

EDITED BY: SEPEHR BAHARESTAN & ALINA PACE

1. Tancock C. In a nutshell: How to write a lay summary [Internet]. Elsevier Connect; 2018. Available from: <https://www.elsevier.com/connect/authors-update/in-a-nutshell-how-to-write-a-lay-summary> [cited 2023 Aug 2].
2. Goldstein CM, Krukowski RA. The Importance of Lay Summaries for Improving Science Communication. *Ann Behav Med*. 2023;57(7):509–10. Available from: doi:10.1093/abm/kaad027.
3. Wada, M., Sixsmith, J., Harwood, G. et al. A protocol for co-creating research project lay summaries with stakeholders: guideline development for Canada's AGE-WELL network. *Res Involv Engagem*. 2020;6(22). Available from: doi:10.1186/s40900-020-00197-3.
4. Rosenberg A, Walker J, Griffiths S, Jenkins R. Plain language summaries: Enabling increased diversity, equity, inclusion and accessibility in scholarly publishing. *Learn Publ*. 2023;36(1):109–18. Available from: doi:10.1002/leap.1524
5. How to Prepare a Lay Abstract [Internet]. CDF. Available from: <https://cdf.ca/en/how-to-prepare-a-lay-abstract/> [cited 2023 Aug 2].
6. Cramm H, Breimer J, Lee L, Burch J, Ashford V, Schaub M. Best practices for writing effective lay summaries. *JMVH*. 2017;3(1):7–20. Available from: doi:10.3138/jmvfh.3.1.004.
7. Chan EY, Maglio SJ. The Voice of Cognition: Active and Passive Voice Influence Distance and Construal. *Pers Soc Psychol Bull*. 2020;46(4):547–58. Available from: doi:10.1177/0146167219867784.
8. Liang X, Su LYF, Yeo SK, Scheufele DA, Brossard D, Xenos M, et al. Building Buzz: (Scientists) Communicating Science in New Media Environments. *JMCQ*. 2014;91(4):772–91. Available from: doi:10.1177/1077699014550092.
9. Thelwall M, Haustein S, Larivière V, Sugimoto CR. Do Altmetrics Work? Twitter and Ten Other Social Web Services. *Bornmann L, editor. PLoS ONE*. 2013;8(5). Available from: doi:10.1371/journal.pone.0064841.
10. Nunn E, Pinfield S. Lay summaries of open access journal articles: engaging with the general public on medical research. *Learn Pub*. 2014;27(3):173–84. Available from: doi:10.1087/20140303.
11. Boutron I, Haneef R, Yavchitz A, Baron G, Novack J, Oransky I, et al. Three randomized controlled trials evaluating the impact of "spin" in health news stories reporting studies of pharmacologic treatments on patients'/caregivers' interpretation of treatment benefit. *BMC Med*. 2019;17(1):105. Available from: doi:10.1186/s12916-019-1330-9.
12. Kuehne LM, Olden JD. Lay summaries needed to enhance science communication. *Proc Natl Acad Sci U S A*. 2015;112(12):3585–3586. Available from doi: 10.1073/pnas.1500882112.
13. Revising Your Manuscript [Internet]. PLOS Medicine. Available from: <https://journals.plos.org/plosmedicine/s/revising-your-manuscript#:~:text=Author%20Summary,-We%20ask%20that&text=Authors%20should%20aim%20to%20highlight,the%20study%20simply%20and%20objectively> [cited 2023 Oct 31].
14. Wen J, Lei L. Adjectives and adverbs in life sciences across 50 years: implications for emotions and readability in academic texts. *Scientometrics*. 2022;127(8):4731–49. Available from: doi:10.1007/s1192-022-04453-z.
15. Plavén-Sigra P, Matheson GJ, Schiffer BC, Thompson WH. The readability of scientific texts is decreasing over time. *eLife*. 2017;6:e27725. Available from: doi:10.7554/eLife.27725.
16. Flesch R. How to write plain English: a book for lawyers and consumers. New York: Barnes & Noble; 1981.
17. King SR, Pewsey E, Shailes S. An inside guide to eLife digests. *eLife*. 2017;6:e25410. Available from: doi: 10.7554/eLife.25410.
18. Plain-language Summaries: Results of the 2016 Elife digest reader survey. Elife Sciences Publications Limited [Internet]; 2017. Available from: <https://elifesciences.org/inside-Elife/19c97b89/plain-language-summaries-results-of-the-2016-Elife-digest-reader-survey> [cited 2023 Sept 10].
19. Wen J, He S, Yi L. Easily readable? Examining the readability of lay summaries published in Autism Research. *Autism Res*. 2023;16(5):935–40. Available from: doi:10.1002/aur.2917.
20. Shiely F, Daly A. Trial lay summaries were not fit for purpose. *J Clin Epidemiol*. 2023;156:105–12. Available from: doi:10.1016/j.jclinepi.2023.02.023.
21. Martínez Silvagnoli L, Shepherd C, Pritchett J, Gardner J. Optimizing Readability and Format of Plain Language Summaries for Medical Research Articles: Cross-sectional Survey Study. *J Med Internet Res*. 2022;24(1):e22122. Available from: doi:10.2196/22122.