

# Bridging the Gap: A Study on Enhancing Accessibility and Accuracy in Scientific Lay Summaries

doi: 10.15173/sw.v1i3.3820

Authors: Tehani Blais<sup>1</sup>, Monalisa Bonilla<sup>2</sup>, Maya Elfizy<sup>3</sup>, Leen Ramahi<sup>4</sup>

<sup>1</sup>McMaster University, Faculty of Science, Bachelor of Science 2026

<sup>2</sup>McMaster University, Faculty of Science, Bachelor of Science (Honours), Origins of Disease, 2026

<sup>3,4</sup>McMaster University, Faculty of Science, Bachelor of Science (Honours) 2027

Illustrator: Isra Chowdhury<sup>3</sup>

<sup>3</sup>McMaster University, Faculty of Science, Bachelor of Science Co-op (Honours) 2026



## INTRODUCTION



The communication of scientific knowledge is essential to the development of future treatments and discoveries, consisting of conferences, symposiums, and research articles. While these methods are accessible to experts in the field, members of the public may find this beyond their comprehension. More than 50% of the public believe that scientists are elitists, and 88% feel science is too complicated<sup>1</sup>. For this reason, lay summaries are essential in promoting public interest and engagement in science. These short abstracts are typically 250 words, concisely summarizing each section of a research paper for a wide range of readers<sup>2</sup>. This broader audience increases research visibility, further aiding in reducing the accidental and intentional spread of false information often produced by the media<sup>3</sup>.

Scientific journals use lay summaries in an effort to engage more readers by including short and clear analogies written in positive language<sup>4</sup>. However, the production of inaccessible and confusing lay summaries is quite common. Many factors contribute to poorly written lay summaries, including the use of jargon - terms scientists use to communicate with experts. Another barrier of lay summaries is the use of passive voice, as many studies have found that readers interpret passive voice passages as more abstract<sup>5</sup>.

Given these concerns, it is worth considering: What is the quality of published lay summaries? Do different journals have different track records for accuracy and accessibility of lay summaries? How can we find a balance within a lay summary that is concise but still detailed and reliable? These questions have been asked before, and this study has been performed various times. However, it was conducted again to determine if there has been any improvement in the quality of lay summaries from recently published papers.

## METHODS

In this study, 250 students in SCICOMM 2A03 analyzed 200 lay summaries across four journals - The Proceedings of the National Academy of Sciences, PLOS Medicine, eLife, and the Journal of Hepatology- using Dr. Katie Moisse's inclusive summary rubric. Dr. Moisse is an Assistant Professor in the Faculty of Science at McMaster University and the primary instructor of SCICOMM 2A03. The rubric included four sections worth 5 points each, with 1 point representing the lowest and 5 points representing the highest score awarded. This makes the total 20 points for each lay summary. The following are sections used in our rubric: accuracy of the study's methods, results, and conclusions summary, accuracy of the study's rationale, implications, and limitations summary, clarity and organization of the lay summary, and the writer's audience and purpose. Each student in our SCICOMM 2A03 class assessed two lay summaries. To ensure consistent results, every lay summary was reviewed by three different students, and the average of all three scores was taken to minimize reviewer bias. Dr. Moisse analyzed this data using Analysis of Variance- Anova, a statistical test useful in assessing the difference between the averages of multiple groups. The resulting data was used in our experiment.

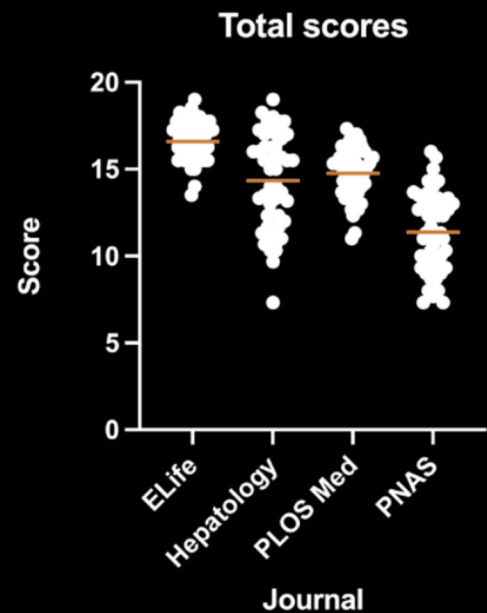
## RESULTS

The total mean rating for the scientific journals eLife, Hepatology, PLOS Med, and PNAS are 16.593, 14.335, 14.757, and 11.373, respectively. Statistical analysis showed significant results between the journals ( $p < 0.001$ ).





This indicates that there is a less than 0.1% chance that our data is due to chance. The only exception is Hepatology versus PLOS Med, showing a p-value of  $p < 0.7171$ , meaning their difference is not statistically significant.



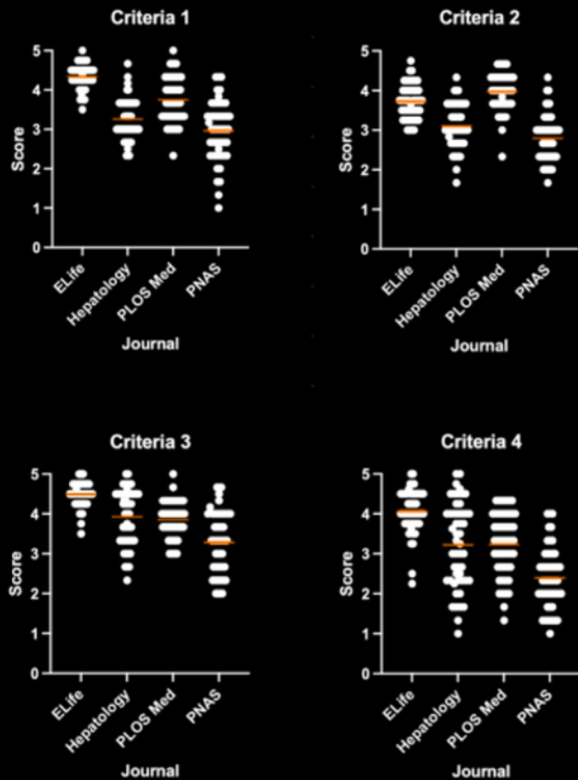
**Figure 1: Mean of Total Scores of Journal Ratings based on SCICOMM 2A03 rubric** 50 peer-reviewed article lay summaries from each journal (n = 4) were scored out of 5 based on four criteria, for a total of 20 points. The highest scoring journal with a mean rating of 16.593 is ELife, then PLOS Med with 14.757, Hepatology with 14.335, and lastly an 11.373 average for PNAS. This data was calculated using an ANOVA.

Criteria 1 was “Does the lay summary summarize the study methods, results, and conclusions accurately?” The rating for eLife is clustered between 4-5. Hepatology and PLOS Med have larger distributions between 2-4. PNAS’s score has majority residing at 3.

Criteria 2 was “Did the lay summary accurately summarize the study rationale, implications, and limitations?” eLife and PLOS both score between 3 and 5 with averages around 4. Hepatology and PNAS show great variability ranging from 1 to just above 4, both averaging around 3. Criteria 3 was “Is the lay summaries’

writing clean, clear, and logically organized?” eLife has the highest average with nearly all articles scoring between 4-5. The 3 other journals have majority of the averages almost reaching 4. The overall score and average of PNAS is slightly lower, at 3.2.

Criteria 4 is “Is the lay summary's writing tailored to its audience and purpose?” eLife has a smaller range and an average of 4. Hepatology and PLOS Med both have very large ranges spread across the graph, with an average around 3. PNAS has a large distribution and has a significantly lower average between 2-3.



**Figure 2: Violin graph of scores for each criterion from every journal and article evaluated** Each journal article (n = 200) is evaluated using four different criteria, focusing on accuracy and accessibility. The four criteria are evaluated on a scale from zero to five, with zero failing to satisfy the criteria and 5 fully fulfilled. This figure demonstrates the distribution of scores evaluated from the four journals.

DISCUSSION

As illustrated in Figure 1, a precise guideline on lay summary formulation leads to higher comprehension of article content. ELife, with the highest mean rating of 16.593, was 1.5 times higher than PNAS- the lowest rated journal. The score distribution is also significant; the variation between eLife’s best and worst article is 5.5, while PNAS has a variation of 8.67. This indicates that eLife’s digests consistently score higher than other journals. The reason for this contrast may be attributed to the lay summary writing process. eLife’s abstracts are written by editors and writers, alongside authors, to accurately and properly translate research findings<sup>6</sup>. On the other hand, PNAS simply requests authors to compose their own summary, offering third party resources as guidelines<sup>7</sup>. The importance of concrete instructions is further emphasized when considering previous research on this topic. The results of this previous, yet similar, study ranked the Journal of Hepatology as last, with a score of 7.6 out of 20<sup>8</sup>. This is a stark difference from its current mean rating of 14.757 out of 20, which can be attributed to its change in abstract guidelines.

CONCLUSION

Our study assesses the quality of lay summaries across various scientific journals, examining their accuracy and accessibility. We hypothesized that journals with clear lay summary guidelines contain higher quality, comprehensive lay summaries. We also hypothesized that differences in the quality, accuracy, and accessibility would exist between the four journals. Our findings support these hypotheses, bringing us to the conclusion that eLife contains the most comprehensive, and therefore highest quality, lay summaries. Several limitations were revealed throughout our study, including

reviewer can be orming data bias, avoided by performing data collection in a blind fashion.

Our findings could potentially influence scientific communication in a positive manner, allowing for greater public scientific exposure. However, several directions remain unstudied. Future studies could explore how effective lay summaries influence a journal's impact factor, as it would reach a wider audience. Additionally, developing a universal guideline for lay summaries could help standardize summary quality across all scientific journals. Science communication is a world of its own, much like speaking a different language. What good can come out of research if no one is there to share it with?

Edited By: Kumkum Anugopal & Fawziah Zubaidah