



Heriot-Watt University
Research Gateway

Characteristics of Papers that Affect Citations in the Journal of Fish Biology

Citation for published version:

Fenton, MA, Fennell, H & Kaiser, MJ 2024, 'Characteristics of Papers that Affect Citations in the Journal of Fish Biology', *Journal of Fish Biology*, vol. 104, no. 1, pp. 310-314. <https://doi.org/10.1111/jfb.15565>

Digital Object Identifier (DOI):

[10.1111/jfb.15565](https://doi.org/10.1111/jfb.15565)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Journal of Fish Biology

Publisher Rights Statement:

© 2023 The Authors.

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Characteristics of papers that affect citations in the *Journal of Fish Biology*

M. A. Fenton | H. L. Fennell  | M. J. Kaiser 

The Lyell Centre, Heriot-Watt University,
Edinburgh, Scotland

Correspondence

M. J. Kaiser, The Lyell Centre, Heriot-Watt
University, Edinburgh, Scotland.

Email: m.kaiser@hw.ac.uk

Abstract

Identifying the factors that influence the citation of articles helps authors improve the impact and reach of their research. Analysis of publications in the *Journal of Fish Biology* between 2008 and 2021 revealed that variables such as the number of keywords, abstract length, number of authors, and page length were associated with higher impact papers. These trends applied to both review and regular papers. These findings suggest that papers that are more informative, have higher numbers of authors, and have more keywords are more likely to be cited. Adoption of some simple “best-practice” behaviors can improve the likelihood that a paper is cited.

KEYWORDS

abstract, authorship, citation, collaborating institutes, impact, keywords, page length, paper, review

An individual audit of a single journal provides a fine-grain analysis that can uncover trends within the field, thereby informing editors on how to improve the submission process and selection of papers to send out for review. Ultimately, improving selection processes and author performance will enhance the profile of the journal, while also indicating its value to potential authors. Previous studies on the impact (likelihood of citation) of articles have identified several factors that may influence the impact of articles: collaboration (individual, institutional, or international), number of references and keywords employed, length of the paper, abstract, title, structure of the title (i.e., the inclusion of nonalphanumeric characters), abstract readability, and author eminence (Didegah & Thelwall, 2013; Buter & van Raan, 2011). The variation in norms and conventions among different journals and research discipline areas may prevent the identification of universal variables associated with increased impact. What holds true for one discipline may be the opposite for another. For example, articles with long titles are more highly cited in medical journals, but are associated with fewer citations in biology and social science journals (Thelwall, 2016). Analysis of the variables that increase citations for articles in a specific journal can provide valuable insights that can be used by researchers within a field who seek to publish within a

specific journal to maximize the impact of their published research. Similarly, such an analysis enables editors to inform authors on how they might maximize the accessibility of their research to the intended audience.

To gain insight into the factors that influence the number of citations received for papers published in the *Journal of Fish Biology*, descriptive data were extracted from all publications within the journal between January 2008 and March 2021. A total of 3413 publications were included in the study, 3154 of which were articles (i.e., not obituaries, book reviews, errata, corrections, or letters). For each article, data on 27 variables were extracted (see Appendix A). Data were imported into R (R Core Team, 2022) and categorized as research articles, review articles, or short communications. The citation counts for each classification were then normalized to allow for comparison among years. Papers were grouped as “high impact,” “medium impact,” or “low impact,” where papers in the top or bottom quartile of normalized citation counts were classified as being “high impact” or “low impact,” respectively, to account for time since publication. All other papers were classified as “medium impact.”

Normalized citation values for each group were tested for normality using a Shapiro–Wilk test. The influence of extracted variables

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Fish Biology* published by John Wiley & Sons Ltd on behalf of Fisheries Society of the British Isles.

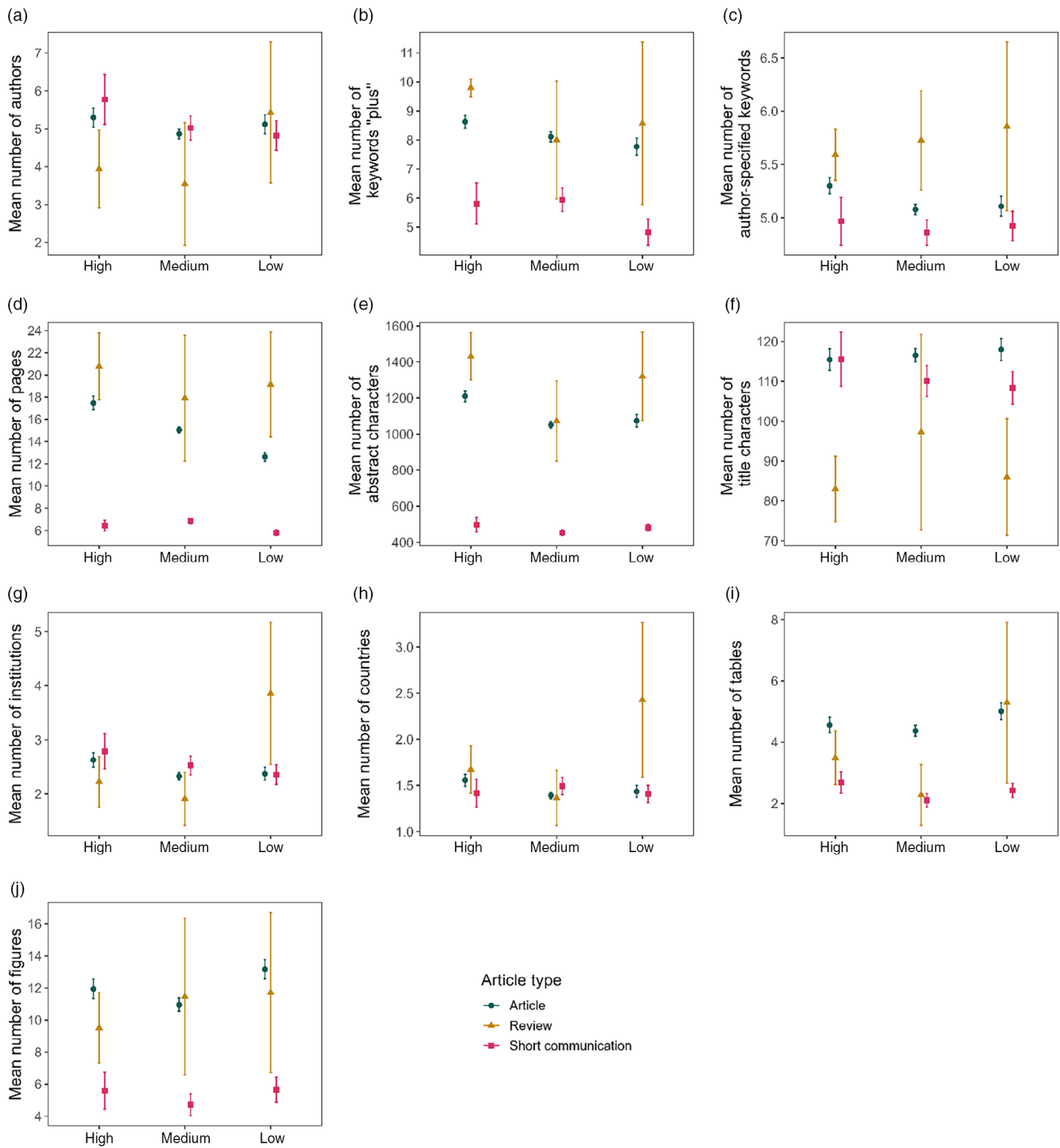


FIGURE 1 The relationship between 10 of the most influential variables extracted from papers published in the *Journal of Fish Biology* (between January 2010 and March 2021) and paper impact (high, medium, low) ($\pm 95\%$ confidence interval). Papers were categorized as high impact if they fell into the top quartile of normalized citation counts and low impact if they fell into the bottom quartile of normalized citation counts; all other papers were classified as medium impact.

on papers was explored using a Pearson correlation matrix to eliminate predictor variables that were auto-correlated or showed no pattern. This initial analysis left 11 variables that affected citation count: author count, country count, institute count, total page count, number of figures, number of tables, number of author-specified keywords, keywords "plus," lead author country, title character count, and

abstract character count (keywords "plus" refers to words or phrases that appear frequently in the titles of the papers referenced in the body of the article, but not in the title of the article itself).

Using the R package "AER" (Klieber & Zeileis, 2008) the data were found to be overdispersed and hence a negative binomial regression was adopted. The "MASS" package was used to estimate a

TABLE 1 The estimated parameters, standard error, Z and *p* values for the best fit negative binomial models describing the relationship between parameters and normalized citation count for each article

Type of article	Parameter	Estimate	Standard error	z value	<i>p</i>
Article	(Intercept)	-1.073	0.134	-8.030	<0.001
	Total pages	0.053	0.003	20.490	<0.001
	Number of author-specified keywords	0.078	0.020	3.892	<0.001
	Number of keywords "plus"	0.022	0.007	3.203	0.001
	Title character count	-0.002	0.001	-3.839	<0.001
	Number of tables	-0.016	0.006	-2.592	<0.001
Review article	(Intercept)	-0.795	0.962	-0.826	0.408
	Number of figures	0.006	0.015	0.435	0.664
	Number of keywords "plus"	0.170	0.089	1.892	0.059
	Title character count	-0.009	0.004	-2.470	0.014
	Abstract character count	0.001	0.000	2.903	0.004
Short communication	(Intercept)	-1.250	0.193	-6.463	<0.001
	Total pages	0.086	0.024	3.503	<0.001
	Number of authors	0.045	0.016	2.779	0.005

negative binomial regression, assessing the extent to which the different explanatory variables influenced citation counts. To determine which variables the model predicted had the most influence on citations, we sequentially removed combinations of variables to find the models with the best fit. We fitted separate models to ensure that the overdispersion parameter was not held constant. The best fit models were then ranked using Akaike's information criterion.

The suitability of the final models was assessed by plotting the model predictions against the observed data. The normality of residuals was tested using the Kolmogorov–Smirnov test and a Q–Q plot. Cook's distance plots were used to check for outliers. Heteroscedasticity was tested using the Levene's test and scatter plots of the standardized residuals while assessing the fitted values and all covariates (Figure 1).

Analysis showed a clear relationship between a number of variables and high impact, medium impact, and low impact (Table 1). For research articles, higher citations were most strongly associated with the number of author-specified key words that were positively correlated with the citation rate of the article ($F = 0.078$, $p \leq 0.001$). This trend could be attributed to how readers locate articles of interest by searching keywords. A small number of keywords lowered the likelihood of an article being found and read. Similarly, a greater number of pages for both research articles and short communications (measured by total number of pages) was associated with a higher citation rate ($F = 0.053$, 0.086 and $p \leq 0.001$, <0.001 , respectively), which indicates that more substantive paper content attracts more citations. Some correlation among variables was also evident, with articles with high author counts being more likely to have a higher number of associated institutes involved in the publication, as well as a greater number of countries.

These results broadly concur with the findings of previous studies that examined the factors influencing the impact of publications, where variables such as the number of keywords, the length of the

paper, the abstract, the title length, and collaboration (individual, institutional, or international) all influence citation rates (Didegah & Thelwall, 2013; Buter & van Raan, 2011).

Previous research across different journals has found that the structure of the title of an article may affect the recovery of an article through keyword searches and journal browsing, and may influence the decision to read an article (Thelwall, 2016). Authors tend to choose a title with the goal of maximizing the impact of a paper (Hudson, 2016) and may employ techniques and special characters to help generate interest. Previous studies found that different components of a title, including its content, structure, and length, may influence the number of citations attracted by a paper, although this is subject to the customs, norms, and preferences of different disciplines (Hudson, 2016; Thelwall, 2016). The presence of colons, hyphens, and other nonalphanumeric characters is common across disciplines and is widely associated with higher citation rates (Buter & van Raan, 2011). A notable exception to this rule, however, is titles containing questions, which are less frequently cited (Hudson, 2016). While analysis of the title structure was beyond the scope of this study, the results suggest that articles and review articles with longer title character counts are associated with a lower citation rate. A similar relationship was found between citation rate and abstract length (character count) for review articles, as well as the length of both research articles and short communications. This leads to the conclusion that overly specific titles are less likely to attract higher citations than titles that are more general and hence attract a broader readership.

In previous studies, multiple authorship was associated with higher citation counts. This is thought to be the result of "specialization gains" as authors with different areas of expertise are able to pool resources to collaborate and combine their skills and strengths (Haslam et al., 2008; Hudson, 2016). The results of this paper's analysis reveal that short communications with a higher number of authors have a higher citation rate than short communications with lower

numbers of authors, but no relationship between citation rate and number of authors was observed for articles and review articles (Table 1).

Within the field of biology, international collaboration is associated with an increased number of citations (Didegah & Thelwall, 2013). However, some studies have found that authors of particular nationalities are associated with higher-impact publications, while others suggest that there is no link between nationality and impact (Greenwald & Schuh, 1994; Haslam et al., 2008). This analysis did not find any correlation between citation count and number of institutions or countries for articles, review articles, or short communications. Other factors that may influence the citation rate of a paper (such as author gender, study region, study design, and study species) are not currently collected by Wiley and therefore these factors could not be included in the analysis. Future data collection on these variables should be considered to understand and address inequalities in gender and publication bias. Previous work by Burdett et al. (2021) has shown that between June 2018 and May 2021 only 35% of lead authors were women in articles published in the *Journal of Fish Biology*. Further research on author gender, study region, study design, and study species could provide further insight into citation trends.

AUTHOR CONTRIBUTIONS

M. A. Fenton and H. L. Fennell contributed equally to the production of the manuscript. M. A. Fenton and H. L. Fennell extracted the data, undertook the data analysis, and led the writing of the paper. M. J. Kaiser conceived the objective of the paper, and contributed to its writing and interpretation of the results.

FUNDING INFORMATION

The authors thank Wiley for funding this analysis.

ORCID

H. L. Fennell  <https://orcid.org/0000-0001-8671-225X>

M. J. Kaiser  <https://orcid.org/0000-0001-8782-3621>

REFERENCES

- Burdett, H., Kelling, I., & Carrigan, M. (2021). #TimesUp: Tackling inequalities in marine and fisheries science. *Journal of Fish Biology*, 100(1), 4–9.
- Buter, R. K., & van Raan, A. F. J. (2011). Non-alphanumeric characters in titles of scientific publications: An analysis of their occurrence and correlation with citation impact. *Journal of Informetrics*, 5(4), 608–617.
- Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, 7(4), 861–873.
- Greenwald, A. G., & Schuh, E. S. (1994). An ethnic bias in scientific citations. *European Journal of Social Psychology*, 24(6), 623–639.
- Haslam, N., Ban, L., Kaufmann, L., Loughnan, S., Peters, K., Whelan, J., & Wilson, S. (2008). What makes an article influential? Predicting impact in social and personality psychology. *Scientometrics*, 76(1), 169–185.
- Hudson, J. (2016). An analysis of the titles of papers submitted to the UK REF in 2014: Authors, disciplines, and stylistic details. *Scientometrics*, 109(2), 871–889.
- Kleiber, C., & Zeileis, A. (2008). *Applied econometrics with R*. Springer-Verlag. <https://CRAN.R-project.org/package=AER>
- R Core Team. (2022). R: A language and environment for statistical computing. R Computing. <https://CRAN.R-project.org/>
- Thelwall, M. (2016). Avoiding obscure topics and generalising findings produces higher impact research. *Scientometrics*, 110(1), 307–320.

How to cite this article: Fenton, M. A., Fennell, H. L., & Kaiser, M. J. (2023). Characteristics of papers that affect citations in the *Journal of Fish Biology*. *Journal of Fish Biology*, 1–5. <https://doi.org/10.1111/jfb.15565>

APPENDIX A

Extracted variables from papers in the *Journal of Fish Biology* broken down by component, variable, and type (number, factor etc.)

Component	Study variable	Type of variable
General	Number of citations	Number
	Year published	Factor
	Type of paper	Factor
Authors	Number of authors	Number
	Gender of lead author	Factor
	Multi-institutional	Yes/No
	Number of institutions	Number
	Country	Factor
	Number of countries	Number
Topic/keywords	Topic	Factor
	Number of keywords	Number
	Keyword specificity	Factor
Structure	Word count	Number
	Pages	Number
	References	Number
	Figures	Number
	Tables	Number
Abstract	Abstract character count	Number
	Abstract word count	Number
Title	Characters per title	Number
	Words per title	Number
	Punctuation in the title	Yes/No
	Punctuation in the title	Yes/No
	Punctuation in the title, other	Yes/No
	Region specified in title	Yes/No
	Study design stated in the title	Yes/No
	Study findings stated in the title	Yes/No
Study species listed in title	Yes/No	