Ecological Entomology

Ecological Entomology (2019), DOI: 10.1111/een.12740

OPINION

Does this title bug (Hemiptera) you? How to write a title that increases your citations

SHANNON M. MURPHY, MAYRA C. VIDAL,[†] CLAUDIA J. HALLAGAN, E. DALE BRODER,[‡] ELIZABETH E. BARNES,[¶] EVA S. HORNALOWELL and JACOB D. WILSON Department of Biological Sciences, University of Denver, Denver, Colorado, U.S.A.

Introduction

The number of times authors are cited can dramatically shape their career (e.g. the H index to measure an author's impact in the field; Hirsh, 2005). The number of scientific publications is increasing almost exponentially (Bornmann & Mutz, 2015), leading to competition among authors to publish papers that are easy to find and attractive to read. Scientists read an average of 1,142 titles each year and 204 abstracts, but they only go on to read *c*. 97 articles per year (Mabe & Amin, 2002), which indicates a severe filtering process based solely on title attractiveness. A title must convey the key topics of the paper while also quickly capturing a reader's attention. The pressure for scientists to 'sell' their work to other scientists is thought to be a driving force changing attributes of titles in scientific publications (e.g. increased use of question marks; Ball, 2009). Thus, crafting an appealing title is an important step in encouraging readership and future citation.

Recent work has explored correlations between title features and citation rates. Shorter titles tend to be cited more often than longer titles in psychology (Subotic & Mukherjee, 2014) and open access journals (Paiva *et al.,* 2012), but longer titles receive more citations in medical journals (Habibzadeh & Yadollahie, 2010; Jacques & Sebire, 2010). The use of special characters can also affect citation rate; colons, questions marks, and hyphens are associated with lower citation rates in open access journals (Paiva *et al.,* 2012) and in management science journals (Nair & Gibbert, 2016), but the research around colons is conflicting, as another review found that their use has no effect on citation rate across a variety of academic fields (Hartley, 2007). Previous research has also found differences in download rates versus citation rates based on how the title is written. For example, titles that include question marks (Jamali & Nikzad, 2011) or that are amusing (Subotic & Mukherjee, 2014) are downloaded more frequently but cited less often. The specificity of the title can also affect citation rate. For example, in ecological research, titles that include the name of the study organism are cited less frequently (Fox & Burns, 2015). Similarly, titles that include the geographic area where the study took place are also cited less frequently across a variety of disciplines (Jacques & Sebire, 2010; Paiva *et al.,* 2012).

How title qualities affect citation rates have been explored for some fields, such as psychology (Subotic & Mukherjee, 2014), management science (Nair & Gibbert, 2016), and ecology (Fox & Burns, 2015; Fox *et al.,* 2016), but not for entomology. We explored titles from 5 years of publication (2005–2009) from two entomology (*Ecological Entomology* and *Entomologia Experimentalis et Applicata*) and two ecology (*Oecologia* and *Oikos*) journals, documenting the presence of title features (Table 1). We tested whether any of these title features were correlated with citation rates.

We hypothesised that titles that address a general scientific phenomenon would have more citations than those that indicated the study of a specific species, as found in ecology journals by Fox and Burns (2015). Based on literature from other fields, we predicted that the use of study organism names (Latin or common) would decrease citation rates. Using Latin names in a title might cause readers who are not immediately familiar with that organism to skip the paper entirely when they are perusing a journal's table of contents, whereas titles that describe a general phenomenon, without indicating the study organism, might draw a wider audience and thus more citations. Moreover, we predicted that titles with common names would not be as heavily penalised as titles with Latin names, which are more unfamiliar to a wider audience. Similarly, we predicted that the use of functional groups or specific geographic locations in titles would also decrease citation rates. We predicted that the use of functional groups would decrease citations because, as with Latin and common names, readers who do not work on a specific functional group (e.g. herbivores, pollinators, predators) might skip papers that mention them. However, the use of functional groups to refer to a study species is broader than using a specific name, and so we also predicted that using functional groups in titles would not affect citation rate as negatively as using Latin or common names. Titles that focus on a geographic

Correspondence: Shannon M. Murphy, Department of Biological Sciences, University of Denver, 2050 E Iliff Ave, Boettcher West 302, Denver, CO 80210. E-mail: shannon.m.murphy@du.edu

[†]Current address: Department of Biology, Syracuse University, Syracuse, NY, U.S.A.

[‡]Current address: Department of Biology, St Ambrose University, Davenport, IA, U.S.A.

[¶]Current address: Department of Entomology, Purdue University, West Lafayette, IN, U.S.A.

Table 1. Features of titles that we recorded for each title. For binomial data, we recorded whether or not the title included the feature (yes/no). If the title included the feature (yes), we separately recorded specific information about that feature (e.g. if the Latin name *Drosophila* was in the title, we recorded 'yes' and then separately recorded *Drosophila*).

Title feature	Data type	If yes, information recorded
Species/genus Latin name	Binomial (yes/no)	Genus and/or species Latin name
Species/genus common name	Binomial (yes/no)	Genus and/or species common name
Family or higher Latin name	Binomial (yes/no)	Family or higher taxonomic classification Latin name
Family or higher common name	Binomial (yes/no)	Family or higher taxonomic classification common name
Functional group	Binomial (yes/no)	Name of functional group
Geographic location	Binomial (yes/no)	Country, state/region, and/or biome
Title attempts humour	Binomial (yes/no)	
Title includes '?'	Binomial (yes/no)	
Two-part title	Binomial (yes/no)	
Number of words	Continuous	Number of words in title

location or a biome may lose the interest of readers who do not work in those specific areas, reducing citation rate (e.g. Paiva *et al.,* 2012), and thus we predicted lower citation rates for titles that included specific geographic information. Given differing results in previous literature about how punctuation, length, and humour affect citation rates, we did not form specific hypotheses for how these traits would affect citation rate in entomological journals. Our findings will hopefully help authors to craft better titles and therefore improve readership and citation rate (see Box 1 for guidance).

BOX 1: Guidance for writing a title.

- We advise that authors avoid including the Latin names of their study system in the title. Species-specific titles may undersell their work and suggest a narrow research focus. Readers who are unfamiliar with the name may skip the paper.
- If desired, Latin names could be included in the key words or abstract so that they will be indexed by databases.
- However, overly broad titles are probably equally unhelpful and may mislead readers about the content of the paper.
- We recommend that authors aim for a 'Goldilocks zone'; the title should describe the interesting research question without driving away potential readers.

Materials and methods

Data collection

We collected titles from four journals [two entomology (*Ecological Entomology* and *Entomologia Experimentalis et Applicata*) and two ecology (*Oecologia* and *Oikos*)] over 5 years of publication (2005–2009). We selected these ecology journals because their impact factors are higher than those of the entomology journals, but within the same realm; all four of these journals are well respected in their fields. We chose this time frame so as to include recently published papers, but also to allow time for other authors to cite the articles. We recorded the titles from each issue of each journal during the 5-year span but excluded any papers that were reviews, responses, letters to the editor, corrigenda, notes, erratum, short communications, special issues, meta-analyses, or other non-empirical research papers. Non-empirical papers such as reviews and meta-analyses may be heavily cited regardless of the title features we studied, and thus we focused on empirical research papers. However, given the size of our dataset, we may have failed to exclude some non-empirical papers; to ensure that any highly cited reviews did not bias our results, we inspected all of the papers that had a mean citation rate of 15 per year or higher and removed any that were not empirical research. Thus, we analysed a total of 3,562 papers: 429 titles from *Ecological Entomology*, 619 from *Entomologia Experimentalis et Applicata*, 1,450 from *Oecologia*, and 1,064 from *Oikos*. For each title, we recorded whether or not the title included a variety of features (Table 1).

During the coding process, we established standardised methods of categorisation. We made sure not to code title features twice when a common name included other information that could conceivably be part of another category. For example, 'Alaskan swallowtail butterfly' was a common name that appeared in two of our titles (Murphy, 2007a,b) that included a geographic location (Alaska) and a family or taxonomically higher common name (butterfly). However, we only categorised 'Alaskan swallowtail butterfly' as a species common name; 'Alaska' was not separately included as a geographic location and 'butterfly' was not separately included as a family or taxonomically higher common name. Similarly, 'coffee leaf miner' was only categorised as a species common name of the leaf miner (coffee leaf miner) and 'coffee' was not separately coded as another common name. Some common names are based on Latin names, but are understood by the general public as common names (e.g. animal, amphibian, reptile, mammal, rodent, aphid, plant, thrips), and we coded these as common names. By contrast, some Latin names are used as common names by scientists, but are not understood by the general public, so we coded them as Latin names (e.g. rotifer, arthropod). Some words could be either a species or an ecosystem (e.g. mangroves) and we categorised them into only one category, based on the title's intended meaning. For example, we treated 'coral' as a common name if it referred to an organism and

treated it as a biome if used to refer to a coral reef (see Appendix S1 for more information about coding). When searching for two-part titles, we looked for any punctuation that divided a sentence, including dashes, colons, question marks, periods, and semicolons.

To accurately count the number of citations for each paper, we downloaded the number of citations per paper on 4 December 2017 using Web of Science and recorded the total number of citations for each title. We then calculated the mean number of citations per year. However, because papers in our dataset differ in age by 5 years, comparing citation rates at the same time point is disadvantageous to the more recent papers because they have had less time to accumulate citations (e.g. if we compare mean citations for all papers in 2012, this would be 7 years after publication for some papers but only 3 years after publication for others). Thus, we also calculated the mean number of citations per year for each paper 6 years after publication (e.g. 2009 papers were analysed from 2009 to 2014). This 6-year window allowed a fair comparison of papers published in different years but over the same relative time frame post-publication. Lastly, we found the impact factor for each journal at the time that the paper was published.

Statistical analysis

To test which title features influence number of citations, we used model selection of generalised linear mixed models. The mean number of citations in 6 years (following publication; see earlier) was our response variable; however, this variable was overdispersed. Following Zuur *et al.* (2010), we used a gamma distribution in our model to deal with the overdispersion problem, and we verified that it was controlled by using the function 'dispersion_glmer' of the package BLMECO (Korner-Nievergelt *et al.,* 2015). Using a gamma distribution, the overdispersion value was 0.73 and we could therefore neglect it. To use gamma distribution, we changed the zero data points of our response variable to 0.01 ($n = 14$ points, 0.39% of all data), as gamma can only be used with positive values. We also tested for multicollinearity of independent variables using the 'vif' function of the USDM package (Naimi *et al.,* 2014); all values were < 2 and thus we did not consider our data to be collinear (following criteria in Zuur *et al.,* 2010).

We performed backward model selection starting with the complete model which included all the binary fixed factors as described in Table 1. The random effects for all models were 'journal' and 'journal impact factor during the year of publication'. We did not include year of publication as a random effect because the variance of this variable was negligible in the model as we are already controlling for time by using the 'mean number of citations in 6 years' as the response variable. Furthermore, there is no effect of year in the mean number of citations in 6 years, as expected (estimate $= -0.002 \pm 0.003$, t-value $= -0.576$, $P = 0.56$). In a step-by-step approach, we chose the variable with the highest P-value > 0.1 to remove the variable from the model. We used an ANOVA to compare the resulting model with the variable removed with the previous model. We only kept variables that had $P < 0.1$ in the ANOVA comparison, thus avoiding possible type 2 error of stepwise model selection. As different types of model selection might have their intrinsic biases (Johnson & Omland, 2004), we also performed model selection using information criteria. We used the function 'dredge' from the package MUMIN v.1.42.1 (Bartoń, 2018) to compare all possible models from our complete model and used the criterion of ΔAIC (akaike information criterion) = 0 to select our model. For this approach, we considered all possible models with the combinations of fixed factors because we did not have prior expectations of which combinations of variables to compare; thus, this approach is also biased as it is a non-informed accession of models (Gotelli & Ellison, 2012). All models were built using the function 'glmer' of the package LME4 to calculate our models (Bates *et al.,* 2015). For the continuous variable 'number of words', we used a generalised additive mixed model (GAMM) because the relationship between citation rate and number of words might not necessarily be linear. We performed our GAMM model using the package MGCV with the function 'gamm' (Wood *et al.,* 2016). In our model, the 'mean number of citations in 6 years' was the response variable, 'number of words' was the smooth term, and 'journal impact factor during the year of publication' was the random effect. We performed a separate model to test if different types of geographic location could have an impact on citation rate, as we combined biome, country, and region, which might have different impacts on citation. The response variable was 'mean number of citations in 6 years', the fixed effects were 'biome', 'region', and 'country', and the random effects were 'journal' and 'journal impact factor during the year of publication'. All analyses were performed in R environment 3.4.1 (R Development Core Team, 2011).

To test if factors influence citation rate differently in the fields of entomology and ecology, we analysed our data for each subject area separately. We followed the same analysis as for the whole dataset described earlier; we started with the complete model and removed step by step the variables with higher $P > 0.1$, and we compared our results from stepwise selection with information criteria selection, as explained earlier. In these models, we did not include journal as a random effect, as there would be only two levels. We tested the influence of the number of words in each subject area, as explained earlier, using GAMM.

Lastly, we performed a sensitivity analysis to test whether our results were driven by over-represented terms. For most of our title features, specific terms only occurred once or at most a few times (e.g. most species Latin names, species common names, family Latin names, family common names, and geographic locations), making a sensitivity analysis impossible. However, within functional group, several entries occurred repeatedly; notably, predator was the most common functional group included in titles ($n = 250$ of 1187 that included functional group) and was commonly used in both entomology and ecology titles. To test if the pattern we found for functional groups was driven by the presence of 'predator' in the title, we removed all data points that had this functional group and reanalysed the influence of functional group on the mean number of citations in 6 years for each subject ($n = 70$ uses of predator in entomology journals or $n = 180$ uses of predator in ecology journals).

Data from this study are archived in the Dryad Digital Repository: https://doi.org/10.5061/dryad.qr77sm4 (Murphy *et al.*, 2019).

Results

When both ecology and entomology journals are considered together in the analysis, titles that did not include species and family Latin names received more citations (47% and 22% more citations, respectively) than titles that did include Latin names ($R^2 = 0.007$, Fig. 1; Table 2). Common names had no effect on citation rate for either species or family (Fig. 1; Appendix S2) nor did the inclusion of functional groups (Fig. 1; Appendix S2). When we considered geographic location, there was no effect on citation rate (Fig. 2; Appendix S2), even when analysing by biome, region, and country separately (country, $t = 1.00$, $n = 3{,}556$, $P = 0.32$; region, $t = -1.87$, $n = 3{,}556$, $P = 0.06$; biome, $t = -1.23$, $n = 3{,}556$, $P = 0.22$). Use of question mark, humour, or division into two parts did not affect citation rate (Fig. 2; Appendix S2), nor did the number of words (Fig. 3; $t = 0.27$, $n = 3{,}556$, $P = 0.79$; linear relationship). Our analysis using information criteria selection found that the model with $\Delta\text{AIC} = 0$ included only species Latin name and family Latin name (Appendix S3).

When we analysed the entomology and ecology journals separately, we found that including species Latin names did indeed have negative effects on citation rate for both subjects (16% and 17% fewer citations for entomology and ecology journals, respectively; $R^2 = 0.018$ for entomology, $R^2 = 0.014$ for ecology; Fig. 4a; Table 2), whereas family or taxonomically higher Latin names had negative effects only for ecology journals (16% fewer; Fig. 4b; Table 2). Interestingly, including the functional group resulted in more citations for entomology papers, but had a marginally negative effect in citation rate for ecology papers (16% more citations for entomology; Fig. 4c; Table 2). Similarly, including geographic location resulted in more citations for entomology papers, but had no significant effect for ecology papers (17% more citations for entomology; Fig. 4d, Table 2). Common names, question marks, whether the title attempted humour or was divided into two parts (Appendix S2), and the number of words (entomology: $t = -1.06$, $n = 1{,}048$, $P = 0.29$; ecology: $t = 0.37$ $n = 2{,}514$, $P = 0.71$) in the title did not affect citation rate for either entomology or ecology journals. Using the information criteria selection approach, the model with $\Delta\text{AIC} = 0$ for the entomology subset included species Latin name, functional group, geographic location, and presence of question mark, while the best model for the ecology subset included family and higher Latin name, species Latin name, functional group, and species common name (Appendix S4). Considering the former, the presence of a question mark was not significant in the selected model ($t = 1.47$, $P = 0.14$), nor was species common name in the latter model ($t = 1.59$, $P = 0.11$). Thus, results with both model selection approaches were similar.

We also tested if the most common functional group (predator) drove the distinct patterns found for ecology and entomology papers. Even when we excluded papers that had predator in the title, entomology papers with a functional group mentioned in the title still had more citations than titles without a functional group ($t = -2.4$, $n = 964$, $P = 0.016$). For ecology papers, when we removed predators, there was no difference in citation rate between papers with and without functional group in the title ($t = 0.84$, $n = 2{,}299$, $P = 0.4$).

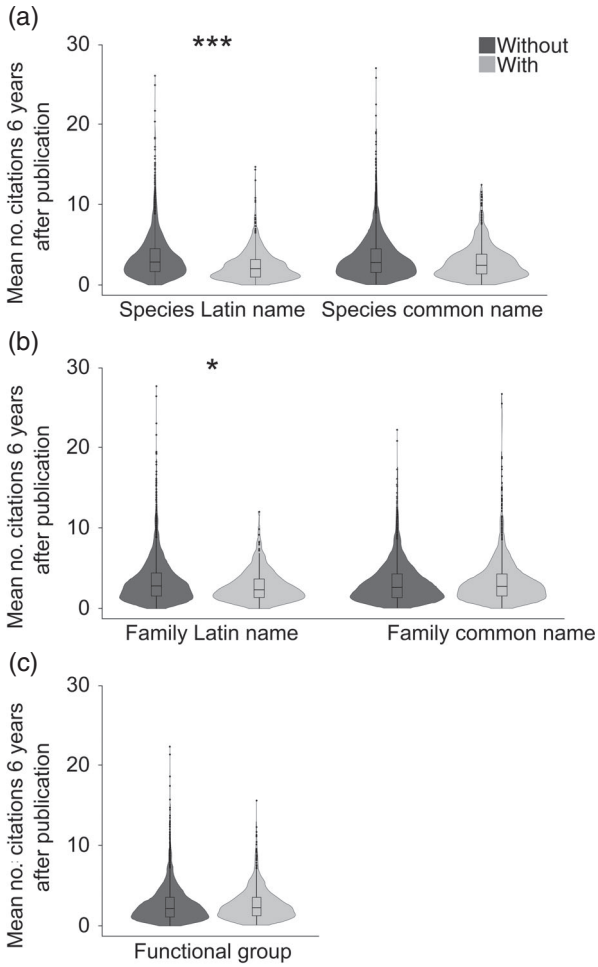

Fig. 1. Mean number of citations 6 years after publication for features that either are or are not included in the title: (a) species Latin name, species common name; (b) family Latin name (or taxonomically higher), family common name (or taxonomically higher); and (c) functional group. Both entomology and ecology journals are shown together. All plots are violin plots overlain on box plots; for the violin plot, the outer width represents the frequency of each value and the widest section represents the mode average (*$P < 0.05$; ***$P < 0.001$).

Discussion

We predicted that the presence of Latin names, common names, functional groups, and geographic location in the title would decrease citation rates of research articles in entomology and ecology journals. We had no specific predictions for how the use of punctuation, whether the title attempted humour, and the number of words would affect citation rate. Interestingly, we found that using a Latin name in a title was the only factor

Table 2. Model summary for the effect of each title feature on the number of citations when all of the journals are considered together, when only entomology journals are analysed, and when only ecology journals are analysed. Title features that remained significant in some or all of the final models are species Latin name, family Latin name (or taxonomically higher), functional group, and geographic location. See Appendix S2 for initial complete model.

	All journals				Entomology				Ecology			
Title feature	Estimate	SE	*t*	*P*	Estimate	SE	*t*	*P*	Estimate	SE	*t*	*P*
Intercept	0.38	0.06	6.5	< 0.001	0.53	0.02	21.58	< 0.001	0.26	0.01	33.1	< 0.001
Species Latin	0.053	0.01	5.11	< 0.001	0.05	0.03	2.15	0.03	0.05	0.01	4.54	< 0.001
Family Latin	0.033	0.02	2.05	0.04	–	–	–	–	0.04	0.02	2.16	0.03
Functional group	–	–	–	–	−0.07	0.02	−2.73	0.006	0.01	0.008	1.91	0.056
Geographic location	–	–	–	–	−0.07	0.03	−2.28	0.023	–	–	–	–

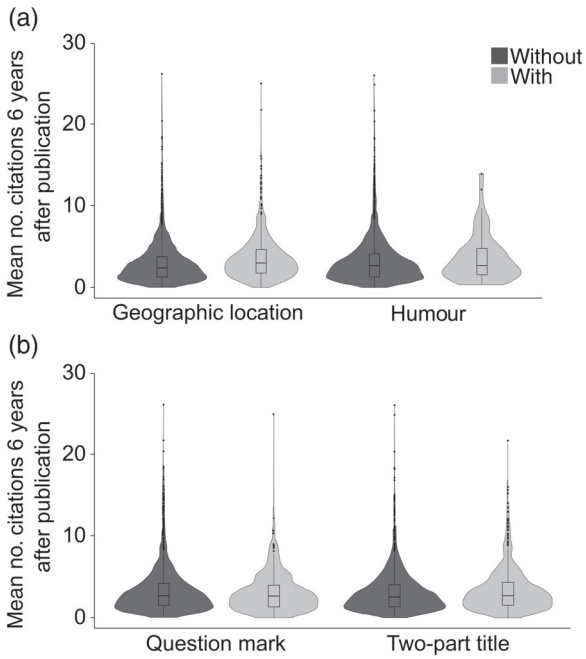

Fig. 2. Mean number of citations 6 years after publication for features that either are or are not included in the title: (a) geographic location, humour; (b) question mark, and two-part title. Both entomology and ecology journals are shown together. All plots are violin plots overlain on box plots; for the violin plot, the outer width represents the frequency of each value and the widest section represents the mode average.

to significantly decrease citations across all journals. Using Latin names in a title probably causes readers who are not familiar with that name to skip the paper entirely when perusing a journal's table of contents, whereas titles that describe a general research area, without indicating the study organism, draw a wider audience and thus more citations. Notably, when considering both ecology and entomology journals together, we found that use of common names, functional group names, geographic location, presence of punctuation, whether the title attempted humour, and title length had no effect on citation rates.

Titles that included species and Latin names were cited less often, as predicted, whereas titles with common names did not affect citation rate of the paper. Previous research has found similar results; papers in the journal *Functional Ecology* that included genus or species Latin names in the titles received fewer citations than titles that included higher taxonomic groupings, common names, or did not mention the study system (Fox & Burns, 2015). Fox and Burns (2015) also showed that including species' Latin names in titles is decreasing over time, and papers that do include them are more likely to be rejected. We found that even including the Latin name of higher taxonomic groupings (i.e. kingdom, phylum, class, order or family) led to fewer citations, perhaps because Latin names tend to be familiar only to a restricted group of people studying those specific organisms. Interestingly, we found that papers from entomology journals that had family or taxonomically higher Latin names in the title were cited at a similar rate as papers without family or taxonomically higher Latin names. We suggest that, as entomology journals are usually read by people familiar with insects and other arthropods, the readers are already familiar with the Latin names for these taxa (e.g. most entomologists are familiar with all of the insect orders). However, the field of ecology is too broad for researchers to be familiar with Latin names for the diverse taxa that ecologists study. For example, researchers who study vertebrates may not be familiar with the Latin names of insect orders, or might not think they are interested in papers about them. Similarly, many entomologists are probably not familiar with Latin names of fish, birds, or reptiles. Yet, many of these scientists study closely related research questions and could learn from similar studies on unfamiliar taxa. Thus, we advise that authors avoid including the Latin names of their study system in the title in order to broaden the readership and recognition of their work.

Our analysis that tested ecology and entomology journals separately revealed different effects of title features on citation rate. The effect of including a functional group (e.g. predator, herbivore, pathogen) in the title differed between ecology and entomology journals; inclusion of functional groups decreased citation rates in ecology journals, as we predicted, but increased citations in entomology journals. However, the pattern in ecology journals was mainly driven by the presence of 'predator' in the title. The pattern of increased citation rate when the title included functional group in entomology papers might arise because including functional group gives more information about the subject of the paper while still keeping the title broad enough to attract a wide audience. To our knowledge, we are the first to ever test the effect of functional groups in titles on the citation

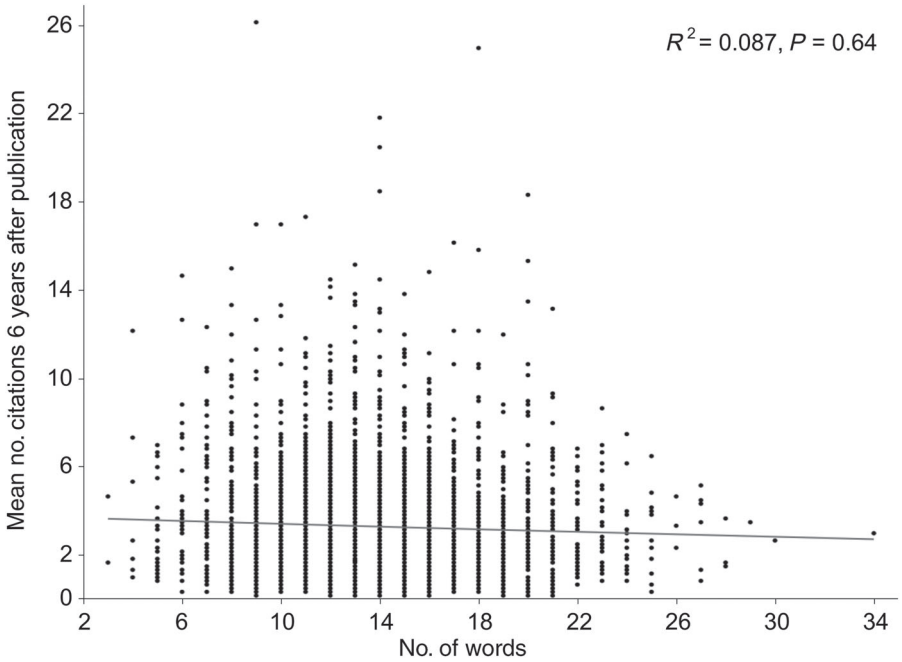

Fig. 3. Relationship between the number of words in the title and the mean number of citations 6 years after publication. Each point represents a title from a publication. Both entomology and ecology journals are shown together.

rate, and thus more research on this pattern is necessary before drawing conclusions. Besides the different effects of functional group, the effects of geographic location also differed between ecology and entomology journals. We found that including geographic location in the title increased citation rate for entomology journals, which was not what we predicted. This result contradicts previous research which found that including geographic location in article titles decreases citation rates in other fields (Jacques & Sebire, 2010; Paiva *et al.,* 2012). However, we should be cautious about this result, as very few papers in entomology journals included geographic location in the title (n = 140 out of 1,048), and therefore this pattern might be due to unbalanced data.

We found no effect on citation rate of the use of a question mark, an attempt at humour, punctuation to separate a title into two parts, or the length of a title. Other researchers have found that the use of a question mark in a title reduced citations in multiple disciplines (Jamali & Nikzad, 2011; Paiva *et al.,* 2012; Nair & Gibbert, 2016). In our study, citation rate was not affected by punctuation dividing titles into two parts, but literature reviews of other fields have found that colon use can impact citation rate in contrasting ways depending on the discipline, and thus the effect of colons on citation rates seems to be discipline-specific (science management: Jacques & Sebire, 2010, medicine: Nair & Gibbert, 2016). Lastly, we found no relationship between title length and citation rate, contradicting previous findings (positive effect: Jacques & Sebire, 2010; Habibzadeh & Yadollahie, 2010; negative effect: Subotic & Mukherjee, 2014). The differential effects of title length on citation rate appear to be similarly discipline-specific as these reviews investigated papers in medical and psychology journals.

Our investigation into the effects of title features on citation rates has some limitations and inspires future questions. For some of the title features we wished to test, we were restricted in the conclusions we could make because of the distribution of our data. For example, very few papers made attempts at humour in the title (n = 57 out of 3,562); thus, the lack of influence of this type of feature on citation rate could be due to reduced representation in our data. The problem of small sample size was even more pronounced when we subdivided the titles into ecology or entomology papers. For example, we could not draw broad conclusions about the patterns we found regarding the presence of geographic location in entomology titles due to imbalance in the data (n = 140 out of 1,048 included a geographic location in the title). Furthermore, we urge caution in interpreting our finding that including functional group in a title increases citation rate for entomology journals; it would be interesting to explore the link between citation rate and the use of functional groups in titles using a larger sample size of papers across more fields, as our study was the first to test the inclusion of functional groups. It would also be interesting to isolate differences in citation rate and download rate, because the rate of download for a journal article might be a better indicator of a title's appeal.

Our survey of entomology and ecology journal titles corroborates investigations of journal titles in other fields and highlights the risk of choosing a title that is too specific. In particular, the use of Latin names in a title reduced citation rate and probably readership for papers. Species-specific titles may undersell the work and suggest to readers that the research cannot be extrapolated beyond a specific study system, when, in fact, entomologists frequently address fundamental questions in biology.

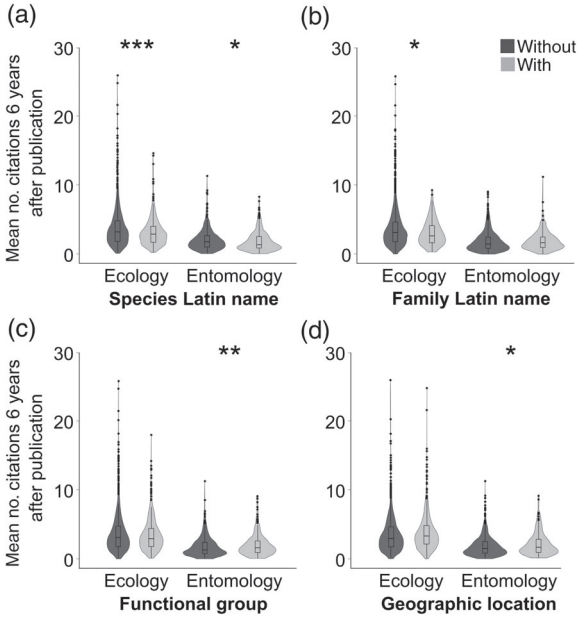

Fig. 4. Mean number of citations 6 years after publication for features that either are or are not included in the title, separated by ecology and entomology journals: (a) species Latin name; (b) family Latin name (or taxonomically higher); (c) functional group; and (d) geographic location. All plots are violin plots overlain on box plots; for the violin plot, the outer width represents the frequency of each value and the widest section represents the mode average (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

However, one should also be wary of overgeneralising by choosing a very broad title that implies far-reaching implications that are inaccurate. As with many tasks in science, writing the perfect title is a 'Goldilocks challenge' where each author must find the perfect balance. Whether or not to include study system information and how to include it will depend on the research question and the publishing journal. Our research provides guidance on how to write your perfect 'Goldilocks' title that will increase recognition and citations of your research.

Acknowledgements

We thank the University of Denver Science Librarian, Meg Eastwood, for help with downloading citation rates for each paper. We thank Laurel Cepero with help gathering data at the beginning of this project. We thank the University of Denver Ecology and Evolutionary Biology group (DUEEBs), the editor, and two anonymous reviewers for helpful feedback that greatly improved an earlier draft of this manuscript. The authors declare there are no conflicts of interest.

Author contributions

SMM conceived the idea for the research. SMM, MCV, CJH, EDB and EEB designed the data collection protocol. All authors collected the data; SMM, MCV and ESHL managed the data; and MCV analysed the data. SMM, MCV, CJH, EDB, JDW and ESHL wrote the first draft of the manuscript, and all authors helped to revise and edit the paper.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1–S4. Supplementary files.

References

Ball, R. (2009) Scholarly communication in transition: the use of question marks in the titles of scientific articles in medicine, life sciences and physics 1966-2005. *Scientometrics*, **79**, 667–679.

Bartoń, K. (2018) *Multi-Model Inference*. R package version 1.42.1.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bornmann, L. & Mutz, R. (2015) Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*, **66**, 2215–2222.

Fox, C. & Burns, C. (2015) The relationship between manuscript title structure and success: editorial decisions and citation performance for an ecological journal. *Ecology and Evolution*, **5**, 1970–1980.

Fox, C., Paine, C. & Sauterey, B. (2016) Citations increase with manuscript length, author number, and references cited in ecology journals. *Ecology and Evolution*, **6**, 7717–7726.

Gotelli, N.J. & Ellison, A.M. (2012) *A Primer of Ecological Statistics*, 2nd edn. Sinauer Associates, Inc, Sunderland, Massachusetts.

Habibzadeh, F. & Yadollahie, M. (2010) Are shorter article titles more attractive for citations? Cross-sectional study of 22 scientific journals. *Scientometrics*, **51**, 165–170.

Hartley, J. (2007) Planning that title: practices and preferences for titles with colons in academic articles. *Library & Information Science Reserach*, **29**, 553–568.

Hirsh, J.E. (2005) An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, **102**, 16569–16572.

Jacques, T. & Sebire, N. (2010) The impact of article titles on citation hits: an analysis of general and specialist medical journals. *Journal of the Royal Society of Medicine Short Reports*, **1**, 1–5.

Jamali, H. & Nikzad, M. (2011) Article title type and its relation with the number of downloads and citations. *Scientometrics*, **88**, 653–661.

Johnson, J.B. & Omland, K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology & Evolution*, **19**, 101–108.

Korner-Nievergelt, F., Rother, T., von Felten, S., Guelat, J., Almasi, B. & Korner-Nievergelt, P. (2015) *Bayesian Data Analysis in Ecology using Linear Models with R, BUGS and Stan*. Elsevier, London, U.K.

Mabe, M. & Amin, M. (2002) Dr Jekyll and Dr Hyde: author-reader asymmetries in scholarly publishing. *ASLIB Proceedings*, **54**, 149–157.

Murphy, S.M. (2007a) The effect of host plant on larval survivorship of the Alaskan swallowtail butterfly (*Papilio machaon aliaska*). *Entomologia Experimentalis et Applicata*, **122**, 109–115.

Murphy, S.M. (2007b) Inconsistent use of host plants by the Alaskan swallowtail butterfly: adult preference experiments suggest labile oviposition strategy. *Ecological Entomology*, **32**, 143–152.

Murphy, S.M., Vidal, M.C., Hallagan, C.J., Broder, E.D., Barnes, E.E., Horna Lowell, E.S. *et al.* (2019) Data from: does this title bug

(Hemiptera) you? How to write a title that increases your citations. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.qr77sm4.

Naimi, B., Hamm, N., Groen, T.A., Skidmore, A.K. & Toxopeus, A.G. (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography*, **37**, 191–203.

Nair, L. & Gibbert, M. (2016) What makes a 'good' title and (how) does it matter for citations? A review and general model of article title attributes in management science. *Scientometrics*, **107**, 1331–1359.

Paiva, C., Lima, J. & Paiva, B. (2012) Articles with short titles describing the results are cited more often. *Clinics*, **67**, 509–513.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 1(2.11.1), p. 409.

Subotic, S. & Mukherjee, B. (2014) Short and amusing: the relationship between title characteristics, downloads, and citations in psychology articles. *Journal of Information Science*, **40**, 115–125.

Wood, S.N., Pya, N. & Saefken, B. (2016) Smoothing parameter and model selection for general smooth models (with discussion). *Journal of the American Statistical Association*, **111**, 1548–1575.

Zuur, A.F., Ieno, E.N. & Elphick, C.S. (2010) A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, **1**, 3–14.

Accepted 30 March 2019

Associate Editor: Francis Gilbert