Computationally reproducing results from meta-analyses in Ecology and Evolutionary Biology using shared code and data

Steven Kambouris [1,2,*], **David P. Wilkinson** [1,2], **Eden T. Smith** [1,3], **Fiona Fidler** [1,2,3]
[1]MetaMelb Research Initiative, The University of Melbourne
[2]School of Agriculture, Food and Ecosystem Sciences, The University of Melbourne
[3]School of Historical and Philosophical Studies, The University of Melbourne
[*]Correspondence: steven.kambouris@unimelb.edu.au

ABSTRACT

Many journals in ecology and evolutionary biology encourage or require authors to make their data and code available alongside articles. In this study we investigated how often this data and code could be used together, when both were available, to computationally reproduce results published in articles. We surveyed the data and code sharing practices of 177 meta-analyses published in ecology and evolutionary biology journals published between 2015–17: 60% of articles shared data only, 1% shared code only, and 15% shared both data and code. In each of the articles which had shared both (n = 26), we selected a target result and attempted to reproduce it. Using the shared data and code files, we successfully reproduced the targeted results in 27–73% of the 26 articles, depending on the stringency of the criteria applied for a successful reproduction. The results from this sample of meta-analyses in the 2015–17 literature can provide a benchmark for future meta-research studies gauging the computational reproducibility of published research in ecology and evolutionary biology.

CONTENTS

1 INTRODUCTION

Over the past decade, meta-research (or metascience) has emerged as the term for the rigorous evaluation of research (1). The emergence of meta-research is related to discussions of replication and reproducibility across multiple disciplines, notably psychology (2), and including ecology and evolutionary biology (3–6). Replication is one focus of meta-research studies in ecology and evolutionary biology (7; 8), but the remit of meta-research encompasses topics such as the extent of selective reporting and publication bias in ecology (9; 10), the prevalence of questionable research practices amongst ecologists (11), and analytic flexibility (12). Closely related to meta-research studies identifying such problems are works and initiatives proposing solutions, based on principles of openness and transparency. Initiatives in the field include the Tools for Transparency in Ecology and Evolution (13), which was followed by the formation of the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology (SORTEE) for ecologists and biologists with an interest in transparency and open science (14).

This study contributes to the meta-research within ecology and evolutionary biology by focusing on computational reproducibility. Computational reproducibility is defined as "obtaining consistent results using the same input data; computational steps, methods, and code; and conditions of analysis" (15, p.46). By this definition, availability of both the data and code underpinning an article is a necessary prerequisite for computational reproducibility. Thus, our study of computational reproducibility is also a study of data and code availability. (Note that if data but not code were available, recalculation of results could still be attempted using the written description of statistical analysis methods to write fresh analysis code. Such an approach has been called "analytical reproducibility" and has been studied separately (16–18). Although analytical reproducibility and computational reproducibility are related concepts, in this study we focused on computational reproducibility.)

If we have the shared data and code for an article, then ideally we should be able to use both to recalculate results that match the published results. The technical difficulty of achieving this in practice is well-recognised, even for researchers returning to their own computer code years later (19). Thus, there have been a number of studies across different disciplines gauging how often results in the published literature can actually be computationally reproduced from data and code. Stodden et al. (20) evaluated the effectiveness of the data and code sharing policy implemented in the journal *Science* in 2011, by attempting to obtain data and code for 204 articles published after the policy change in order to computationally reproduce their results. They obtained data and code for 44% of articles in the sample and were able to successfully reproduce results for 26% of the sample. Wood et al. (21) assessed the computational reproducibility of 109 articles published in 2014 from journals in development, economics, and public health. Their study, described as a "research audit exercise" found that a lack of available data and code meant that reproduction could not be attempted for 71% of articles in the sample. They were able to reproduce results identical to or within rounding of the original results for 27 articles, and found only minor differences in another 5 articles. In psychology, Obels et al. (22) considered a set of 62 Registered Reports published over 2014–18, and found 36 (58%) had shared data and code, making them suitable candidates for attempting computational reproducibility. They successfully reproduced the main results of 21 of these 36 articles, which was 58% of the attempts made. More recently, Crüwell et al. (23) audited 14 articles published in a 2019 issue of *Psychological Science*, all of which had been awarded an Open Data Badge (`https://www.cos.io/initiatives/badges`) signifying that the article authors had shared the data for reproducing their results. Crüwell et al. (23) found that while all 14 articles did share data, only 6 shared code. Their attempts to computationally reproduce results from this issue found that one article was exactly reproducible, and three were reproducible with only minor differences. In ecology and evolutionary biology, ArchMiller et al. (24) attempted to computationally reproduce a sample of 80 studies published in the *The Journal of Wildlife Management* and *Wildlife Society Bulletin*. They were able to obtain data and code for 19 of the studies, and mostly or fully reproduce the results for 13 of them.

Such results reinforce the centrality of data and code sharing to computational reproducibility. Data sharing is a well-established topic in ecology and evolutionary biology, with numerous efforts to facilitate and improve data sharing, coming from both individual researchers and institutions such as journals. Journals have recognised and stressed the importance of data archiving (25–27). Researchers have created guides and compiled advice for how to best approach data archiving and sharing (5; 28; 29). There have also been efforts to review the effectiveness of data archiving policies and assess how the field is doing (30–32). Code availability in ecology and evolutionary biology has also been studied: Mislan et al. (33) surveyed 96 ecology journals in 2015, and found that only a small minority (14%) required code to be made available alongside published articles (in contrast to 38% of journals requiring data be made available alongside published articles). Culina et al. (34) repeated this survey in 2020 and found that of the same 96 journals, 75% mandated or encouraged making code available. However, despite this now common journal policy, Culina et al. (34) also found that only 27% of a sample of 346 ecology articles published 2015–19 actually shared code.

Aims and scope

We conducted computational reproducibility attempts on a sample of meta-analyses published in ecology and evolution journals over 2015–17 (the restriction to meta-analyses is explained in Section S1 of the Appendix). Our focus was on using shared data and code files to reproduce specific results. The primary outcome of the reproducibility attempts is the calculation of an overall computational reproducibility "success rate", similar to Stodden et al. (20).

This study commenced in late 2017 following rising interest in meta-research within ecology and evolutionary biology, including interest in data- and code-sharing specifically (32; 33). While the results of this study are not a reflection of what the rate of computational reproducibility in more recent ecology and evolutionary biology literature might be, they do provide a benchmark of the state of computational reproducibility during the period 2015–17, and provide a point of comparison for other evaluations of computational reproducibility over earlier or more recent periods.

We surveyed the data- and code-sharing rates of the applicable meta-analysis literature. We only counted as "shared" data/code that was reported as already available, rather than data/code that was (potentially) available upon request. It is possible that some authors of the meta-analyses included in this study may have shared their data and code in response to a request. However, a request for data or code requires an interaction between the requesting party and the article authors, and there is a possibility that the request will not be successful, for a variety of reasons (e.g., the authors are no longer contactable via the contact details provided in the article, the authors do not respond in a timely manner, the authors respond but refuse for some reason, or the authors respond but can no longer find the data and code). We decided not to request data or code from article authors in this study, because requesting data/code would introduce a element of the study that may not be reproducible by others: the success or failure of any requests would rely on factors such as timing, existing connections (of lack thereof) with authors, and the purpose behind the request.

2 METHODS

Our study had four stages: first, we obtained a sample of published meta-analyses from ecology and evolution journals; second, we assessed each meta-analysis for data- and code-sharing; third, we selected results to be reproduced using the shared data and code; and finally, we attempted to reproduce the selected results.

We curated a set of meta-analyses to survey by conducting a Scopus abstract and citation database search (see details in Section S2 of the Appendix). The search query, conducted on 20th December 2017, searched article titles, abstracts, and keywords for the string "meta-anal*", subject to two constraints. The first constraint restricted results to articles published between 2015 and 2017, inclusive. The second constraint restricted results to articles published in one of 21 ecology and evolution journal titles (identified by ISSN), which are

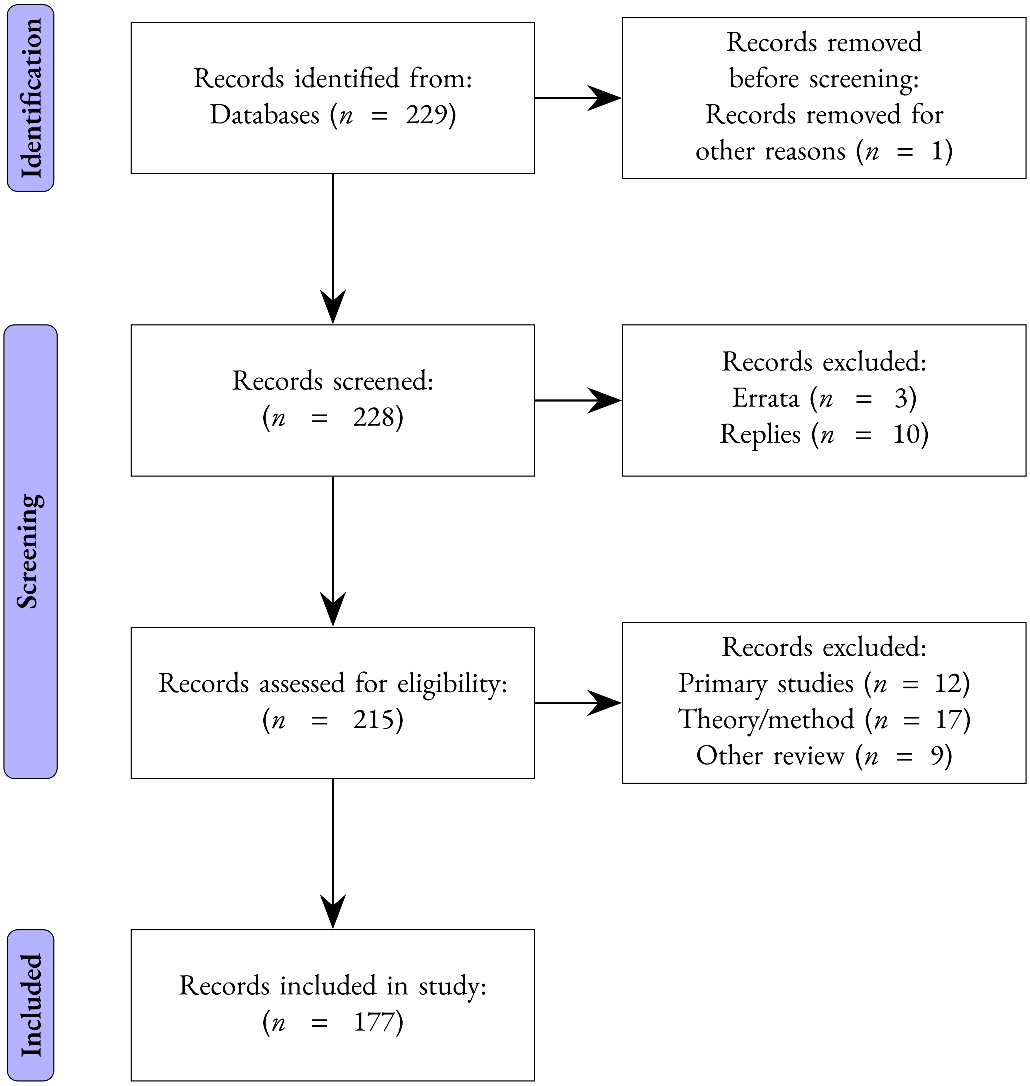

Figure 1: PRISMA-style flow diagram depicting the article selection process.

the same journal titles as used for the survey of meta-analyses conducted in Nakagawa and Santos (35).

The search yielded 229 results. One irrelevant result was found to have been included in the results due to a Scopus database error and was immediately excluded, leaving 228 results.

The search results were coded to retain only those articles which were actual meta-analysis studies, details of the coding scheme used are in Section S2 of the Appendix. The final set of ecology and evolutionary biology meta-analyses, the basis of the rest of this study, was a set of 177 articles coded as reporting to be meta-analyses. Figure 1 shows a PRISMA-style flow diagram for this study.

Recording code and data sharing in each article

Each meta-analysis article in the set of 177 was assessed for data and code sharing using the coding scheme detailed in Section S4 of the Appendix. It was expected that "data" (curated, formatted information—both numeric and text-based—that was the "raw material" for reported calculations and analyses) would be presented in one or more formatted computer files (e.g., in comma separated values format), possibly accompanied by additional computer documents containing metadata or explanations of the data files' contents. Following Mislan et al. (33) and Culina et al. (34), we regarded "code" as referring to computer code, specifically analysis code, designed to do tasks such as importing and manipulating data and performing statistical calculations based on data (e.g., calculating summary statistics or fitting models). Code may have been written in a programming language (e.g., R or Python) or it may have been syntax designed to be run by a dedicated statistical analysis software package such as SPSS, SAS, or STATA.

"Sharing" meant that the authors of the article had made data and code files available alongside the publication of the article. For journals which were not Open Access, data and code files provided as supplementary materials on publishers' websites were typically hidden behind subscriber paywalls along with the articles and were not available to everyone. We made the decision that data and code provided in this way counted as having been shared for the purposes of this study. It is for this reason that we have chosen to refer to "shared" data and code rather than "open" data and code, since "open" carries with it connotations about availability and accessibility that may not apply to data and code files provided as supplementary material behind a publisher's paywall.

We also reviewed the methods section of each article for references to the use of software. If an article did not report any details of software used, we reviewed supplementary documentation if supplied. The review process is detailed in Section S5 of the Appendix.

Selecting target results for computational reproduction

For each article in the subset of meta-analysis articles with both shared data and code, we identified a numeric "target" result that would be the basis of the computational reproduction attempt. Selecting a single result from an article involved subjective judgment, and could potentially be manipulated to increase or decrease the chance of success of reproducing each result. To mitigate this risk, we used the following process to identify a target result: our target result would be the first meta-analytic summary effect (consisting of the point estimate, the sample size, and the measure of uncertainty such as a confidence interval) reported in the results section of each article. The reasoning for this strategy is as follows: (i) in general, summary effects are commonly reported in meta-analyses, and so this would identify like results across articles; and (ii) identifying the first reported result is a consistent method of selection across articles that minimises (but does not eliminate) the need for interpretation and therefore reduces the risk of bias. A procedure which allowed for results to be deliberately chosen for computational reproduction could potentially be selected on the basis of perceived ease of reproduction.

In practice, identifying and extracting the first reported meta-analysis summary effect was complicated by two factors. First, articles presented results in different ways: some articles reported results in the body of the text while others referred to a table or figure. We extracted numerical values directly from in-text results and from results presented in tables. For results presented graphically in figures, we extracted numerical results using the software package WebPlotDigitizer version 4.4 for the Windows platform. We rounded all values extracted from figures to two decimal places. Figures required additional interpretation if they plotted multiple summary effects. In these cases, we prioritised extracting the "overall" summary effect if it existed, and otherwise selected the "first" plotted result, according to the layout of the figure (e.g., either the leftmost or topmost result). Frequently, a result was reported in-text and also expressed in a figure/table; we prioritised extracting in-text results over results reported in figures/tables. The second factor was that not all meta-analysis articles actually reported a summary effect result. In these cases, we extracted numerical values for the first-reported result of any kind associated with the meta-analysis.

Reproducing results and results comparison

For each article, we assessed the shared data and code for its relevance to the identified target result using the following strategy: (i) where available, we consulted documentation accompanying the data and code files; (ii) we examined any comments made within the code syntax files; (iii) where available, we examined the metadata of data files; (iv) we examined the contents of data files directly, looking for clues in variable names and data formats; (v) we examined the syntax of code files directly, looking for clues in function names and the kinds of calculations made. This approach was sufficient to discern with confidence whether the data and code files were applicable to the re-calculation of the target result. We went ahead with attempting to reproduce the target result for each article where both the shared data and code were found to be relevant.

In cases where the code and/or data was not relevant to the identified target result, we stopped attempting to reproduce those particular target results. Rather than do nothing further with these cases, we returned to the article and identified an alternative target result that was relevant to the shared data and code and reported the results of these reproduction attempts separately.

Each reproduction attempt was packaged as a reproducible document written in RMarkdown contained within a controlled computational environment using Docker (details are in Section S9 of the Appendix). Where code could be successfully run, reproduced target results were compared with the originally published values. For each target result (which consisted of a set of numbers e.g., summary effect estimate, confidence interval bounds, and sample size), we followed the method used in Hardwicke et al. (17) and quantified the difference between the original published value and reproduced value by calculating the relative error, expressed as a percentage: $\delta = 100 \times |x_R - x_O| / |x_O|$, where x_O is the original reported result value and x_R

is the reproduced result value. Note that the relative error is undefined when the original value is zero, and can have a large value when $|x_R - x_O|$ is greater than $|x_O|$. Following Hardwicke et al. (17), we distinguished between three categories of error, exact matches ($\partial = 0\%$), minor numerical discrepancies ($0\% < \partial < 10\%$) and major numerical discrepancies ($\partial \geq 10\%$). Although we calculated the relative error for all target values, for reporting purposes we introduced a category of matches to the rounding precision of the original result: if an original result value was 1.51 (reported to two decimal places), we considered reproduced values of 1.50 and 1.52 (±0.01) to be matched to rounding precision.

3 RESULTS

The 177 meta-analyses were located within the 21 journals as shown in Table 1. The table also shows the total number of articles from each journal returned by the literature search. Note that neither *Evolutionary Ecology* or *The Quarterly Review of Biology* were found to have published any articles which reported to be meta-analyses over 2015–17 (the literature search did not return any results at all from the journal *Evolutionary Ecology*). The journal found to have the most meta-analyses during 2015–17 was *Biological Reviews*, followed by *Oikos*. The meta-analyses in the sample were fairly evenly spread across the three years searched, as shown in Table 2. Note that six articles have a publication year of 2018; these articles had all been published online during 2017 (and so were picked up in the literature search), but at the time of the literature search had not yet been assigned to a journal issue. These six were subsequently published in journal issues dated in 2018. We kept these six journal articles and regarded them as published in 2017.

Rates of data and code sharing

When articles were reviewed for data sharing (as per the coding scheme described in Section S4 of the Appendix), a clear majority of 78% or 138 meta-analyses indicated that data had been shared in some manner. Despite the positive indication, in five cases data files could not actually be obtained. This meant that the effective data sharing rate among this sample of meta-analysis articles was 75% (133 out of 177).

The rates of code sharing were much lower in comparison to data sharing: we were able to obtain code files for 16% of meta-analysis articles (28 out of 177). This was one less than the number of articles which had indicated code was available. Of the 28 articles with code, 26 had shared data too, meaning that 15% of articles (26 of 177) in this sample shared both data and code. Section S6 of the Appendix breaks down data and code sharing rates by journal.

Journal Title	Meta-analysis N	Meta-analysis %	Other N	Other %	Total N	Total %
Biological Reviews	24	13.6	5	9.8	29	12.7
Oikos	22	12.4	2	3.9	24	10.5
Ecology Letters	19	10.7	1	2.0	20	8.8
New Phytologist	18	10.2	5	9.8	23	10.1
Ecology	13	7.3	9	17.6	22	9.6
Journal of Applied Ecology	10	5.6	2	3.9	12	5.3
Molecular Ecology	10	5.6	5	9.8	15	6.6
Oecologia	10	5.6	1	2.0	11	4.8
Functional Ecology	9	5.1	1	2.0	10	4.4
Journal of Ecology	7	4.0	0	0.0	7	3.1
Journal of Animal Ecology	6	3.4	3	5.9	9	3.9
Ecological Monographs	5	2.8	0	0.0	5	2.2
Behavioral Ecology	4	2.3	3	5.9	7	3.1
Evolution	4	2.3	0	0.0	4	1.8
Journal of Evolutionary Biology	4	2.3	10	19.6	14	6.1
Animal Behaviour	3	1.7	2	3.9	5	2.2
Behavioral Ecology and Sociobiology	3	1.7	0	0.0	3	1.3
Ecological Applications	3	1.7	0	0.0	3	1.3
The American Naturalist	3	1.7	1	2.0	4	1.8
The Quarterly Review of Biology	0	0.0	1	2.0	1	0.4
Evolutionary Ecology	0	0.0	0	0.0	0	0.0
Total	*177*	*100.0*	*51*	*100.0*	*228*	*100.0*

Table 1: Breakdown of the 177 identified meta-analysis articles by journal title.

Publication Year	N	%
2015	56	31.6
2016	61	34.5
2017	60	33.9
Total	*177*	*100.0*

Table 2: Breakdown of the 177 identified meta-analysis articles by publication year. Articles with publication year 2017 includes six articles which were first published online in 2017 before being assigned to a journal issue dated in 2018.

Characteristics of shared data and code

Figure 2 lists the locations of the shared data files for the 133 articles. The majority of articles that shared data did so on the journal publisher's website (58%, $n = 77$): in these cases, the data file(s) had been uploaded as supplementary material to the article. The Dryad Digital Repository (36) was the next most common location to share data (35% or 46 articles), followed by the Figshare (8%, $n = 11$) and Zenodo (1.5%, $n = 2$) repositories. One article was judged to have shared the data for its meta-analyses in tables presented within the published article itself: the article mentioned that the effect sizes and other details for all the individual studies included in the meta-analysis calculations were provided across two tables.

Figure 3 shows the types (formats) of data files shared by the 133 articles. The most common format for

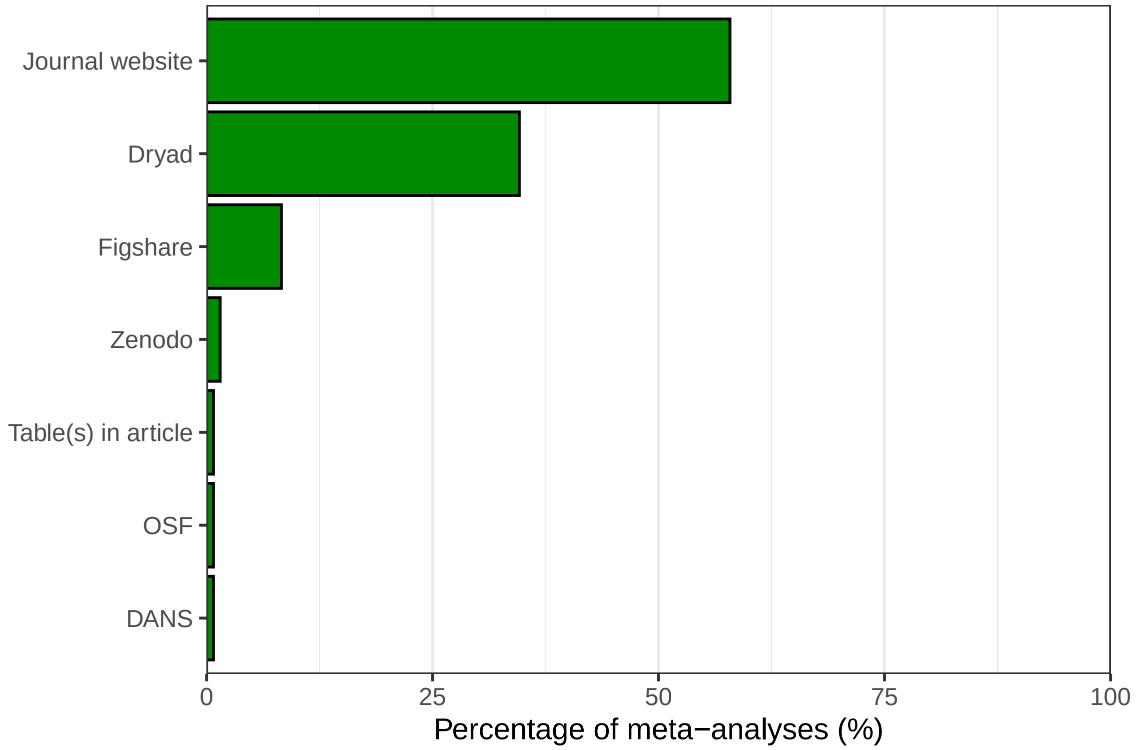

Figure 2: Breakdown of the locations where articles shared data online. Note that some articles shared data files in more than one location; both locations were counted, so the percentages indicated add up to more than 100%.

data files was a Microsoft Excel spreadsheet (44%, $n = 59$); this included both the binary XLS format and the Open XML XLSX format. The next most common format was the comma separated values (CSV) format (25%, $n = 33$). Data in a variety of plain text formats was shared by 15% of articles ($n = 20$): this included files containing phylogenetic data in NEXUS or Newick tree format. A substantial minority of articles shared tabular data in document formats like Microsoft Word Document formats DOC and DOCX (17%, $n = 22$), Portable Document Format PDF (14%, $n = 19$), Hypertext Markup Language HTML (2%, $n = 3$), and one article shared data in Rich Text Format RTF (1%). Two articles shared data files with a binary format: one article shared a data file in RData format, a binary file used by the R language, and one article shared multiple data files in a proprietary binary format associated with data logging equipment.

Table 3 breaks down the type (i.e., language or compatible software environment) of code shared by the 28 meta-analysis articles which shared code. The majority of articles shared R code (26 out of 28, 93%): 25 shared only R code, and one article shared R code and C++ code, which were designed to work together. The remaining two articles shared FORTRAN code and Python code respectively.

Software mentioned in articles

Overall, 171 meta-analysis articles (97%) mentioned at least one specific software package that was used during the study, whether mentioned in the article text or in supplementary material. The R software environ-

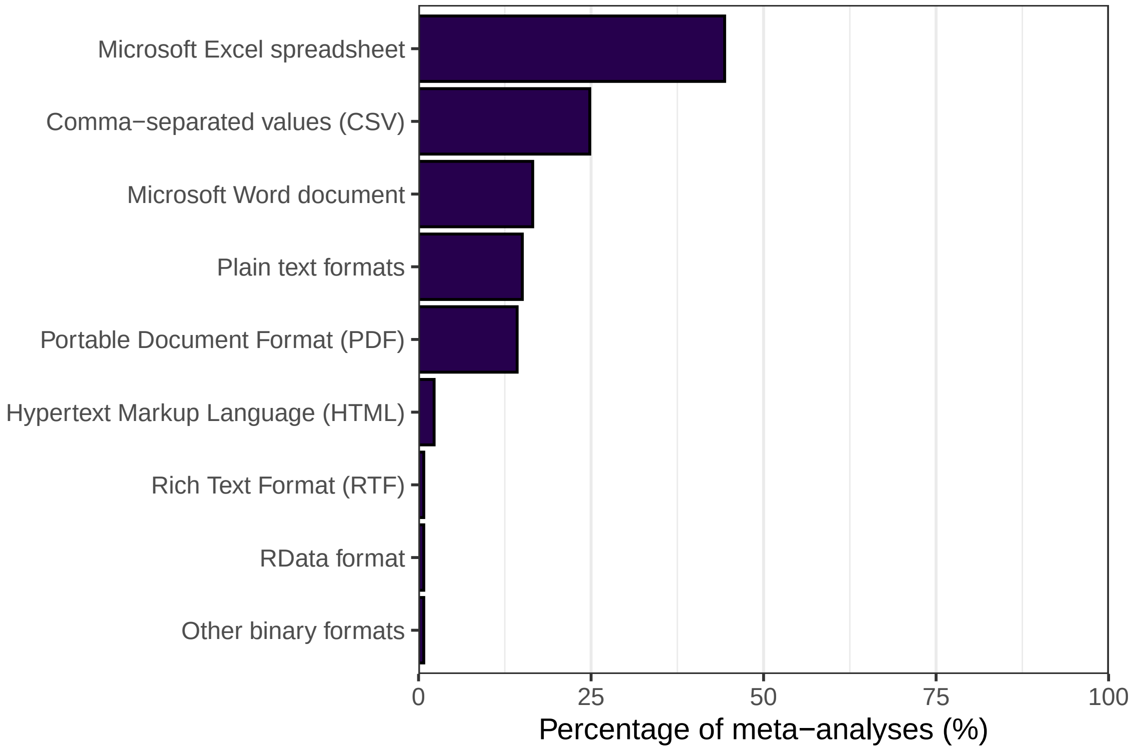

Figure 3: Breakdown of the types of file format shared by each article. Some articles shared data files of more than one type, and both types of file were counted (multiple files of the same file format only counted as one). This means that the percentages will add up to more than 100%.

Type of code shared	N	%
FORTRAN	1	3.6
Python	1	3.6
R	25	89.3
R and C++	1	3.6
Total	28	100.0

Table 3: The 28 code-sharing meta-analysis articles broken down by type of code shared.

ment was the most commonly mentioned software package with nearly 80% of articles mentioning R. The next most commonly mentioned piece of software was MetaWin; 11% of articles mentioned using it. The specialised meta-analysis software package CMA was mentioned by two articles, or 1% of the sample. The full list of all software packages mentioned is in Section S7 of the Appendix. Due to the popularity of R in this sample, and the specifics of its package system, R and R packages were summarised separately from the non-R software packages.

There were 144 mentions of software packages that were not the R software environment or an R package. The majority of these mentions were accompanied by a reference: 83 (58%) included a complete citation that appeared in the article's reference section, and 39 (27%) included a short in-text reference. The short in-text references included simple mentions of the software publisher or author, and/or a URL to the software's website. Only 15% of these software package mentions had no citation of any kind. A majority of these software package mentions (95, or 66%) also specified which version of the software package was used.

Nearly 80% (141) of meta-analysis articles mentioned using the R software environment. The majority of these mentions of R included a citation: 86 (61%) included the citation in the reference section and 21 (15%) included a short in-text reference. The version of R used was mentioned in 88 (62%) articles (see Table S9 in Section S7 of the Appendix). In total, there were 257 mentions of specific R packages: 220 (86%) included a full citation and 3 (1%) a short in-text reference. The most common R package mentioned was the *metafor* package (37), mentioned by 75 articles (53% of the articles which mentioned R). Package versions were mentioned in 58 (23%) cases. A table listing all R packages mentioned in articles is provided in Table S8 in Section S7 of the Appendix.

Reproducing target results

We used the subset of 26 articles with both shared data and shared code for the reproduction attempts. For each article we selected a target result; in 22 cases, we were able to identify what we termed a "summary effect" result: a mean, correlation, or model parameter such as slope derived from the data collected for the meta-analysis. These target results are detailed in Table 4. In the other 4 cases, the articles did not report such a result, but instead a variety of different results from an eclectic set of analyses. These other results are specified by article in Tables S10-S15 in Section S8 of the Appendix.

ID	Study	Result source	Effect size type	N	Estimate	Uncertainty
MA016	(38)	in text (p.1100)	Pearson's *r*	49	-0.83	$<$ 0.001 (*p*-value)
MA060	(39)	in text (p.674)	Fisher *z*-transformation	37	0.044	(-0.174, 0.289) (95% HPDI)
MA062	(40)	in text (p.1115)	Hedges' *d*	37	-0.205	(-0.444, 0.035) (95% CI)
MA065	(41)	in text (p.80)	Hedges' *g*	703	-8.42	(-10.73, -6.63) (95% CI)
MA067	(42)	in text (p.306)	Hedges' *g*	52	-0.21	0.07 (SE), -2.7 (*z*-score), 0.006 (*p*-value)
MA068	(43)	in text (p.14)	odds ratio	75	1.82	(1.37, 2.41) (95% HPDI)
MA071	(44)	Figure 3A (p.538)	response ratio	50	-0.26	(-1.02, 0.51) (95% CI)
MA074	(45)	in text (pp.2795-2796)	Pearson's *r*	43	0.183	(0.089, 0.274) (95% CI)
MA081	(46)	in text (p.5351)	slope parameter	1296	1.30	(0.95, 1.66) (95% CI)
MA091	(47)	in text (p.2556)	Cohen's *d*	65	0.56	(0.42, 0.69) (95% CI)
MA095	(48)	Figure 3A (pp.1495-1496)	Fisher *z*-transformation	25	0.76	(0.61, 0.91) (95% CI)
MA126	(49)	in text (p.83)	log odds ratio	n.s.	-1.11	0.49 (SE), -2.28 (*z*-score), 0.023 (*p*-value), (-2.06, -0.15) (95% CI)
MA145	(50)	in text (p.366)	Fisher *z*-transformation	118	-0.08	(-0.22, 0.03) (95% HPDI), 38 ($N_{studies}$), 25 ($N_{species}$)
MA147	(51)	in text (p.66-69)	percentage	49	0.13	0.030 (SE), (0.074, 0.19) (95% CI)
MA155	(52)	in text (p.565)	Pearson's *r*	n.s.	0.51	0.01 (*p*-value)
MA188	(53)	in text (p.653)	log response ratio	818	-0.363	(-0.408, -0.318) (95% CI)
MA191	(54)	in text (p.92)	slope parameter	553	0.86	(0.77, 0.94) (95% CI)
MA198	(55)	in text (p.4595)	Fisher *z*-transformation	79	-0.41	(-0.55, -0.27) (95% CI)
MA202	(56)	in text (pp.1072-1073)	Hedges' *d*	329	-0.330	(-0.503, -0.156) (95% CI)
MA211	(57)	Figure 2 (p.374)	log response ratio	3298	0.24	(0.23, 0.25) (95% CI)
MA213	(58)	in text (p.2004)	difference in means	654	-0.07	0.362 (*p*-value)
MA229	(59)	Figure 3 (p.256)	log response ratio	57	0.40	(0.24, 0.53) (95% CI)

Table 4: Details of the 22 summary effect target results selected for reproduction attempts. In the table, the following abbreviations are used: CI – confidence interval; HPDI – highest posterior density interval; SE – standard error; n.s. – not stated

There were 173 separate values across the 26 target results from the articles with both data and code, with an average of 6.7 values making up each target result. This included summary effect estimate values, sample size values, measures of uncertainty such as lower and upper bounds of confidence intervals described in Table 4, and other values described in Tables S10-S15 in Section S9 of the Appendix.

Table 5 summarises the relevance of the articles' shared code to the target results: Of the 22 articles with summary effect target results, 19 had relevant code and one had partially relevant code. Of the 4 articles with other target results, one had relevant code and two had partially relevant code. The remaining cases did not have relevant code. "Not relevant" meant that the shared code performed calculations or analyses that were unrelated to the calculation of the target result selected for reproduction or any meta-analysis results (the code conducted simulations or analysed experimental data instead.) "Partially relevant" code performed calculations or analyses that related to meta-analysis results, but not the target result selected for reproduction. The "not relevant" and "partially relevant" code could not be used to reproduce the target result.

We judged 20 out of 26 articles with shared data and code (77%) to have code relevant to the target result and attempted to reproduce those 20 results.

We attempted to reproduce the 108 target results associated with the 20 articles with relevant code. The reproduction attempt for each article was fully documented in a report; refer to Section S9 of the Appendix for details. We regarded the 65 target results associated with the six articles with irrelevant/partially relevant code as failed attempts (we return to these articles in the next section). Table 6 summarises the results of the reproduction attempts of the target results.

Table 6 shows that just under 50% of target results could either be reproduced either exactly or differed only by the rounding precision of the original value (rounding or floating point errors could explain these discrepancies). Of the remaining target results, thirteen differed from the original value by less than 10%, three reproduced values differed from the original value by 10% or more, and there were six target results from three articles that could not be reproduced at all; the circumstances of these six failures are described in Table 7.

The summary of the reproduction attempts in Table 6 counts every target result value separately, whether an effect size point estimate, a lower or upper bound of a confidence interval, or a sample size. Calculating a reproducibility success rate over the total number of values in this way does not consider that the sets of values from each article are inter-dependent, and so the success or failure in reproducing one value from an article may not be considered to be independent of the success or failure in reproducing another value from

ID	Study	Result type	Code relevance
MA016	(38)	summary effect	not relevant
MA060	(39)	summary effect	relevant
MA062	(40)	summary effect	relevant
MA065	(41)	summary effect	relevant
MA067	(42)	summary effect	relevant
MA068	(43)	summary effect	partially relevant
MA071	(44)	summary effect	relevant
MA074	(45)	summary effect	relevant
MA081	(46)	summary effect	relevant
MA091	(47)	summary effect	relevant
MA092	(60)	other result	not relevant
MA094	(61)	other result	partially relevant
MA095	(48)	summary effect	relevant
MA126	(49)	summary effect	relevant
MA129	(62)	other result	relevant
MA145	(50)	summary effect	relevant
MA147	(51)	summary effect	relevant
MA155	(52)	summary effect	not relevant
MA188	(53)	summary effect	relevant
MA191	(54)	summary effect	relevant
MA198	(55)	summary effect	relevant
MA202	(56)	summary effect	relevant
MA211	(57)	summary effect	relevant
MA212	(63)	other result	partially relevant
MA213	(58)	summary effect	relevant
MA229	(59)	summary effect	relevant

Table 5: Summary of reviews to gauge the relevance of shared code to each target result.

Outcome of target result reproduction attempt	N	%
Original and reproduced values match exactly	75	43.4
Original and reproduced values differ by rounding precision	11	6.4
Original and reproduced values differ by less than 10%	13	7.5
Original and reproduced values differ by 10% or more	3	1.7
Failed, could not calculate any value for target result	6	3.5
Failed, code not relevant to target result	65	37.6
Total	173	100.0

Table 6: Breakdown of the reproduction attempt outcomes for the 173 target results.

the same article. The possibility of dependency of reproduction success between the different target values within an article is examined in Section S9 of the Appendix.

ID	Study	Target result(s)	Description
MA081	(46)	2 values (upper and lower confidence interval limits)	The code uses bootstrapping to calculate the reported confidence interval, but we encountered an error: the bootstrapping procedure as coded creates random data from which the bootstrapped value cannot be calculated, making it impossible to complete the bootstrap calculation.
MA211	(57)	4 values (summary effect estimate, upper and lower confidence interval limits, sample size)	There is a mismatch between the supplied data and code: the code that would clearly calculate the target results attempts to subset the supplied data using a variable that does not appear anywhere in any shared data files.

Table 7: Descriptions of the failures to reproduce target results.

The original and reproduced values for the summary effect size target results are compared in Table 8. Apart from one failure to reproduce a summary effect size (MA211), the reproduced values were close to the originally reported values. All reproduced summary effect sizes were in the same direction as the original. There were nine exact matches between original and reproduced values. Of those that were not exact matches, six (MA060, MA062, MA071, MA191, MA198, MA229) differed by the rounding precision of the original values, and so were off by ±0.001 (where reported to 3 decimals places) or ±0.01 (where reported to 2 decimal places). Also, five cases with discrepancies (MA060, MA062, MA065, MA198, MA202) used methods which relied on random number generation (Markov chain Monte Carlo and multiple imputation). The code for these articles did not include information about setting a random seed, and so it was not possible to recover the precise target result value as originally calculated by the code.

ID	Study	Effect size type	Original	Reproduced	Percent error (%)
MA060	(39)	Fisher *z*-transformation	0.044	0.043	2.27*
MA062	(40)	Hedges' *d*	-0.205	-0.204	0.49*
MA065	(41)	Hedges' *g*	-8.42	-8.87	5.34
MA067	(42)	Hedges' *g*	-0.21	-0.21	0.00
MA071	(44)	response ratio	-0.26	-0.27	3.85*
MA074	(45)	Pearson's *r*	0.183	0.185	1.09
MA081	(46)	slope parameter	1.30	1.30	0.00
MA091	(47)	Cohen's *d*	0.56	0.56	0.00
MA095	(48)	Fisher *z*-transformation	0.76	0.76	0.00
MA126	(49)	log odds ratio	-1.11	-1.11	0.00
MA145	(50)	Fisher *z*-transformation	-0.08	-0.08	0.00
MA147	(51)	percentage	0.13	0.13	0.00
MA188	(53)	Log response ratio	-0.363	-0.363	0.00
MA191	(54)	allometric slope parameter	0.86	0.85	1.16*
MA198	(55)	Fisher *z*-transformation	-0.41	-0.42	2.44*
MA202	(56)	Hedges' *d*	-0.330	-0.340	3.03
MA211	(57)	log response ratio	0.24		
MA213	(58)	difference in means	-0.07	-0.07	0.00
MA229	(59)	log response ratio	0.40	0.39	2.50*

Table 8: Original and reproduced values of target summary effect sizes, for articles with relevant code. Percent errors marked with * indicate that these results differed only by the rounding precision of the original values.

A full table showing comparisons of original and reproduced values for all target results is provided in Table S16 in Section S9 of the Appendix.

Reproducing target results when code not relevant

The previous section identified six cases where the code shared with the article was only partially relevant or not relevant to the article's meta-analysis results. There were three cases with shared code judged partially relevant, and three cases with shared code judged not relevant (these cases are described in detail in Section S10 of the Appendix).

As described earlier, the target results for these articles were regarded as failed reproduction attempts. However, we reviewed the code and data for these articles again, with the following in mind: (i) where the shared code was at least partially relevant to the meta-analysis in the article, could the code that *had* been shared be used to reproduce an alternative meta-analysis target result, and (ii) where the shared code was clearly not relevant to the meta-analysis, was the shared data and meta-analysis methods description in the article enough to allow us to write code to successfully reproduce the selected target result. The results of assessing two articles fitting scenario (i) are described in Section S10 of the Appendix; one article's code despite being partially relevant was judged unworkable and so was treated as part of scenario (ii) along with the three articles with code not relevant.

Table 9 breaks down the outcomes of the analytical reproduction attempts when writing new R code: we were able to calculate a value to compare to the original for all target results from the four articles considered. There were 44 exact matches between original and reproduced values (75%), and of the non-exact matches, one differed by the rounding precision of the original value, ten (17%) reproduced values were within 10% of the original values, and three (5%) reproduced values were more than 10% from the original values. The was also one case of a non-numeric text string not matching the original text string.

Outcome of target result reproduction attempt	N	%
Original and reproduced values match exactly	44	74.6
Original and reproduced values differ by rounding precision	1	1.7
Original and reproduced values differ by less than 10%	10	16.9
Original and reproduced values differ by 10% or more	3	5.1
Original and reproduced values differ (non-numeric target result)	1	1.7
Total	59	100.0

Table 9: Breakdown of reproduction attempt outcomes for 59 target results from articles with irrelevant code. The irrelevant code shared by four articles (MA016, MA092, MA155, and MA212) required the writing of entirely new code to attempt to reproduce their target results. In this table, "N" refers to the number of reproduction attempts falling into each outcome category, and "%" expresses this as percentage out of all 59 of these attempts.

As these results show, the reproduction attempts using newly-written R code were largely accurate, even though they did not constitute a computational reproducibility attempt evaluating both the shared data and code of the articles, as was the case for the results in the previous section.

Computational reproduction success rates

The overall computational reproducibility success rate for this study depends on how it is defined. Different definitions lead to different values for the numerator and denominator in the calculation. We considered the success rate in terms of the number of meta-analysis articles with successful reproductions of the target results. Since multiple target result values were identified in each of the 26 articles with shared data and code, the reproduction success on each individual target result value needed to be collapsed into a single result at the article level. There were different approaches to this, with varying levels of strictness.

Table 10 reports the overall computational reproducibility success rates for different collapsing approaches across two scenarios: (i) when all six code-irrelevant cases were considered failures by default (and thus only the 20 articles with target result-relevant code could be potential successes), and (ii) when the reproduction attempts from both the 20 articles with target result-relevant code *and* the four articles where we wrote new R code were included in the success calculations (the two articles where alternative target results were selected in order to evaluate the shared code were still considered failures by default). In addition, for each scenario, two success rates were calculated: one which expressed the number of successful article reproduction attempts as a percentage of all 177 meta-analysis articles in the sample, and the other which expressed the number of successful article reproduction attempts as a percentage of the subset of 26 meta-analysis articles which shared code and data.

Depending on the level of stringency applied to count as a success, the success rate for the code-relevant cases only was in the range of 4.0–10.7% of all articles in the sample (or 26.9–73.1% of articles with code and data). Including the cases where new code was written for the code-irrelevant cases raised the success rate, with a range of 5.1–13.0% of all articles in the sample (or 34.6–88.5% of the articles with code and data).

4 DISCUSSION

In their study of the availability of code in ecology, Culina et al. (34) estimated the proportion of the ecology literature surveyed that was *potentially* computationally reproducible. The threshold for articles to be potentially reproducible was that (seemingly) all the code and data required to reproduce results was shared, with the assumption that in practice shared code as well as data was required for reproducibility. They found that 20% of literature published in 2015–16 and 21% published in 2018–19 was potentially reproducible.

	All code-irrelevant cases considered failures			Including attempts where new code was written for code-irrelevant cases		
Result for article	*N*	Success rate (%), all	Success rate (%), subset	*N*	Success rate (%), all	Success rate (%), subset
All target result values match original exactly	7	4.0	26.9	9	5.1	34.6
At least 50% of target result values match original exactly	13	7.3	50.0	16	9.0	61.5
All target result values match original exactly or to rounding precision	9	5.1	34.6	11	6.2	42.3
At least 50% of target result values match original exactly or to rounding precision	17	9.6	65.4	21	11.9	80.8
All target result values within 10% of original	15	8.5	57.7	17	9.6	65.4
At least 50% of target result values within 10% of original	19	10.7	73.1	23	13.0	88.5

Table 10: Reproducibility success rates at the article level for different collapsing criteria. In this table, *N* is the number of articles meeting each collapsing criterion, "success rate (%), all" expresses *N* as a percentage of all 177 meta-analysis articles in the sample, and "success rate (%), subset" expresses *N* as a percentage of the subset of 26 articles with shared data and code. In the first three columns of this table, the articles with data and code judged irrelevant to the target results were considered failures by default. In the last three columns, reproduction attempts where we wrote new code to reproduce the target results were included in success calculations.

In this study, we found that 14.7% of articles in our 2015–17 sample (26/177) shared both code and data. Thus, under a definition of computational reproducibility that requires both data and code (used in both Culina et al. (34) and here) we found that 15% of articles had the *potential* to have results computationally reproduced.

Comparing this result to the results in Culina et al. (34) is not entirely like for like, since different sets of journals and time periods were surveyed and this study was restricted to meta-analyses exclusively while Culina et al. (34) was not. Nevertheless, both studies generally agree that the potential for ecology literature to be computationally reproducible was low during the period 2015–17, using the reasonable threshold of 20% as a "low" occurrence rate.

Of course, this study went further than Culina et al. (34) and actually attempted to computationally reproduce results. As seen in Table 10, failures to reproduce results and the discovery that some code was irrelevant resulted in an *actual* computational reproducibility rate of 4.0–10.7% (depending on the criterion for success applied). This actual success rate(s) can be compared with the success rate observed in ArchMiller et al. (24): 8 out of the 74 suitable articles (published 2016–18) reviewed were found to be fully reproducible, and a further 5 out of 74 articles partially computationally reproducible, for a success rate of 11% (fully reproducible only) or 18% (fully and partially reproducible). (Although 74 out of an original 80 articles were reviewed in total, the researchers could only obtain data and code and thus make a reproducibility attempt for 19 of those articles.) The difference in methods for reporting reproducibility success differed between ArchMiller et al. (24) and this study, making a direct comparison difficult to interpret: ArchMiller et al. (24) rated the computational reproducibility of articles on a five-point scale which required some qualitative judgment by the researchers, while this study has reported multiple success rates according to different sets of quantitative criteria for success. In addition, in the ArchMiller et al. (24) study, authors of the original articles were contacted to request data and code, which might have contributed towards the higher success rate observed.

In Culina et al. (34), ArchMiller et al. (24), and this study, the low rates of reproducibility (potential or actual) were driven by the low rates of ecology and evolutionary biology articles with both shared data and code. While presenting results in the context of all articles surveyed is clearly warranted, calculating computational reproducibility success rates in this way masks the extent to which data and code, once obtained, can be used to successfully reproduce results. As seen in Table 10, among the subset of articles where computational reproduction was actually attempted, the success rates are much higher as the denominator has been reduced from 177 to 26. Thus, when both data and code were available for an article, *all* target results could be matched *exactly* in 27% of cases. Relaxing the threshold required to rounding precision rather than strictly exact, all target results could be matched in 35% of cases. Although it is still interesting to investigate precisely why the shared data and code do not produce the exact same results more often than this, these results are heartening: the availability of data and code did allow for the exact or close reproduction of results in a substantial fraction of cases. And while this study has included strict criteria for what counts as a success, the level of stringency researchers place on the accuracy and precision of reproduced results will depend on their specific purposes. In a hypothetical circumstance where reproducing all results to within 10% of the original values were acceptable, the clear majority (58%) of articles with data and code in this study would meet this criterion.

The results mentioned above do not include the cases where we wrote new code for those articles where the shared code turned out to be irrelevant to the target result. If these attempts were included in the success rate calculations, the results would improve as shown in Table 10. However, the inclusion of these results as "computational reproducibility" attempts does not fit with our initial definition of computational reproducibility, which posits both data and code be used to recalculate a result. We regarded writing new analysis code from a description of the methods to be a different category of task ("analytic reproducibility"). Conducting analytic reproducibility attempts (based on a sample of the meta-analysis articles which shared data only, for example) in addition to the four attempts in this study was beyond the scope of this study.

Although our canonical computational reproducibility attempts made use of existing code that had been shared to re-run an analysis, we still needed to write bespoke code in order to facilitate the attempt. All attempts required custom code for minor matters like specifying input file locations and re-directing analysis output. Occasionally, custom code was required for more substantial tasks such as processing the shared data files before they could be analysed by the shared code. This frequent need for such additional effort by the researcher conducting the computational reproduction is well recognised in other studies of computational reproducibility. The reproducibility project described in Wood et al. (21) had an expectation that replication code and data received would be "ready-to-run"; they used the term "push button replication" to describe computational reproducibility attempts, which suggests an ideal scenario where an independent researcher can simply "push the button and reproduce the published results" (21, p.2). However, this was rarely attainable in practice, and to get code working, researchers sometimes had to escalate from minor code troubleshooting (e.g., installing required libraries, or changing the version of a software package used) to "[changing] commands in Stata to allow the code to run, updating commands to the current version of the software, and even correcting typos in an attempt to reproduce the original results" (21, p.7). This was recognised separately in Stodden et al. (20), who classified the different levels of effort required when attempting to reproduce results from 22 articles. The classification captured the escalation of effort required from minor difficulties or tweaks (such as installing required software libraries, or adjusting code to work on a different computational system) to major, tedious difficulties (such as needing to write code to re-format data or fill in missing steps) (20, Table 4). We encountered similar difficulties to those described in Stodden et al. (20) and Wood et al. (21), and although we have sought to make a clear distinction between computational reproducibility and analytic reproducibility by contrasting "running existing code" with "writing new code", we acknowledge that in practice this distinction may become blurred in cases of computational reproducibility attempts requiring new code to be written. Further scrutiny of the definition of "computational reproducibility" in the light of the results of this study is included in Section S11 of the Appendix.

Limitations

A limitation of this study is that the observed rate(s) of computational reproducibility were possibly underestimated. By design, this study did not attempt to contact article authors seeking access to data and code. Although other studies (20; 24) report mixed success with receiving data and code from authors, it is still the case that assistance from original authors could have lifted the rate of obtained data and code for articles, and in turn potentially the overall reproducibility rate(s).

We did not record the time spent on each reproduction attempt, despite some attempts taking much longer than others. Given that researcher time, effort, and opportunity cost are important considerations, this is perhaps a lost chance to have provided additional information about the activity of reproduction.

Although the strategy of selecting only a single target result to reproduce per article made it feasible to attempt to reproduce results from more articles, it did not provide a measurement of the reproducibility of entire articles. Thus, on the basis of these investigations we cannot claim that any of these articles are entirely "reproducible". Despite this limitation this strategy can be considered in the context of a "triage" approach: a hypothetical article identified as failing such a relatively simple reproducibility check likely has issues with the data, code, or the reliability of published results that must be addressed before further time/effort is expended, or before any results are taken to be accurate for particular purposes.

5 CONCLUSION

This study, like Wood et al. (21) and Crüwell et al. (23), is an example of an audit of the computational reproducibility of the literature that ought to be a regular, ongoing part of the broader project of meta-research to bolster the credibility of results within disciplines. Such checks are an effective gauge the efficacy of data- and code-sharing practices and policies, as well as providing assurance on the accuracy of published results. Our methods for conducting the reproduction attempts can be used as a template for computational reproducibility projects, and which can be expanded upon as required. Our results can be a benchmark and point of comparison for the success rates of other computational reproducibility attempts, at other times and for different types of studies.

We reported the success rate of computational reproducibility of one type of study (meta-analysis) published during 2015–17. The low rate of code sharing among articles published during this period was the principal limitation on the number of possible reproduction attempts. From this, improvement in computational reproducibility would then depend on researchers sharing their code alongside their data when publishing. Journal and funder policies mandating code sharing are clearly one key element of achieving higher rates of code sharing; another would be to equip researchers with the knowledge they need to produce (re-)usable code that can be shared with confidence. On this point, there are a number of resources specifically for ecology and evolutionary biology researchers. The guide to reproducible code published by the British Ecological Society (64) provides a overview of working reproducibly at all stages of a research project, from initial organisation and structuring of code files to the archiving of a completed project. For the actual writing of code, the introduction to writing “clean code” by Filazzola and Lortie (65) emphasises the formatting and organisation of code to facilitate clear communication of code’s purpose and function. There is also an effort to alert researchers to tools that can make reproducible work easier: Braga et al. (66) have compiled a list of 12 ways researchers in ecology and evolutionary biology can use online code repository GitHub, from the straightforward archiving of code and data files to using it to coordinating code development across a team of collaborators.

While widespread availability of code would undoubtedly assist audit studies investigating computational reproducibility post-publication, the success rate of such studies would be further improved (perhaps substantially so) if code was reviewed before publication, perhaps as part of peer review as discussed by Fernández-Juricic (67). Ivimey-Cook et al. (68) provide a comprehensive primer of code review at all stages of a research project, outlining a workflow for conducting effective reviews. Implementing code review into the research process (whether as part of formal peer review or not) would require a change in current research practices and the allocation of resources; the costs of this would need to be compared against the advantages of enhancing the reproducibility of reported results.

Given the initiatives to improve researchers’ code, in concert with journal policies mandating data and code sharing, and the growing awareness of a role for code review, there is reason to be optimistic that future studies of computational reproducibility in ecology and evolutionary biology will not only find higher rates of success, but will be easier for meta-researchers to conduct.

DATA AVAILABILITY STATEMENT

The data and code files to reproduce all results reported in this article are available on Zenodo at `https://doi.org/10.5281/zenodo.8114702`.

The reproducibility reports created made use of data and code files shared alongside the published meta-analysis articles surveyed in this study; these data/code files are not included in the above Zenodo repository.

AUTHOR CONTRIBUTIONS

SK: Conceptualization (lead); Methodology (lead); Investigation (lead); Formal Analysis (lead); Writing - Original Draft (lead); Writing - Review & Editing (equal).

DW: Supervision (supporting); Writing - Review & Editing (equal).

ETS: Supervision (supporting); Writing - Review & Editing (equal).

FF: Supervision (lead); Writing - Review & Editing (equal).

AUTHOR COMPETING INTERESTS

The authors disclose that they have no competing interests.

REFERENCES

[1] Fidler F, Chee YE, Wintle BC, Burgman MA, McCarthy MA, Gordon A. Metaresearch for Evaluating Reproducibility in Ecology and Evolution. BioScience. 2017;67(3):282–289. doi:10.1093/biosci/biw159.

[2] Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015;349(6251):aac4716. doi:10.1126/science.aac4716.

[3] Nakagawa S, Parker TH. Replicating research in ecology and evolution: feasibility, incentives, and the cost-benefit conundrum. BMC Biology. 2015;13(1):88. doi:10.1186/s12915-015-0196-3.

[4] Schnitzer SA, Carson WP. Would Ecology Fail the Repeatability Test? BioScience. 2016;66(2):98–99. doi:10.1093/biosci/biv176.

[5] Ihle M, Winney IS, Krystalli A, Croucher M. Striving for transparent and credible research: practical guidelines for behavioral ecologists. Behavioral Ecology. 2017;28(2):348–354. doi:10.1093/beheco/arx003.

[6] Shavit A, Ellison AM, editors. Stepping in the Same River Twice: Replication in Biological Research. New Haven: Yale University Press; 2017.

[7] Kelly CD. Rate and success of study replication in ecology and evolution. PeerJ. 2019;7:e7654. doi:10.7717/peerj.7654.

[8] Fraser H, Barnett A, Parker TH, Fidler F. The role of replication studies in ecology. Ecology and Evolution. 2020;10(12):5197–5207. doi:10.1002/ece3.6330.

[9] Yang Y, Sánchez-Tójar A, O'Dea RE, Noble DWA, Koricheva J, Jennions MD, et al. Publication bias impacts on effect size, statistical power, and magnitude (Type M) and sign (Type S) errors in ecology and evolutionary biology. BMC Biology. 2023;21(1):71. doi:10.1186/s12915-022-01485-y.

[10] Kimmel K, Avolio ML, Ferraro PJ. Empirical evidence of widespread exaggeration bias and selective reporting in ecology. Nature Ecology & Evolution. 2023;7(9):1525–1536. doi:10.1038/s41559-023-02144-3.

[11] Fraser H, Parker T, Nakagawa S, Barnett A, Fidler F. Questionable research practices in ecology and evolution. PLOS ONE. 2018;13(7):e0200303. doi:10.1371/journal.pone.0200303.

[12] Gould E, Fraser HS, Parker TH, Nakagawa S, Griffith SC, Vesk PA, et al. Same data, different analysts: variation in effect sizes due to analytical decisions in ecology and evolutionary biology. EcoEvoRxiv; 2023. Available from: `https://ecoevorxiv.org/repository/view/6000/`.

[13] Parker TH, Forstmeier W, Koricheva J, Fidler F, Hadfield JD, Chee YE, et al. Transparency in Ecology and Evolution: Real Problems, Real Solutions. Trends in Ecology & Evolution. 2016;31(9):711–719. doi:10.1016/j.tree.2016.07.002.

[14] O'Dea RE, Parker TH, Chee YE, Culina A, Drobniak SM, Duncan DH, et al. Towards open, reliable, and transparent ecology and evolutionary biology. BMC Biology. 2021;19(1):68. doi:10.1186/s12915-021-01006-3.

[15] National Academies of Sciences, Engineering, and Medicine. Reproducibility and Replicability in Science. Washington, DC: The National Academies Press; 2019.

[16] Hardwicke TE, Mathur MB, MacDonald K, Nilsonne G, Banks GC, Kidwell MC, et al. Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal Cognition. Royal Society Open Science. 2018;5(8):180448. doi:10.1098/rsos.180448.

[17] Hardwicke TE, Bohn M, MacDonald K, Hembacher E, Nuijten MB, Peloquin BN, et al. Analytic reproducibility in articles receiving open data badges at the journal Psychological Science: an observational study. Royal Society Open Science. 2021;8(1):201494. doi:10.1098/rsos.201494.

[18] Minocher R, Atmaca S, Bavero C, McElreath R, Beheim B. Estimating the reproducibility of social learning research published between 1955 and 2018. Royal Society Open Science. 2021;8(9):210450. doi:10.1098/rsos.210450.

[19] Perkel JM. Challenge to scientists: does your ten-year-old code still run? Nature. 2020;584(7822):656–658. doi:10.1038/d41586-020-02462-7.

[20] Stodden V, Seiler J, Ma Z. An empirical analysis of journal policy effectiveness for computational reproducibility. Proceedings of the National Academy of Sciences. 2018;115(11):2584–2589. doi:10.1073/pnas.1708290115.

[21] Wood BDK, Müller R, Brown AN. Push button replication: Is impact evaluation evidence for international development verifiable? PLOS ONE. 2018;13(12):e0209416. doi:10.1371/journal.pone.0209416.

[22] Obels P, Lakens D, Coles NA, Gottfried J, Green SA. Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology. Advances in Methods and Practices in Psychological Science. 2020;3(2):229–237. doi:10.1177/2515245920918872.

[23] Crüwell S, Apthorp D, Baker BJ, Colling L, Elson M, Geiger SJ, et al. What's in a Badge? A Computational Reproducibility Investigation of the Open Data Badge Policy in One Issue of Psychological Science. Psychological Science. 2023;34(4):512–522. doi:10.1177/09567976221140828.

[24] ArchMiller AA, Johnson AD, Nolan J, Edwards M, Elliott LH, Ferguson JM, et al. Computational Reproducibility in The Wildlife Society's Flagship Journals. The Journal of Wildlife Management. 2020;84(5):1012–1017. doi:https://doi.org/10.1002/jwmg.21855.

[25] Moore AJ, Mcpeek MA, Rausher MD, Rieseberg L, Whitlock MC. The need for archiving data in evolutionary biology. Journal of Evolutionary Biology. 2010;23(4):659–660. doi:10.1111/j.1420-9101.2010.01937.x.

[26] Whitlock M, McPeek M, Rausher M, Rieseberg L, Moore A. Data Archiving. The American Naturalist. 2010;175(2):145–146. doi:10.1086/650340.

[27] Simmons LW. Mandatory data archiving in Behavioral Ecology. Behavioral Ecology. 2016;27(1):1–1. doi:10.1093/beheco/arw001.

[28] Culina A, Baglioni M, Crowther TW, Visser ME, Woutersen-Windhouwer S, Manghi P. Navigating the unfolding open data landscape in ecology and evolution. Nature Ecology & Evolution. 2018;2(3):420–426. doi:10.1038/s41559-017-0458-2.

[29] Whitlock MC. Data archiving in ecology and evolution: best practices. Trends in Ecology & Evolution. 2011;26(2):61–65. doi:10.1016/j.tree.2010.11.006.

[30] Caetano DS, Aisenberg A. Forgotten treasures: the fate of data in animal behaviour studies. Animal Behaviour. 2014;98:1–5. doi:10.1016/j.anbehav.2014.09.025.

[31] Miller SE, Jernigan CM, Legan AW, Miller CH, Tumulty JP, Walton A, et al. Animal behavior missing from data archives. Trends in Ecology & Evolution. 2021;36(11):960–963. doi:10.1016/j.tree.2021.07.008.

[32] Roche DG, Kruuk LEB, Lanfear R, Binning SA. Public Data Archiving in Ecology and Evolution: How Well Are We Doing? PLOS Biology. 2015;13(11):e1002295. doi:10.1371/journal.pbio.1002295.

[33] Mislan KAS, Heer JM, White EP. Elevating The Status of Code in Ecology. Trends in Ecology & Evolution. 2016;31(1):4–7. doi:10.1016/j.tree.2015.11.006.

[34] Culina A, Berg Ivd, Evans S, Sánchez-Tójar A. Low availability of code in ecology: A call for urgent action. PLOS Biology. 2020;18(7):e3000763. doi:10.1371/journal.pbio.3000763.

[35] Nakagawa S, Santos ESA. Methodological issues and advances in biological meta-analysis. Evolutionary Ecology. 2012;26(5):1253–1274. doi:10.1007/s10682-012-9555-5.

[36] Dryad. Dryad; 2021. Available from: `https://datadryad.org/stash/`.

[37] Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010;36(1):1–48. doi:10.18637/jss.v036.i03.

[38] Xu X, Medvigy D, Joseph Wright S, Kitajima K, Wu J, Albert LP, et al. Variations of leaf longevity in tropical moist forests predicted by a trait-driven carbon optimality model. Ecology Letters. 2017;20(9):1097–1106. doi:10.1111/ele.12804.

[39] Winternitz J, Abbate JL, Huchard E, Havlíček J, Garamszegi LZ. Patterns of MHC-dependent mate selection in humans and nonhuman primates: a meta-analysis. Molecular Ecology. 2017;26(2):668–688. doi:10.1111/mec.13920.

[40] Grueber CE, Gray LJ, Morris KM, Simpson SJ, Senior AM. Intergenerational effects of nutrition on immunity: a systematic review and meta-analysis. Biological Reviews. 2018;93(2):1108–1124. doi:10.1111/brv.12387.

[41] Noble DWA, Stenhouse V, Schwanz LE. Developmental temperatures and phenotypic plasticity in reptiles: a systematic review and meta-analysis. Biological Reviews. 2018;93(1):72–97. doi:10.1111/brv.12333.

[42] Risely A, Klaassen M, Hoye BJ. Migratory animals feel the cost of getting sick: A meta-analysis across species. Journal of Animal Ecology. 2017;87(1):301–314. doi:10.1111/1365-2656.12766.

[43] Ronget V, Gaillard J, Coulson T, Garratt M, Gueyffier F, Lega J, et al. Causes and consequences of variation in offspring body mass: meta-analyses in birds and mammals. Biological Reviews. 2017;93(1):1–27. doi:10.1111/brv.12329.

[44] Sievers M, Hale R, Parris KM, Swearer SE. Impacts of human-induced environmental change in wetlands on aquatic animals. Biological Reviews. 2017;93(1):529–554. doi:10.1111/brv.12358.

[45] Harts AMF, Booksmythe I, Jennions MD. Mate guarding and frequent copulation in birds: A meta-analysis of their relationship to paternity and male phenotype. Evolution. 2016;70(12):2789–2808. doi:10.1111/evo.13081.

[46] Jaffé R, Pope N, Acosta AL, Alves DA, Arias MC, De la Rúa P, et al. Beekeeping practices and geographic distance, not land use, drive gene flow across tropical bees. Molecular Ecology. 2016;25(21):5345–5358. doi:10.1111/mec.13852.

[47] Lemoine NP, Hoffman A, Felton AJ, Baur L, Chaves F, Gray J, et al. Underappreciated problems of low replication in ecological field studies. Ecology. 2016;97(10):2554–2561. doi:10.1002/ecy.1506.

[48] Gibert A, Gray EF, Westoby M, Wright IJ, Falster DS. On the link between functional traits and growth rate: meta-analysis shows effects change with plant size, as predicted. Journal of Ecology. 2016;104(5):1488–1503. doi:10.1111/1365-2745.12594.

[49] Anderson JT. Plant fitness in a rapidly changing world. New Phytologist. 2016;210(1):81–87. doi:10.1111/nph.13693.

[50] Moore FR, Shuker DM, Dougherty L. Stress and sexual signaling: a systematic review and meta-analysis. Behavioral Ecology. 2016;27(2):363–371. doi:10.1093/beheco/arv195.

[51] Holman L. Bet hedging via multiple mating: A meta-analysis. Evolution. 2016;70(1):62–71. doi:10.1111/evo.12822.

[52] Strader ME, Aglyamova GV, Matz MV. Red fluorescence in coral larvae is associated with a diapause-like state. Molecular Ecology. 2016;25(2):559–569. doi:10.1111/mec.13488.

[53] Senior AM, Nakagawa S, Lihoreau M, Simpson SJ, Raubenheimer D. An Overlooked Consequence of Dietary Mixing: A Varied Diet Reduces Interindividual Variance in Fitness. The American Naturalist. 2015;186(5):649–659. doi:10.1086/683182.

[54] Voje KL. Scaling of Morphological Characters across Trait Type, Sex, and Environment: A Meta-analysis of Static Allometries. The American Naturalist. 2015;187(1):89–98. doi:10.1086/684159.

[55] Paz-Vinas I, Loot G, Stevens VM, Blanchet S. Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. Molecular Ecology. 2015;24(18):4586–4604. doi:10.1111/mec.13345.

[56] Mehrabi Z, Tuck SL. Relatedness is a poor predictor of negative plant–soil feedbacks. New Phytologist. 2015;205(3):1071–1075. doi:10.1111/nph.13238.

[57] Yuan ZY, Chen HYH. Negative effects of fertilization on plant nutrient resorption. Ecology. 2015;96(2):373–380. doi:10.1890/14-0140.1.

[58] Colautti RI, Lau JA. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. Molecular Ecology. 2015;24(9):1999–2017. doi:10.1111/mec.13162.

[59] Gamfeldt L, Lefcheck JS, Byrnes JEK, Cardinale BJ, Duffy JE, Griffin JN. Marine biodiversity and ecosystem functioning: what's known and what's next? Oikos. 2015;124(3):252–265. doi:10.1111/oik.01549.

[60] Xu X, Medvigy D, Powers JS, Becknell JM, Guan K. Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. New Phytologist. 2016;212(1):80–95. doi:10.1111/nph.14009.

[61] Turney S, Buddle CM. Pyramids of species richness: the determinants and distribution of species diversity across trophic levels. Oikos. 2016;125(9):1224–1232. doi:10.1111/oik.03404.

[62] Crouzeilles R, Curran M. Which landscape size best predicts the influence of forest cover on restoration success? A global meta-analysis on the scale of effect. Journal of Applied Ecology. 2016;53(2):440–448. doi:10.1111/1365-2664.12590.

[63] Valls A, Coll M, Christensen V. Keystone species: toward an operational concept for marine biodiversity conservation. Ecological Monographs. 2015;85(1):29–47. doi:10.1890/14-0306.1.

[64] Cooper N, Hsing PY, editors. Reproducible Code. BES Guides to Better Science. London, UK: British Ecological Society; 2017.

[65] Filazzola A, Lortie C. A call for clean code to effectively communicate science. Methods in Ecology and Evolution. 2022;13(10):2119–2128. doi:10.1111/2041-210X.13961.

[66] Braga PHP, Hébert K, Hudgins EJ, Scott ER, Edwards BPM, Sánchez Reyes LL, et al. Not just for programmers: How GitHub can accelerate collaborative and reproducible research in ecology and evolution. Methods in Ecology and Evolution. 2023;14(6):1364–1380. doi:10.1111/2041-210X.14108.

[67] Fernández-Juricic E. Why sharing data and code during peer review can enhance behavioral ecology research. Behavioral Ecology and Sociobiology. 2021;75(7):103. doi:10.1007/s00265-021-03036-x.

[68] Ivimey-Cook ER, Pick JL, Bairos-Novak KR, Culina A, Gould E, Grainger M, et al. Implementing code review in the scientific workflow: Insights from ecology and evolutionary biology. Journal of Evolutionary Biology. 2023;36(10):1347–1356. doi:10.1111/jeb.14230.

[69] Lau J, Rothstein HR, Stewart GB. History and Progress of Meta-analysis. In: Koricheva J, Gurevitch J, Mengersen K, editors. Handbook of Meta-analysis in Ecology and Evolution. Princeton University Press; 2013. p. 407–419.

[70] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Chichester, U.K.: Wiley; 2009.

[71] Green S, Higgins JP, Alderson P, Clarke M, Mulrow CD, Oxman AD. Introduction. In: Fellow JPHSSV, Director SGF, editors. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons, Ltd; 2008. p. 1–9.

[72] Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLOS Medicine. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097.

[73] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71.

[74] O'Dea RE, Lagisz M, Jennions MD, Koricheva J, Noble DWA, Parker TH, et al. Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. Biological Reviews. 2021;96(5):1695–1722. doi:10.1111/brv.12721.

[75] Arnqvist G, Wooster D. Meta-analysis: synthesizing research findings in ecology and evolution. Trends in Ecology & Evolution. 1995;10(6):236–240. doi:10.1016/S0169-5347(00)89073-4.

[76] Gurevitch J, Hedges LV. Statistical Issues in Ecological Meta-Analyses. Ecology. 1999;80(4):1142–1149. doi:10.1890/0012-9658(1999)080[1142:SIIEMA]2.0.CO;2.

[77] Gurevitch J, Hedges LV. Meta-analysis: Combining the results of independent experiments. In: Scheiner SM, Gurevitch J, editors. Design and analysis of ecological experiments. 2nd ed. New York, NY: Oxford University Press; 2001. p. 347–369.

[78] Koricheva J, Gurevitch J, Mengersen K, editors. Handbook of Meta-analysis in Ecology and Evolution. Princeton University Press; 2013.

[79] Gates S. Review of methodology of quantitative reviews using meta-analysis in ecology. Journal of Animal Ecology. 2002;71(4):547–557. doi:10.1046/j.1365-2656.2002.00634.x.

[80] Vetter D, Rücker G, Storch I. Meta-analysis: A need for well-defined usage in ecology and conservation biology. Ecosphere. 2013;4(6):1–24. doi:10.1890/ES13-00062.1.

[81] Koricheva J, Gurevitch J. Uses and misuses of meta-analysis in plant ecology. Journal of Ecology. 2014;102(4):828–844. doi:10.1111/1365-2745.12224.

[82] ArchMiller AA, Bauer EF, Koch RE, Wijayawardena BK, Anil A, Kottwitz JJ, et al. Formalizing the definition of meta-analysis in Molecular Ecology. Molecular Ecology. 2015;24(16):4042–4051. doi:10.1111/mec.13264.

[83] Nakagawa S, Noble DWA, Senior AM, Lagisz M. Meta-evaluation of meta-analysis: ten appraisal questions for biologists. BMC Biology. 2017;15(1):18. doi:10.1186/s12915-017-0357-7.

[84] Bowles TM, Jackson LE, Loeher M, Cavagnaro TR. Ecological intensification and arbuscular mycorrhizas: a meta-analysis of tillage and cover crop effects. Journal of Applied Ecology. 2017;54(6):1785–1793. doi:10.1111/1365-2664.12815.

[85] Mori AS, Tatsumi S, Gustafsson L. Landscape properties affect biodiversity response to retention approaches in forestry. Journal of Applied Ecology. 2017;54(6):1627–1637. doi:10.1111/1365-2664.12888.

[86] Charlebois JA, Sargent RD. No consistent pollinator-mediated impacts of alien plants on natives. Ecology Letters. 2017;20(11):1479–1490. doi:10.1111/ele.12831.

[87] Martin-StPaul N, Delzon S, Cochard H. Plant resistance to drought depends on timely stomatal closure. Ecology Letters. 2017;20(11):1437–1447. doi:10.1111/ele.12851.

[88] Romano A, Saino N, Møller AP. Viability and expression of sexual ornaments in the barn swallow Hirundo rustica: a meta-analysis. Journal of Evolutionary Biology. 2017;30(10):1929–1935. doi:10.1111/jeb.13151.

[89] Davidson KE, Fowler MS, Skov MW, Doerr SH, Beaumont N, Griffin JN. Livestock grazing alters multiple ecosystem properties and services in salt marshes: a meta-analysis. Journal of Applied Ecology. 2017;54(5):1395–1405. doi:10.1111/1365-2664.12892.

[90] Grant JWA, Weir LK, Steingrímsson SÓ. Territory size decreases minimally with increasing food abundance in stream salmonids: Implications for population regulation. Journal of Animal Ecology. 2017;86(6):1308–1316. doi:10.1111/1365-2656.12737.

[91] Horswill C, O'Brien SH, Robinson RA. Density dependence and marine bird populations: are wind farm assessments precautionary? Journal of Applied Ecology. 2017;54(5):1406–1414. doi:10.1111/1365-2664.12841.

[92] Auer SK, Killen SS, Rezende EL. Resting vs. active: a meta-analysis of the intra- and inter-specific associations between minimum, sustained, and maximum metabolic rates in vertebrates. Functional Ecology. 2017;31(9):1728–1738. doi:10.1111/1365-2435.12879.

[93] Buchanan AL, Hermann SL, Lund M, Szendrei Z. A meta-analysis of non-consumptive predator effects in arthropods: the influence of organismal and environmental characteristics. Oikos. 2017;126(9):1233–1240. doi:10.1111/oik.04384.

[94] Soria M, Leigh C, Datry T, Bini LM, Bonada N. Biodiversity in perennial and intermittent rivers: a meta-analysis. Oikos. 2017;126(8):1078–1089. doi:10.1111/oik.04118.

[95] Delavaux CS, Smith-Ramesh LM, Kuebbing SE. Beyond nutrients: a meta-analysis of the diverse effects of arbuscular mycorrhizal fungi on plants and soils. Ecology. 2017;98(8):2111–2119. doi:10.1002/ecy.1892.

[96] LaManna JA, Martin TE. Logging impacts on avian species richness and composition differ across latitudes and foraging and breeding habitat preferences. Biological Reviews. 2017;92(3):1657–1674. doi:10.1111/brv.12300.

[97] Leal LC, Peixoto PEC. Decreasing water availability across the globe improves the effectiveness of protective ant–plant mutualisms: a meta-analysis. Biological Reviews. 2017;92(3):1785–1794. doi:10.1111/brv.12307.

[98] Romano A, Costanzo A, Rubolini D, Saino N, Møller AP. Geographical and seasonal variation in the intensity of sexual selection in the barn swallow Hirundo rustica: a meta-analysis. Biological Reviews. 2017;92(3):1582–1600. doi:10.1111/brv.12297.

[99] Anthelme F, Meneses RI, Valero NNH, Pozo P, Dangles O. Fine nurse variations explain discrepancies in the stress-interaction relationship in alpine regions. Oikos. 2017;126(8):1173–1183. doi:10.1111/oik.04248.

[100] Hindrikson M, Remm J, Pilot M, Godinho R, Stronen AV, Baltrūnaité L, et al. Wolf population genetics in Europe: a systematic review, meta-analysis and suggestions for conservation and management. Biological Reviews. 2017;92(3):1601–1629. doi:10.1111/brv.12298.

[101] Knapp JL, Bartlett LJ, Osborne JL. Re-evaluating strategies for pollinator-dependent crops: How useful is parthenocarpy? Journal of Applied Ecology. 2017;54(4):1171–1179. doi:10.1111/1365-2664.12813.

[102] Gázquez A, Beemster GTS. What determines organ size differences between species? A meta-analysis of the cellular basis. New Phytologist. 2017;215(1):299–308. doi:10.1111/nph.14573.

[103] Hendershot JN, Read QD, Henning JA, Sanders NJ, Classen AT. Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. Ecology. 2017;98(7):1757–1763. doi:10.1002/ecy.1829.

[104] Farji-Brener AG, Werenkraut V. The effects of ant nests on soil fertility and plant performance: a meta-analysis. Journal of Animal Ecology. 2017;86(4):866–877. doi:10.1111/1365-2656.12672.

[105] Hitchcock DJ, Varpe Ø, Andersen T, Borgå K. Effects of reproductive strategies on pollutant concentrations in pinnipeds: a meta-analysis. Oikos. 2017;126(6):772–781. doi:10.1111/oik.03955.

[106] Miller SE, Barrueto M, Schluter D. A comparative analysis of experimental selection on the stickleback pelvis. Journal of Evolutionary Biology. 2017;30(6):1165–1176. doi:10.1111/jeb.13085.

[107] Rowiński PK, Rogell B. Environmental stress correlates with increases in both genetic and residual variances: A meta-analysis of animal studies. Evolution. 2017;71(5):1339–1351. doi:10.1111/evo.13201.

[108] Yue K, Fornara DA, Yang W, Peng Y, Peng C, Liu Z, et al. Influence of multiple global change drivers on terrestrial carbon storage: additive effects are common. Ecology Letters. 2017;20(5):663–672. doi:10.1111/ele.12767.

[109] Wood KA, Ponting J, D'Costa N, Newth JL, Rose PE, Glazov P, et al. Understanding intrinsic and extrinsic drivers of aggressive behaviour in waterbird assemblages: a meta-analysis. Animal Behaviour. 2017;126:209–216. doi:10.1016/j.anbehav.2017.02.008.

[110] Greenwood S, Ruiz-Benito P, Martínez-Vilalta J, Lloret F, Kitzberger T, Allen CD, et al. Tree mortality across biomes is promoted by drought intensity, lower wood density and higher specific leaf area. Ecology Letters. 2017;20(4):539–553. doi:10.1111/ele.12748.

[111] Martin PA, Newton AC, Bullock JM. Impacts of invasive plants on carbon pools depend on both species' traits and local climate. Ecology. 2017;98(4):1026–1035. doi:10.1002/ecy.1711.

[112] Holtmann B, Lagisz M, Nakagawa S. Metabolic rates, and not hormone levels, are a likely mediator of between-individual differences in behaviour: a meta-analysis. Functional Ecology. 2017;31(3):685–696. doi:10.1111/1365-2435.12779.

[113] Catano CP, Dickson TL, Myers JA. Dispersal and neutral sampling mediate contingent effects of disturbance on plant beta-diversity: a meta-analysis. Ecology Letters. 2017;20(3):347–356. doi:10.1111/ele.12733.

[114] Poulin R, Pérez-Ponce de León G. Global analysis reveals that cryptic diversity is linked with habitat but not mode of life. Journal of Evolutionary Biology. 2017;30(3):641–649. doi:10.1111/jeb.13034.

[115] Usui T, Butchart SHM, Phillimore AB. Temporal shifts and temperature sensitivity of avian spring migratory phenology: a phylogenetic meta-analysis. Journal of Animal Ecology. 2017;86(2):250–261. doi:10.1111/1365-2656.12612.

[116] Hrycik AR, Almeida LZ, Höök TO. Sub-lethal effects on fish provide insight into a biologically-relevant threshold of hypoxia. Oikos. 2017;126(3):307–317. doi:10.1111/oik.03678.

[117] Rossetti MR, Tscharntke T, Aguilar R, Batáry P. Responses of insect herbivores and herbivory to habitat fragmentation: a hierarchical meta-analysis. Ecology Letters. 2017;20(2):264–272. doi:10.1111/ele.12723.

[118] Sorenson GH, Dey CJ, Madliger CL, Love OP. Effectiveness of baseline corticosterone as a monitoring tool for fitness: a meta-analysis in seabirds. Oecologia. 2017;183(2):353–365. doi:10.1007/s00442-016-3774-3.

[119] Foo YZ, Nakagawa S, Rhodes G, Simmons LW. The effects of sex hormones on immune function: a meta-analysis. Biological Reviews. 2017;92(1):551–571. doi:10.1111/brv.12243.

[120] Hossie T, Landolt K, Murray DL. Determinants and co-expression of anti-predator responses in amphibian tadpoles: a meta-analysis. Oikos. 2017;126(2):n/a–n/a. doi:10.1111/oik.03305.

[121] Booksmythe I, Mautz B, Davis J, Nakagawa S, Jennions MD. Facultative adjustment of the offspring sex ratio and male attractiveness: a systematic review and meta-analysis. Biological Reviews. 2017;92(1):108–134. doi:10.1111/brv.12220.

[122] Ameye M, Allmann S, Verwaeren J, Smagghe G, Haesaert G, Schuurink RC, et al. Green leaf volatile production by plants: a meta-analysis. New Phytologist. 2017; p. n/a–n/a. doi:10.1111/nph.14671.

[123] Vidal MC, Murphy SM. Bottom-up vs. top-down effects on terrestrial insect herbivores: a meta-analysis. Ecology Letters. 2018;21(1):138–150. doi:10.1111/ele.12874.

[124] Deng Q, McMahon DE, Xiang Y, Yu CL, Jackson RB, Hui D. A global meta-analysis of soil phosphorus dynamics after afforestation. New Phytologist. 2017;213(1):181–192. doi:10.1111/nph.14119.

[125] Lee MR, Bernhardt ES, van Bodegom PM, Cornelissen JHC, Kattge J, Laughlin DC, et al. Invasive species' leaf traits and dissimilarity from natives shape their impact on nitrogen cycling: a meta-analysis. New Phytologist. 2017;213(1):128–139. doi:10.1111/nph.14115.

[126] Zvereva EL, Zverev V, Kruglova OY, Kozlov MV. Strategies of chemical anti-predator defences in leaf beetles: is sequestration of plant toxins less costly than de novo synthesis? Oecologia. 2017;183(1):93–106. doi:10.1007/s00442-016-3743-x.

[127] Xiao Z, Wang X, Koricheva J, Kergunteuil A, Le Bayon RC, Liu M, et al. Earthworms affect plant growth and resistance against herbivores: A meta-analysis. Functional Ecology. 2018;32(1):150–160. doi:10.1111/1365-2435.12969.

[128] Camenzind T, Hättenschwiler S, Treseder KK, Lehmann A, Rillig MC. Nutrient limitation of soil microbial processes in tropical forests. Ecological Monographs. 2018;88(1):4–21. doi:10.1002/ecm.1279.

[129] Yang H, Zhang Q, Koide RT, Hoeksema JD, Tang J, Bian X, et al. Taxonomic resolution is a determinant of biodiversity effects in arbuscular mycorrhizal fungal communities. Journal of Ecology. 2017;105(1):219–228. doi:10.1111/1365-2745.12655.

[130] Molnar JL, Diogo R, Hutchinson JR, Pierce SE. Reconstructing pectoral appendicular muscle anatomy in fossil fish and tetrapods over the fins-to-limbs transition. Biological Reviews. 2017; p. n/a–n/a. doi:10.1111/brv.12386.

[131] Gruner DS, Bracken MES, Berger SA, Eriksson BK, Gamfeldt L, Matthiessen B, et al. Effects of experimental warming on biodiversity depend on ecosystem type and local species composition. Oikos. 2016;126(1):8–17. doi:10.1111/oik.03688.

[132] Dynarski KA, Houlton BZ. Nutrient limitation of terrestrial free-living nitrogen fixation. New Phytologist. 2018;217(3):1050–1061. doi:10.1111/nph.14905.

[133] Junker RR, Kuppler J, Amo L, Blande JD, Borges RM, van Dam NM, et al. Covariation and phenotypic integration in chemical communication displays: biosynthetic constraints and eco-evolutionary implications. New Phytologist. 2017; p. n/a–n/a. doi:10.1111/nph.14505.

[134] Rohner PT, Teder T, Esperk T, Lüpold S, Blanckenhorn WU. The evolution of male-biased sexual size dimorphism is associated with increased body size plasticity in males. Functional Ecology. 2017;In press:n/a–n/a. doi:10.1111/1365-2435.13004.

[135] Goodell K, Parker IM. Invasion of a dominant floral resource: effects on the floral community and pollination of native plants. Ecology. 2017;98(1):57–69. doi:10.1002/ecy.1639.

[136] Poorter H, Fiorani F, Pieruschka R, Wojciechowski T, Putten WH, Kleyer M, et al. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. New Phytologist. 2016;212(4):838–855. doi:10.1111/nph.14243.

[137] Edge CB, Houlahan JE, Jackson DA, Fortin M. The response of amphibian larvae to environmental change is both consistent and variable. Oikos. 2016;125(12):1700–1711. doi:10.1111/oik.03166.

[138] Gill AL, Finzi AC. Belowground carbon flux links biogeochemical cycles and resource-use efficiency at the global scale. Ecology Letters. 2016;19(12):1419–1428. doi:10.1111/ele.12690.

[139] Buoro M, Olden JD, Cucherousset J. Global Salmonidae introductions reveal stronger ecological effects of changing intraspecific compared to interspecific diversity. Ecology Letters. 2016;19(11):1363–1371. doi:10.1111/ele.12673.

[140] Besson AA, Lagisz M, Senior AM, Hector KL, Nakagawa S. Effect of maternal diet on offspring coping styles in rodents: a systematic review and meta-analysis. Biological Reviews. 2016;91(4):1065–1080. doi:10.1111/brv.12210.

[141] Flick AJ, Acevedo MA, Elderd BD. The negative effects of pathogen-infected prey on predators: a meta-analysis. Oikos. 2016;125(11):1554–1560. doi:10.1111/oik.03458.

[142] Tsen EWJ, Sitzia T, Webber BL. To core, or not to core: the impact of coring on tree health and a best-practice framework for collecting dendrochronological information from living trees. Biological Reviews. 2016;91(4):899–924. doi:10.1111/brv.12200.

[143] Derroire G, Balvanera P, Castellanos-Castro C, Decocq G, Kennard DK, Lebrija-Trejos E, et al. Resilience of tropical dry forests – a meta-analysis of changes in species diversity and composition during secondary succession. Oikos. 2016;125(10):1386–1397. doi:10.1111/oik.03229.

[144] Hidding B, Bakker ES, Hootsmans MJM, Hilt S. Synergy between shading and herbivory triggers macrophyte loss and regime shifts in aquatic systems. Oikos. 2016;125(10):1489–1495. doi:10.1111/oik.03104.

[145] Jessop TS, Lane ML, Teasdale L, Stuart-Fox D, Wilson RS, Careau V, et al. Multiscale Evaluation of Thermal Dependence in the Glucocorticoid Response of Vertebrates. The American Naturalist. 2016;188(3):342–356. doi:10.1086/687588.

[146] Katz DSW. The effects of invertebrate herbivores on plant population growth: a meta-regression analysis. Oecologia. 2016;182(1):43–53. doi:10.1007/s00442-016-3602-9.

[147] Fuzessy LF, Cornelissen TG, Janson C, Silveira FAO. How do primates affect seed germination? A meta-analysis of gut passage effects on neotropical plants. Oikos. 2016;125(8):1069–1080. doi:10.1111/oik.02986.

[148] Cooke J, Leishman MR. Consistent alleviation of abiotic stress with silicon addition: a meta-analysis. Functional Ecology. 2016;30(8):1340–1357. doi:10.1111/1365-2435.12713.

[149] Yoon S, Read Q. Consequences of exotic host use: impacts on Lepidoptera and a test of the ecological trap hypothesis. Oecologia. 2016;181(4):985–996. doi:10.1007/s00442-016-3560-2.

[150] Coetzee BWT, Chown SL. A meta-analysis of human disturbance impacts on Antarctic wildlife. Biological Reviews. 2016;91(3):578–596. doi:10.1111/brv.12184.

[151] Östman Ö, Eklöf J, Eriksson BK, Olsson J, Moksnes PO, Bergström U. Top-down control as important as nutrient enrichment for eutrophication effects in North Atlantic coastal ecosystems. Journal of Applied Ecology. 2016;53(4):1138–1147. doi:10.1111/1365-2664.12654.

[152] He Q, Silliman BR. Consumer control as a common driver of coastal vegetation worldwide. Ecological Monographs. 2016;86(3):278–294. doi:10.1002/ecm.1221.

[153] Delgado-Baquerizo M, Maestre FT, Reich PB, Trivedi P, Osanai Y, Liu YR, et al. Carbon content and climate variability drive global soil bacterial diversity patterns. Ecological Monographs. 2016;86(3):373–390. doi:10.1002/ecm.1216.

[154] Thom D, Seidl R. Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests. Biological Reviews. 2016;91(3):760–781. doi:10.1111/brv.12193.

[155] Liu J, Wu N, Wang H, Sun J, Peng B, Jiang P, et al. Nitrogen addition affects chemical compositions of plant tissues, litter and soil organic matter. Ecology. 2016;97(7):1796–1806. doi:10.1890/15-1683.1.

[156] Liao H, D'Antonio CM, Chen B, Huang Q, Peng S. How much do phenotypic plasticity and local genetic variation contribute to phenotypic divergences along environmental gradients in widespread invasive plants? A meta-analysis. Oikos. 2016;125(7):905–917. doi:10.1111/oik.02372.

[157] Arceo-Gómez G, Ashman TL. Invasion status and phylogenetic relatedness predict cost of heterospecific pollen receipt: implications for native biodiversity decline. Journal of Ecology. 2016;104(4):1003–1008. doi:10.1111/1365-2745.12586.

[158] Oduor AMO, Leimu R, Kleunen Mv. Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. Journal of Ecology. 2016;104(4):957–968. doi:10.1111/1365-2745.12578.

[159] Ficetola GF, Maiorano L. Contrasting effects of temperature and precipitation change on amphibian phenology, abundance and performance. Oecologia. 2016;181(3):683–693. doi:10.1007/s00442-016-3610-9.

[160] Kriengwatana B, Spierings MJ, ten Cate C. Auditory discrimination learning in zebra finches: effects of sex, early life conditions and stimulus characteristics. Animal Behaviour. 2016;116:99–112. doi:10.1016/j.anbehav.2016.03.028.

[161] Samia DSM, Blumstein DT, Stankowich T, Cooper WE. Fifty years of chasing lizards: new insights advance optimal escape theory. Biological Reviews. 2016;91(2):349–366. doi:10.1111/brv.12173.

[162] Smilanich AM, Fincher RM, Dyer LA. Does plant apparency matter? Thirty years of data provide limited support but reveal clear patterns of the effects of plant chemistry on herbivores. New Phytologist. 2016;210(3):1044–1057. doi:10.1111/nph.13875.

[163] Boyle WA, Sandercock BK, Martin K. Patterns and drivers of intraspecific variation in avian life history along elevational gradients: a meta-analysis. Biological Reviews. 2016;91(2):469–482. doi:10.1111/brv.12180.

[164] Blount JD, Vitikainen EIK, Stott I, Cant MA. Oxidative shielding and the cost of reproduction. Biological Reviews. 2016;91(2):483–497. doi:10.1111/brv.12179.

[165] Wesner JS. Contrasting effects of fish predation on benthic versus emerging prey: a meta-analysis. Oecologia. 2016;180(4):1205–1211. doi:10.1007/s00442-015-3539-4.

[166] Moore MP, Riesch R, Martin RA. The predictability and magnitude of life-history divergence to ecological agents of selection: a meta-analysis in livebearing fishes. Ecology Letters. 2016;19(4):435–442. doi:10.1111/ele.12576.

[167] Rowen E, Kaplan I. Eco-evolutionary factors drive induced plant volatiles: a meta-analysis. New Phytologist. 2016;210(1):284–294. doi:10.1111/nph.13804.

[168] Thurman TJ, Barrett RDH. The genetic consequences of selection in natural populations. Molecular Ecology. 2016;25(7):1429–1448. doi:10.1111/mec.13559.

[169] Hébert MP, Beisner BE, Maranger R. A meta-analysis of zooplankton functional traits influencing ecosystem function. Ecology. 2016;97(4):1069–1080. doi:10.1890/15-1084.1.

[170] Lorenz S, Martinez-Fernández V, Alonso C, Mosselman E, García de Jalón D, González del Tánago M, et al. Fuzzy cognitive mapping for predicting hydromorphological responses to multiple pressures in rivers. Journal of Applied Ecology. 2016;53(2):559–566. doi:10.1111/1365-2664.12569.

[171] Dougherty LR, Shuker DM. Variation in pre- and post-copulatory sexual selection on male genital size in two species of lygaeid bug. Behavioral Ecology and Sociobiology. 2016;70(4):625–637. doi:10.1007/s00265-016-2082-6.

[172] van Katwijk MM, Thorhaug A, Marbà N, Orth RJ, Duarte CM, Kendrick GA, et al. Global analysis of seagrass restoration: the importance of large-scale planting. Journal of Applied Ecology. 2016;53(2):567–578. doi:10.1111/1365-2664.12562.

[173] McCary MA, Mores R, Farfan MA, Wise DH. Invasive plants have different effects on trophic structure of green and brown food webs in terrestrial ecosystems: a meta-analysis. Ecology Letters. 2016;19(3):328–335. doi:10.1111/ele.12562.

[174] Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecology Letters. 2016;19(3):336–347. doi:10.1111/ele.12564.

[175] Allen DC, Wesner JS. Synthesis: comparing effects of resource and consumer fluxes into recipient food webs using meta-analysis. Ecology. 2016;97(3):594–604. doi:10.1890/15-1109.1.

[176] Zvereva EL, Kozlov MV. The costs and effectiveness of chemical defenses in herbivorous insects: a meta-analysis. Ecological Monographs. 2016;86(1):107–124. doi:10.1890/15-0911.1.

[177] Barton KE. Tougher and thornier: general patterns in the induction of physical defence traits. Functional Ecology. 2016;30(2):181–187. doi:10.1111/1365-2435.12495.

[178] Street SE, Cross CP, Brown GR. Exaggerated sexual swellings in female nonhuman primates are reliable signals of female fertility and body condition. Animal Behaviour. 2016;112(Supplement C):203–212. doi:10.1016/j.anbehav.2015.11.023.

[179] Roca IT, Desrochers L, Giacomazzo M, Bertolo A, Bolduc P, Deschesnes R, et al. Shifting song frequencies in response to anthropogenic noise: a meta-analysis on birds and anurans. Behavioral Ecology. 2016;27(5):1269–1274. doi:10.1093/beheco/arw060.

[180] Vico G, Manzoni S, Nkurunziza L, Murphy K, Weih M. Trade-offs between seed output and life span – a quantitative comparison of traits between annual and perennial congeneric species. New Phytologist. 2016;209(1):104–114. doi:10.1111/nph.13574.

[181] Daskin JH, Pringle RM. Does primary productivity modulate the indirect effects of large herbivores? A global meta-analysis. Journal of Animal Ecology. 2016;85(4):857–868. doi:10.1111/1365-2656.12522.

[182] Gunton RM, Pöyry J. Scale-specific spatial density dependence in parasitoids: a multi-factor meta-analysis. Functional Ecology. 2016;30(9):1501–1510. doi:10.1111/1365-2435.12627.

[183] German RN, Thompson CE, Benton TG. Relationships among multiple aspects of agriculture's environmental impact and productivity: a meta-analysis to guide sustainable agriculture. Biological Reviews. 2017;92(2):716–738. doi:10.1111/brv.12251.

[184] Shantz AA, Lemoine NP, Burkepile DE. Nutrient loading alters the performance of key nutrient exchange mutualisms. Ecology Letters. 2016;19(1):20–28. doi:10.1111/ele.12538.

[185] Dillingham PW, Moore JE, Fletcher D, Cortés E, Curtis KA, James KC, et al. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. Ecological Applications. 2016;26(1):322–333. doi:10.1890/14-1990.

[186] Wood KA, O'Hare MT, McDonald C, Searle KR, Daunt F, Stillman RA. Herbivore regulation of plant abundance in aquatic ecosystems. Biological Reviews. 2017;92(2):1128–1141. doi:10.1111/brv.12272.

[187] Deng Q, Hui D, Luo Y, Elser J, Wang YP, Loladze I, et al. Down-regulation of tissue N:P ratios in terrestrial plants by elevated CO2. Ecology. 2015;96(12):3354–3362. doi:10.1890/15-0217.1.

[188] Garamszegi LZ, Markó G, Szász E, Zsebők S, Azcárate M, Herczeg G, et al. Among-year variation in the repeatability, within- and between-individual, and phenotypic correlations of behaviors in a natural population. Behavioral Ecology and Sociobiology. 2015;69(12):2005–2017. doi:10.1007/s00265-015-2012-z.

[189] Lüpold S, Simmons LW, Tomkins JL, Fitzpatrick JL. No evidence for a trade-off between sperm length and male premating weaponry. Journal of Evolutionary Biology. 2015;28(12):2187–2195. doi:10.1111/jeb.12742.

[190] Moreno-Mateos D, Meli P, Vara-Rodríguez MI, Aronson J. Ecosystem response to interventions: lessons from restored and created wetland ecosystems. Journal of Applied Ecology. 2015;52(6):1528–1537. doi:10.1111/1365-2664.12518.

[191] Katano I, Doi H, Eriksson BK, Hillebrand H. A cross-system meta-analysis reveals coupled predation effects on prey biomass and diversity. Oikos. 2015;124(11):1427–1435. doi:10.1111/oik.02430.

[192] Graham S, Chapuis E, Meconcelli S, Bonel N, Sartori K, Christophe A, et al. Size-assortative mating in simultaneous hermaphrodites: an experimental test and a meta-analysis. Behavioral Ecology and Sociobiology. 2015;69(11):1867–1878. doi:10.1007/s00265-015-1999-5.

[193] Goessling JM, Kennedy H, Mendonça MT, Wilson AE. A meta-analysis of plasma corticosterone and heterophil : lymphocyte ratios – is there conservation of physiological stress responses over time? Functional Ecology. 2015;29(9):1189–1196. doi:10.1111/1365-2435.12442.

[194] Romero GQ, Gonçalves-Souza T, Vieira C, Koricheva J. Ecosystem engineering effects on species diversity across ecosystems: a meta-analysis. Biological Reviews. 2015;90(3):877–890. doi:10.1111/brv.12138.

[195] Nielsen JM, Popp BN, Winder M. Meta-analysis of amino acid stable nitrogen isotope ratios for estimating trophic position in marine organisms. Oecologia. 2015;178(3):631–642. doi:10.1007/s00442-015-3305-7.

[196] James J, Slater FM, Vaughan IP, Young KA, Cable J. Comparing the ecological impacts of native and invasive crayfish: could native species' translocation do more harm than good? Oecologia. 2015;178(1):309–316. doi:10.1007/s00442-014-3195-0.

[197] Esteban R, Barrutia O, Artetxe U, Fernández-Marín B, Hernández A, García-Plazaola JI. Internal and external factors affecting photosynthetic pigment composition in plants: a meta-analytical approach. New Phytologist. 2015;206(1):268–280. doi:10.1111/nph.13186.

[198] Becker DJ, Streicker DG, Altizer S. Linking anthropogenic resources to wildlife–pathogen dynamics: a review and meta-analysis. Ecology Letters. 2015;18(5):483–495. doi:10.1111/ele.12428.

[199] Starko S, Claman BZ, Martone PT. Biomechanical consequences of branching in flexible wave-swept macroalgae. New Phytologist. 2015;206(1):133–140. doi:10.1111/nph.13182.

[200] Hsu YH, Schroeder J, Winney I, Burke T, Nakagawa S. Are extra-pair males different from cuckolded males? A case study and a meta-analytic examination. Molecular Ecology. 2015;24(7):1558–1571. doi:10.1111/mec.13124.

[201] Mazza CA, Ballaré CL. Photoreceptors UVR8 and phytochrome B cooperate to optimize plant growth and defense in patchy canopies. New Phytologist. 2015;207(1):4–9. doi:10.1111/nph.13332.

[202] Jackson MC. Interactions among multiple invasive animals. Ecology. 2015;96(8):2035–2041. doi:10.1890/15-0171.1.

[203] Arct A, Drobniak SM, Cichoń M. Genetic similarity between mates predicts extrapair paternity—a meta-analysis of bird studies. Behavioral Ecology. 2015;26(4):959–968. doi:10.1093/beheco/arv004.

[204] Boudreau SA, Anderson SC, Worm B. Top-down and bottom-up forces interact at thermal range extremes on American lobster. Journal of Animal Ecology. 2015;84(3):840–850. doi:10.1111/1365-2656.12322.

[205] Yang LY, Machado CA, Dang XD, Peng YQ, Yang DR, Zhang DY, et al. The incidence and pattern of copollinator diversification in dioecious and monoecious figs. Evolution. 2015;69(2):294–304. doi:10.1111/evo.12584.

[206] Siefert A, Violle C, Chalmandrier L, Albert CH, Taudiere A, Fajardo A, et al. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. Ecology Letters. 2015;18(12):1406–1419. doi:10.1111/ele.12508.

[207] Beninde J, Veith M, Hochkirch A. Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. Ecology Letters. 2015;18(6):581–592. doi:10.1111/ele.12427.

[208] Albertson LK, Allen DC. Meta-analysis: abundance, behavior, and hydraulic energy shape biotic effects on sediment transport in streams. Ecology. 2015;96(5):1329–1339. doi:10.1890/13-2138.1.

[209] Frankham R. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Molecular Ecology. 2015;24(11):2610–2618. doi:10.1111/mec.13139.

[210] Slot M, Kitajima K. General patterns of acclimation of leaf respiration to elevated temperatures across biomes and plant types. Oecologia. 2015;177(3):885–900. doi:10.1007/s00442-014-3159-4.

[211] Jahnke M, Olsen JL, Procaccini G. A meta-analysis reveals a positive correlation between genetic diversity metrics and environmental status in the long-lived seagrass Posidonia oceanica. Molecular Ecology. 2015;24(10):2336–2348. doi:10.1111/mec.13174.

[212] Iacarella JC, Dick JTA, Alexander ME, Ricciardi A. Ecological impacts of invasive alien species along temperature gradients: testing the role of environmental matching. Ecological Applications. 2015;25(3):706–716. doi:10.1890/14-0545.1.

[213] Quesnelle PE, Lindsay KE, Fahrig L. Relative effects of landscape-scale wetland amount and landscape matrix quality on wetland vertebrates: a meta-analysis. Ecological Applications. 2015;25(3):812–825. doi:10.1890/14-0362.1.

[214] Dougherty LR, Shuker DM. The effect of experimental design on the measurement of mate choice: a meta-analysis. Behavioral Ecology. 2015;26(2):311–319. doi:10.1093/beheco/aru125.

[215] Sistla SA, Appling AP, Lewandowska AM, Taylor BN, Wolf AA. Stoichiometric flexibility in response to fertilization along gradients of environmental and organismal nutrient richness. Oikos. 2015;124(7):949–959. doi:10.1111/oik.02385.

[216] Jauni M, Gripenberg S, Ramula S. Non-native plant species benefit from disturbance: a meta-analysis. Oikos. 2015;124(2):122–129. doi:10.1111/oik.01416.

[217] Bunn RA, Ramsey PW, Lekberg Y. Do native and invasive plants differ in their interactions with arbuscular mycorrhizal fungi? A meta-analysis. Journal of Ecology. 2015;103(6):1547–1556. doi:10.1111/1365-2745.12456.

[218] Wang X, Taub DR, Jablonski LM. Reproductive allocation in plants as affected by elevated carbon dioxide and other environmental changes: a synthesis using meta-analysis and graphical vector analysis. Oecologia. 2015;177(4):1075–1087. doi:10.1007/s00442-014-3191-4.

[219] Albert A, Auffret AG, Cosyns E, Cousins SAO, D'hondt B, Eichberg C, et al. Seed dispersal by ungulates as an ecological filter: a trait-based meta-analysis. Oikos. 2015;124(9):1109–1120. doi:10.1111/oik.02512.

[220] Mijangos JL, Pacioni C, Spencer PBS, Craig MD. Contribution of genetics to ecological restoration. Molecular Ecology. 2015;24(1):22–37. doi:10.1111/mec.12995.

[221] Tamburello L, Maggi E, Benedetti-Cecchi L, Bellistri G, Rattray AJ, Ravaglioli C, et al. Variation in the impact of non-native seaweeds along gradients of habitat degradation: a meta-analysis and an experimental test. Oikos. 2015;124(9):1121–1131. doi:10.1111/oik.02197.

[222] Ferreira V, Castagneyrol B, Koricheva J, Gulis V, Chauvet E, Graça MAS. A meta-analysis of the effects of nutrient enrichment on litter decomposition in streams. Biological Reviews. 2015;90(3):669–688. doi:10.1111/brv.12125.

[223] Stephens AEA, Westoby M. Effects of insect attack to stems on plant survival, growth, reproduction and photosynthesis. Oikos. 2015;124(3):266–273. doi:10.1111/oik.01809.

[224] Anderegg WRL. Spatial and temporal variation in plant hydraulic traits and their relevance for climate change impacts on vegetation. New Phytologist. 2015;205(3):1008–1014. doi:10.1111/nph.12907.

[225] Fuiman LA, Connelly TL, Lowerre-Barbieri SK, McClelland JW. Egg boons: central components of marine fatty acid food webs. Ecology. 2015;96(2):362–372. doi:10.1890/14-0571.1.

[226] Baeten L, Davies TJ, Verheyen K, Calster HV, Vellend M. Disentangling dispersal from phylogeny in the colonization capacity of forest understorey plants. Journal of Ecology. 2015;103(1):175–183. doi:10.1111/1365-2745.12333.

[227] Taranu ZE, Gregory-Eaves I, Leavitt PR, Bunting L, Buchaca T, Catalan J, et al. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. Ecology Letters. 2015;18(4):375–384. doi:10.1111/ele.12420.

[228] Pintor LM, Byers JE. Do native predators benefit from non-native prey? Ecology Letters. 2015;18(11):1174–1180. doi:10.1111/ele.12496.

[229] Harper KA, Macdonald SE, Mayerhofer MS, Biswas SR, Esseen PA, Hylander K, et al. Edge influence on vegetation at natural and anthropogenic edges of boreal forests in Canada and Fennoscandia. Journal of Ecology. 2015;103(3):550–562. doi:10.1111/1365-2745.12398.

[230] Culina A, Radersma R, Sheldon BC. Trading up: the fitness consequences of divorce in monogamous birds. Biological Reviews. 2015;90(4):1015–1034. doi:10.1111/brv.12143.

[231] Bracken MES, Hillebrand H, Borer ET, Seabloom EW, Cebrian J, Cleland EE, et al. Signatures of nutrient limitation and co-limitation: responses of autotroph internal nutrient concentrations to nitrogen and phosphorus additions. Oikos. 2015;124(2):113–121. doi:10.1111/oik.01215.

[232] Périquet S, Fritz H, Revilla E. The Lion King and the Hyaena Queen: large carnivore interactions and coexistence. Biological Reviews. 2015;90(4):1197–1214. doi:10.1111/brv.12152.

[233] Elliott KH, Hare JF, Vaillant ML, Gaston AJ, Ropert-Coudert Y, Anderson WG. Ageing gracefully: physiology but not behaviour declines with age in a diving seabird. Functional Ecology. 2015;29(2):219–228. doi:10.1111/1365-2435.12316.

[234] Lafuente A, Pérez-Palacios P, Doukkali B, Molina-Sánchez MD, Jiménez-Zurdo JI, Caviedes MA, et al. Unraveling the effect of arsenic on the model Medicago–Ensifer interaction: a transcriptomic meta-analysis. New Phytologist. 2015;205(1):255–272. doi:10.1111/nph.13009.

[235] Dryad. Joint Data Archiving Policy (JDAP); 2020. Available from: `https://datadryad.org/docs/JointDataArchivingPolicy.pdf`.

[236] Rausher MD, McPeek MA, Moore AJ, Rieseberg L, Whitlock MC. Data Archiving. Evolution. 2010;64(3):603–604. doi:10.1111/j.1558-5646.2009.00940.x.

[237] Rieseberg L, Vines T, Kane N. Editorial and retrospective 2010. Molecular Ecology. 2010;19(1):1–22. doi:10.1111/j.1365-294X.2009.04450.x.

[238] Fox CW, Irschick DJ, Knapp AK, Thompson K, Baker L, Meyer J. Functional ecology: moving forward into a new era of publishing. Functional Ecology. 2014;28(2):291–292. doi:10.1111/1365-2435.12254.

[239] Sandhu L, Baker L. Journal of Ecology is part of new BES data archiving policy; 2014. Available from: `https://jecologyblog.com/2014/01/20/journal-of-ecology-is-part-of-new-bes-data-archiving-policy/`.

[240] John Wiley & Sons I. Author Compliance Tool; 2021. Available from: `https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/author-compliance-tool.html`.

[241] Magee AF, May MR, Moore BR. The Dawn of Open Access to Phylogenetic Data. PLOS ONE. 2014;9(10):e110268. doi:10.1371/journal.pone.0110268.

[242] Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive Meta-Analysis; 2013. Available from: `https://www.meta-analysis.com`.

[243] Rosenberg MS, Adams DC, Gurevitch J. Metawin: Statistical software for meta-analysis with resampling tests. Sunderland, MA: Sinauer Associates; 1997.

[244] Hadfield JD. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software. 2010;33(1):1–22. doi:10.18637/jss.v033.i02.

[245] Knuth DE. Literate Programming. The Computer Journal. 1984;27(2):97–111. doi:10.1093/comjnl/27.2.97.

[246] Claerbout JF, Karrenbach M. Electronic documents give reproducible research a new meaning. In: SEG Technical Program Expanded Abstracts 1992. Society of Exploration Geophysicists; 1992. p. 601–604.

[247] Buckheit JB, Donoho DL. WaveLab and Reproducible Research. In: Antoniadis A, Oppenheim G, editors. Wavelets and Statistics. New York, NY: Springer; 1995. p. 55–81. Available from: `https://doi.org/10.1007/978-1-4612-2544-7_5`.

[248] Gentleman R, Temple Lang D. Statistical Analyses and Reproducible Research. Journal of Computational and Graphical Statistics. 2007;16(1):1–23. doi:10.1198/106186007X178663.

[249] R Core Team. R: A Language and Environment for Statistical Computing; 2022. Available from: `https://www.R-project.org/`.

[250] Boettiger C. An introduction to Docker for reproducible research. ACM SIGOPS Operating Systems Review. 2015;49(1):71–79. doi:10.1145/2723872.2723882.

[251] Nüst D, Sochat V, Marwick B, Eglen SJ, Head T, Hirst T, et al. Ten simple rules for writing Dockerfiles for reproducible data science. PLOS Computational Biology. 2020;16(11):e1008316. doi:10.1371/journal.pcbi.1008316.

[252] Boettiger C, Eddelbuettel D. An Introduction to Rocker: Docker Containers for R. The R Journal. 2017;9(2):527–536.

[253] Marwick B, Boettiger C, Mullen L. Packaging Data Analytical Work Reproducibly Using R (and Friends). The American Statistician. 2018;72(1):80–88. doi:10.1080/00031305.2017.1375986.

[254] Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. Journal of Geophysical Research: Biogeosciences. 2009;114(G1). doi:10.1029/2008JG000812.

[255] Rödel HG, Bora A, Kaetzke P, Khaschei M, Hutzelmeyer H, von Holst D. Over-winter survival in subadult European rabbits: weather effects, density dependence, and the impact of individual characteristics. Oecologia. 2004;140(4):566–576. doi:10.1007/s00442-004-1616-1.

[256] Barber-Meyer SM, Mech LD, White PJ. Elk Calf Survival and Mortality Following Wolf Restoration to Yellowstone National Park. Wildlife Monographs. 2008;169(1):1–30. doi:10.2193/2008-004.

S1 META-ANALYSIS IN ECOLOGY AND EVOLUTIONARY BIOLOGY

Any study reporting numerical results (i.e., not just meta-analyses) can potentially be the subject of an attempt to computationally reproduce results, so why focusing on meta-analyses? First, there is a practical imperative here: narrowing the scope of this study makes it tractable. The ecology and evolution literature is vast and varied, and although the literature at large could be sampled to arrive at a manageable subset of articles to assess, the screening process to identify suitable candidate articles and exclude irrelevant ones would be arduous without some sort of guiding principle. In that sense, "meta-analysis" is just one of many potential ways to winnow down the literature, in that it is a quantitative method that produces numerical results that can (in principle) be subject to a computational reproducibility attempt. But still, why narrow the scope to meta-analyses in particular? To contextualise our answer to this, we start with a brief review of meta-analysis in ecology and evolutionary biology.

Meta-analysis in ecology and evolutionary biology

Meta-analysis, a set of statistical methods for combining the results of multiple studies, is a widely-used tool for research synthesis in medicine, the social sciences, and natural sciences (69). Meta-analysis has been used for decades in disciplines such as psychology, education, and especially medicine, where it has become a core tool for assessing the evidence of treatments, in particular via Cochrane systematic reviews (70; 71). In addition to an enormous literature on methods of meta-analysis, guidelines such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) have been developed to standardise how meta-analyses are performed and reported (72; 73). PRISMA has been extended to be relevant to meta-analyses in ecology and evolutionary biology specifically (74).

Compared to some other disciplines, meta-analysis was adopted by ecology and evolutionary biology relatively recently, but has grown substantially, from a handful of meta-analyses published in the early 1990s to over 500 meta-analyses published in 2010 (69). In addition to meta-analyses themselves, there have been numerous methodological papers and handbooks covering how meta-analytical methods can be applied in circumstances specific to ecology/evolutionary biology (e.g., 75–78). Most relevantly for this paper, there have also been reviews of how meta-analyses have been conducted within the fields of ecology and evolutionary biology.

An early review of methods in meta-analysis (79) reviewed 29 meta-analyses published between 1991 and 1998, and is a useful baseline to track how methodology has progressed since the early years of meta-analysis in ecology. Overall, it was observed that techniques used in the medical meta-analysis literature had not been adopted in ecology, in particular methods of assessing publication bias (only 34% of meta-analyses accounted for publication bias, and all calculated a Rosenthal fail-safe number; none used superior alternatives such as funnel plots, regression or the "trim and fill" method—refer to (35) for summaries of these methods). 76% of meta-analyses used the *Q* statistic to explore heterogeneity in effect sizes, and 17% included a sensitivity analysis of some kind. 28% of meta-analysis provided no information on how the primary studies were located. In terms of the effect sizes used by the meta-analyses, 55% used standardised mean differences, 31% used the Pearson correlation coefficient, and 7% used response ratios.

(35) conducted a survey of 100 "biological" meta-analyses (i.e., meta-analyses in ecology and evolutionary biology) published over 2009–2011. They found that only 17% controlled for phylogenetic relatedness between species, and 49% used methods to identify and/or assess the impact of publication bias (specifically, to identify publication bias, about 40% of these meta-analyses used funnel plots, about 10% used a correlation-based method, less than 10% used a quantile plot-based method, and about 5% used a regression-based method. For assessing the impact of publication bias, about 30% calculated Rosenthal fail-safe numbers and less than 10% used the "trim and fill" method). In terms of the effect sizes used by meta-analyses, about 60% used standardised mean differences, about 20% used correlation coefficients and the remainder used other measures.

(80) specifically addressed the point about a need for the term "meta-analysis" to be well-defined in the ecology and conservation biology literature. They examined 133 nominal meta-analyses, and applied a two-stage rating system based on the technical requirements for a meta-analysis according to the medical meta-analysis literature. They found that only 45% of the meta-analyses satisfied all requirements in the first stage of rating, and 25% of the meta-analyses satisfied none. In the second stage of rating which involved only 83 of the meta-analyses which had scored sufficiently highly in the first stage, only a single meta-analysis satisfied all second stage requirements, and 22% of the second stage meta-analyses satisfied none. The authors recommended that going forward, "meta-analyses" in ecology journals ought to include the seven technical requirements outlined in their review.

(81) performed a similar review to both (35) and (80), but focused on meta-analyses in plant ecology. They developed a 16-item rubric to assess the quality of meta-analytic methods, based on previous reviews of meta-analytic methods with some additions and refinements. Each item in the rubric listed meta-analyses that were exemplars of relevant method(s). They used a 14 of these items to assess the methods used in a sample of 322 meta-analyses in plant ecology. The results were mixed: although meta-analyses in plant ecology were highly likely to list all primary studies included in the analysis (87% of meta-analyses in the sample) or

explore caused of heterogeneity in results (89%), only 32% reported the full details of their literature searches, only 31% considered publication bias, and only 25% conducted a sensitivity analysis. 5% of meta-analyses considered changes in study effect sizes over time, and 11% took phylogenetic relatedness into account.

(82) used a 17-item rubric very similar to (81) to evaluate the methods used by 18 meta-analyses published in the journal *Molecular Ecology*. The goal of this paper was to formalise the definition of meta-analysis for editors, authors, reviewers, and readers of the journal. They found that only 22% of studies met the standard expected for an effective meta-analysis, which required satisfying at least 15 of the 17 items, and 56% of meta-analyses satisfied 9 or fewer of the 17 items. The adherence to the different methods was mixed: While 100% of the meta-analyses provided a list of the primary studies included and documented the meta-data extracted from each, and 94% mentioned the inclusion/exclusion criteria used for selected studies, 50% included details of the literature search terms used and only 22% provided details of both the databases searched and dates the searches were conducted. Only 33% of meta-analyses took publication bias into account, 22% quantified the heterogeneity between effect sizes, and only 33% explicitly reported whether they were using a fixed effect or random effects model.

The common theme of heterogeneity in the methods labelled as "meta-analysis" in the ecology literature has led to the emergence of what might be called the "formal" (80; 81) or "narrow" (83) definition of meta-analysis: (81) define meta-analysis "a set of statistical methods for combining outcomes (effect sizes) across different data sets addressing the same research question to examine patterns of response across these data sets and sources of heterogeneity in outcomes", although they do also note that there is no single agreed-upon checklist for assessing whether a given meta-analysis is using the correct methods for this purpose. What does seem clear is that the methods and procedures of the formal/narrow meta-analysis are those mentioned in the checklists/rubrics/rating systems of (80), (81), and (82).

Focusing on meta-analysis in this study

The first and primary reason for choosing to focus on meta-analysis is this: despite the findings outlined in the previous section, meta-analyses are (relatively) uniform in their statistical methods and data, and so restricting the study to meta-analyses allows for the assessment of "like" studies. This has a few different dimensions that speak to the tractability of the study:

- The great variety in quantitative and statistical methods employed across the entire ecology and evolution literature (with the accompanying variety in computational resource requirements) means that failure to computationally reproduce one study but not another could be a result of radically different scales of computational requirements, which is a confounding factor we'd like to avoid as much as possible, due to limited resources. Potential ways of dealing with this (e.g., screening articles to preclude studies with "too high" computational resource requirements) seem too subjective and difficult to operationalise. Choosing a single type of study, meta-analysis, acts to reduce the likely variation in computational resource requirements.

- In general, meta-analytic models are fitted using relatively small data sets (in the order of tens or hundreds, perhaps thousands, of data points as opposed to "big data" with millions of data points) and require modest computational resources (i.e., can be easily run on a desktop or laptop computer with no high performance computing resources required).
- Meta-analyses in particular benefit from the existence of standards for reporting, e.g., PRISMA. Relevantly for this study, this includes standards around the reporting/sharing of data. While a given meta-analysis may not be obliged to strictly adhere to all PRIMSA reporting guidelines, the existence of such guidelines makes it more likely that different studies can be assessed on a like basis than if no such guidelines or standards existed.

The second reason is, as mentioned earlier, meta-analysis has become an important part of the fields of ecology and evolution. To the extent that meta-analyses become regarded (for better or worse) as a higher standard of evidence, it commensurately raises the stakes of meta-analytic results. In that context, being able to assure the results of meta-analyses through computational reproduction has some value.

S2 LITERATURE SEARCH

We set about curating a set of meta-analyses to survey by conducting a Scopus abstract and citation database search (we accessed the Scopus database via the University of Melbourne library's subscription). The search query, conducted on 20th December 2017, searched article titles, abstracts, and keywords for the string "meta-anal*", subject to two constraints. The first constraint restricted results to articles published between 2015 and 2017, inclusive. The second constraint restricted results to articles published in one of 21 ecology and evolution journal titles (identified by ISSN). The journal titles included are as follows: *The American Naturalist*, *Animal Behaviour*, *Behavioral Ecology*, *Behavioral Ecology and Sociobiology*, *Biological Reviews*, *Ecological Applications*, *Ecological Monographs*, *Ecology*, *Ecology Letters*, *Evolution*, *Evolutionary Ecology*, *Functional Ecology*, *Journal of Animal Ecology*, *Journal of Applied Ecology*, *Journal of Ecology*, *Journal of Evolutionary Biology*, *Molecular Ecology*, *New Phytologist*, *Oecologia*, *Oikos*, *Quarterly Review of Biology*.

The Scopus search string used was as follows:

```
TITLE-ABS-KEY ( meta-anal* )  AND  ( PUBYEAR  =  2015
OR  PUBYEAR  =  2016  OR  PUBYEAR  =  2017 )
AND  ISSN ( 0003-0147  OR  0003-3472  OR  1045-2249  OR  0340-5443
OR  1464-7931  OR  1051-0761  OR  0012-9615  OR  0012-9658
OR  1461-023x  OR  0014-3820  OR  0269-7653  OR  0269-8463
OR  0021-8790  OR  0021-8901  OR  0022-0477  OR  1010-061x
OR  0962-1083  OR  0028-646x  OR  0029-8549  OR  0030-1299
OR  0033-5770 )
```

This list of ecology and evolution journal titles is the same as used for the survey of meta-analyses conducted in (35). This choice was made to (i) be assured of searching journals that actively published meta-analyses, and (ii) keep the study tractable: (35) yielded 390 studies from their three-year (2009–11) search of these journal titles and kept the 100 most recent meta-analyses, so that gave an indication of the approximate number of meta-analysis studies we would need to review. It is unclear if this set of journal titles can be considered a “representative” sample of all ecology and evolutionary biology journals; one obvious factor is that not all journals would necessarily consider meta-analyses to be within their scope. However, it seems clear that the list of journals used for this study is not particularly aberrant, at least: for example, (33) reviewed the data and code release policies of 96 “ecology” journals indexed by Web of Science, and the list of 96 journals reviewed includes 17 of the 21 titles surveyed by (35).

Identifying meta-analyses

The search results returned articles which contained the string “meta-anal*” somewhere in the article’s title, abstract, or list of keywords. However, not all such articles will necessarily be meta-analyses. The next step was to screen the articles to obtain a sample of “meta-analyses”. As the review of the ecological meta-analysis methodology literature foreshadowed, this was not straightforward.

The articles were screened using a two-step process: first, some types of articles were checked for and when found, put aside. These article types were (i) errata or corrigenda notices, and (ii) letters or comments in reply to a previously published article. Since errata and comments rely heavily on the context provided by the article they are in reference to (which may or may not be a meta-analysis, and which may or may not be in scope of the literature search), they were considered not suitable to include as “meta-analyses”. We considered these article types to be straightforward to identify (due to clear cues in their title, and other contextual clues such as being included in a comments/letters section of a journal issue), and so removed them from consideration without a formal review of their contents.

The second step involved evaluating the remaining articles in the following way: rather than checking they meet a particular set of methodological requirements, meta-analyses were identified by confirming that an article merely includes a *claim* that it is a meta-analysis (or that a meta-analysis was conducted, or words to that effect) or not. This approach to identifying meta-analyses was intended to be as generous as possible and methodologically agnostic.

Identifying a claim that an article is/conducts a meta-analysis still requires judgment and interpretation, and is subjective. To make the claim identification process transparent, we constructed and employed a simple coding scheme with eight items to summarise the "evidence" in support of each article claiming to be a meta-analysis. The coding scheme is outlined in Table S1. This scheme records the use of the term "meta-analysis" in crucial places in the article (title, abstract, keywords if the article includes them), as well as the quoted text of any actual claim found within the body of the article text.

For items 1–4 and 7, the value "Y" indicates an unambiguous "yes" to the question/contention posed in the column "Description" of Table S1, and the value "N" indicates an unambiguous "no". For items 1–4, the value "U" was available to indicate situations where the mention of "meta-analysis" was somehow unclear. For item 3 only, the value "N/A" was used to indicate that an article did not include any keywords.

Items 5 and 6 record the most substantive piece of evidence: text, directly quoted from the article, which contains the claim to be a meta-analysis (if the claim can be found). Item 7 contains the final judgment of whether the article can be considered to include a claim or not (either yes "Y" or no "N"), and item 8 records any additional notes about the judgment.

All articles remaining after the first step were coded using this scheme. In practice, this meant searching the text of each article for the string "meta" (this word fragment was chosen to avoid issues with the matching of the hyphen in "meta-analysis"), and reviewing all matches in order to answer the coding scheme items.

Articles were considered as claiming to be a meta-analysis if the value of item 7 (Claim in article) in the coding scheme was "Y" . Articles found not to include such a claim (a value of "N" for item 7) were put aside.

The final set of ecology and evolutionary biology meta-analyses, to be the basis of the rest of this study, is simply the set of 177 articles coded as containing claims to be meta-analyses. The bibliographic details of all 177 meta-analysis articles are listed in Table S2.

Index	Field	Values	Description
1	Claim in title	Y, N, U	Does the article include the term “meta-analysis” in its title?
2	Claim in abstract	Y, N, U	Does the article include the term “meta-analysis” in its abstract?
3	Claim in keywords	Y, N, U, N/A	If applicable, is “meta-analysis” one of the article’s keywords?
4	Claim in body text	Y, N, U	Does the article body text contain a claim to be a meta-analysis?
5	Quote of claim	open text	The actual text of the claim as it appears in the article.
6	Quote page number	open text	Page number(s) the quote appears on.
7	Claim in article	Y, N	An overall judgment of whether or not the article claims to be a meta-analysis.
8	Notes	open text	Any additional notes about the article’s meta-analysis claim status.

Table S1: The eight item coding scheme used for determining whether an article claims to be a meta-analysis. In the Values column, “Y” indicates “yes”, “N” indicates “no”, “U” indicates “unclear”, and “N/A” indicates “not applicable”.

Table S2: References for all 177 meta-analysis articles in the data set used in this study.

ID	Study
MA001	Bowles TM, Jackson LE, Loeher M, Cavagnaro TR. Ecological intensification and arbuscular mycorrhizas: a meta-analysis of tillage and cover crop effects. Journal of Applied Ecology. 2017;54(6):1785–1793. doi:10.1111/1365-2664.12815
MA003	Mori AS, Tatsumi S, Gustafsson L. Landscape properties affect biodiversity response to retention approaches in forestry. Journal of Applied Ecology. 2017;54(6):1627–1637. doi:10.1111/1365-2664.12888
MA005	Charlebois JA, Sargent RD. No consistent pollinator-mediated impacts of alien plants on natives. Ecology Letters. 2017;20(11):1479–1490. doi:10.1111/ele.12831
MA006	Martin-StPaul N, Delzon S, Cochard H. Plant resistance to drought depends on timely stomatal closure. Ecology Letters. 2017;20(11):1437–1447. doi:10.1111/ele.12851
MA009	Romano A, Saino N, Møller AP. Viability and expression of sexual ornaments in the barn swallow Hirundo rustica: a meta-analysis. Journal of Evolutionary Biology. 2017;30(10):1929–1935. doi:10.1111/jeb.13151
MA010	Davidson KE, Fowler MS, Skov MW, Doerr SH, Beaumont N, Griffin JN. Livestock grazing alters multiple ecosystem properties and services in salt marshes: a meta-analysis. Journal of Applied Ecology. 2017;54(5):1395–1405. doi:10.1111/1365-2664.12892
MA011	Grant JWA, Weir LK, Steingrímsson SÓ. Territory size decreases minimally with increasing food abundance in stream salmonids: Implications for population regulation. Journal of Animal Ecology. 2017;86(6):1308–1316. doi:10.1111/1365-2656.12737
MA013	Horswill C, O’Brien SH, Robinson RA. Density dependence and marine bird populations: are wind farm assessments precautionary? Journal of Applied Ecology. 2017;54(5):1406–1414. doi:10.1111/1365-2664.12841

ID	Study
MA014	Auer SK, Killen SS, Rezende EL. Resting vs. active: a meta-analysis of the intra- and inter-specific associations between minimum, sustained, and maximum metabolic rates in vertebrates. Functional Ecology. 2017;31(9):1728–1738. doi:10.1111/1365-2435.12879
MA015	Buchanan AL, Hermann SL, Lund M, Szendrei Z. A meta-analysis of non-consumptive predator effects in arthropods: the influence of organismal and environmental characteristics. Oikos. 2017;126(9):1233–1240. doi:10.1111/oik.04384
MA016	Xu X, Medvigy D, Joseph Wright S, Kitajima K, Wu J, Albert LP, et al. Variations of leaf longevity in tropical moist forests predicted by a trait-driven carbon optimality model. Ecology Letters. 2017;20(9):1097–1106. doi:10.1111/ele.12804
MA017	Soria M, Leigh C, Datry T, Bini LM, Bonada N. Biodiversity in perennial and intermittent rivers: a meta-analysis. Oikos. 2017;126(8):1078–1089. doi:10.1111/oik.04118
MA018	Delavaux CS, Smith-Ramesh LM, Kuebbing SE. Beyond nutrients: a meta-analysis of the diverse effects of arbuscular mycorrhizal fungi on plants and soils. Ecology. 2017;98(8):2111–2119. doi:10.1002/ecy.1892
MA019	LaManna JA, Martin TE. Logging impacts on avian species richness and composition differ across latitudes and foraging and breeding habitat preferences. Biological Reviews. 2017;92(3):1657–1674. doi:10.1111/brv.12300
MA020	Leal LC, Peixoto PEC. Decreasing water availability across the globe improves the effectiveness of protective ant–plant mutualisms: a meta-analysis. Biological Reviews. 2017;92(3):1785–1794. doi:10.1111/brv.12307
MA021	Romano A, Costanzo A, Rubolini D, Saino N, Møller AP. Geographical and seasonal variation in the intensity of sexual selection in the barn swallow Hirundo rustica: a meta-analysis. Biological Reviews. 2017;92(3):1582–1600. doi:10.1111/brv.12297
MA022	Anthelme F, Meneses RI, Valero NNH, Pozo P, Dangles O. Fine nurse variations explain discrepancies in the stress-interaction relationship in alpine regions. Oikos. 2017;126(8):1173–1183. doi:10.1111/oik.04248
MA023	Hindrikson M, Remm J, Pilot M, Godinho R, Stronen AV, Baltrūnaité L, et al. Wolf population genetics in Europe: a systematic review, meta-analysis and suggestions for conservation and management. Biological Reviews. 2017;92(3):1601–1629. doi:10.1111/brv.12298
MA024	Knapp JL, Bartlett LJ, Osborne JL. Re-evaluating strategies for pollinator-dependent crops: How useful is parthenocarpy? Journal of Applied Ecology. 2017;54(4):1171–1179. doi:10.1111/1365-2664.12813
MA025	Gázquez A, Beemster GTS. What determines organ size differences between species? A meta-analysis of the cellular basis. New Phytologist. 2017;215(1):299–308. doi:10.1111/nph.14573
MA026	Hendershot JN, Read QD, Henning JA, Sanders NJ, Classen AT. Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. Ecology. 2017;98(7):1757–1763. doi:10.1002/ecy.1829
MA027	Farji-Brener AG, Werenkraut V. The effects of ant nests on soil fertility and plant performance: a meta-analysis. Journal of Animal Ecology. 2017;86(4):866–877. doi:10.1111/1365-2656.12672

ID	Study
MA028	Hitchcock DJ, Varpe Ø, Andersen T, Borgå K. Effects of reproductive strategies on pollutant concentrations in pinnipeds: a meta-analysis. Oikos. 2017;126(6):772–781. doi:10.1111/oik.03955
MA029	Miller SE, Barrueto M, Schluter D. A comparative analysis of experimental selection on the stickleback pelvis. Journal of Evolutionary Biology. 2017;30(6):1165–1176. doi:10.1111/jeb.13085
MA031	Rowiński PK, Rogell B. Environmental stress correlates with increases in both genetic and residual variances: A meta-analysis of animal studies. Evolution. 2017;71(5):1339–1351. doi:10.1111/evo.13201
MA033	Yue K, Fornara DA, Yang W, Peng Y, Peng C, Liu Z, et al. Influence of multiple global change drivers on terrestrial carbon storage: additive effects are common. Ecology Letters. 2017;20(5):663–672. doi:10.1111/ele.12767
MA035	Wood KA, Ponting J, D'Costa N, Newth JL, Rose PE, Glazov P, et al. Understanding intrinsic and extrinsic drivers of aggressive behaviour in waterbird assemblages: a meta-analysis. Animal Behaviour. 2017;126:209–216. doi:10.1016/j.anbehav.2017.02.008
MA036	Greenwood S, Ruiz-Benito P, Martínez-Vilalta J, Lloret F, Kitzberger T, Allen CD, et al. Tree mortality across biomes is promoted by drought intensity, lower wood density and higher specific leaf area. Ecology Letters. 2017;20(4):539–553. doi:10.1111/ele.12748
MA037	Martin PA, Newton AC, Bullock JM. Impacts of invasive plants on carbon pools depend on both species' traits and local climate. Ecology. 2017;98(4):1026–1035. doi:10.1002/ecy.1711
MA038	Holtmann B, Lagisz M, Nakagawa S. Metabolic rates, and not hormone levels, are a likely mediator of between-individual differences in behaviour: a meta-analysis. Functional Ecology. 2017;31(3):685–696. doi:10.1111/1365-2435.12779
MA039	Catano CP, Dickson TL, Myers JA. Dispersal and neutral sampling mediate contingent effects of disturbance on plant beta-diversity: a meta-analysis. Ecology Letters. 2017;20(3):347–356. doi:10.1111/ele.12733
MA040	Poulin R, Pérez-Ponce de León G. Global analysis reveals that cryptic diversity is linked with habitat but not mode of life. Journal of Evolutionary Biology. 2017;30(3):641–649. doi:10.1111/jeb.13034
MA041	Usui T, Butchart SHM, Phillimore AB. Temporal shifts and temperature sensitivity of avian spring migratory phenology: a phylogenetic meta-analysis. Journal of Animal Ecology. 2017;86(2):250–261. doi:10.1111/1365-2656.12612
MA042	Hrycik AR, Almeida LZ, Höök TO. Sub-lethal effects on fish provide insight into a biologically-relevant threshold of hypoxia. Oikos. 2017;126(3):307–317. doi:10.1111/oik.03678
MA044	Rossetti MR, Tscharntke T, Aguilar R, Batáry P. Responses of insect herbivores and herbivory to habitat fragmentation: a hierarchical meta-analysis. Ecology Letters. 2017;20(2):264–272. doi:10.1111/ele.12723
MA045	Sorenson GH, Dey CJ, Madliger CL, Love OP. Effectiveness of baseline corticosterone as a monitoring tool for fitness: a meta-analysis in seabirds. Oecologia. 2017;183(2):353–365. doi:10.1007/s00442-016-3774-3
MA046	Foo YZ, Nakagawa S, Rhodes G, Simmons LW. The effects of sex hormones on immune function: a meta-analysis. Biological Reviews. 2017;92(1):551–571. doi:10.1111/brv.12243

ID	Study
MA048	Hossie T, Landolt K, Murray DL. Determinants and co-expression of anti-predator responses in amphibian tadpoles: a meta-analysis. Oikos. 2017;126(2):n/a–n/a. doi:10.1111/oik.03305
MA049	Booksmythe I, Mautz B, Davis J, Nakagawa S, Jennions MD. Facultative adjustment of the offspring sex ratio and male attractiveness: a systematic review and meta-analysis. Biological Reviews. 2017;92(1):108–134. doi:10.1111/brv.12220
MA052	Ameye M, Allmann S, Verwaeren J, Smagghe G, Haesaert G, Schuurink RC, et al. Green leaf volatile production by plants: a meta-analysis. New Phytologist. 2017; p. n/a–n/a. doi:10.1111/nph.14671
MA053	Vidal MC, Murphy SM. Bottom-up vs. top-down effects on terrestrial insect herbivores: a meta-analysis. Ecology Letters. 2018;21(1):138–150. doi:10.1111/ele.12874
MA054	Deng Q, McMahon DE, Xiang Y, Yu CL, Jackson RB, Hui D. A global meta-analysis of soil phosphorus dynamics after afforestation. New Phytologist. 2017;213(1):181–192. doi:10.1111/nph.14119
MA055	Lee MR, Bernhardt ES, van Bodegom PM, Cornelissen JHC, Kattge J, Laughlin DC, et al. Invasive species' leaf traits and dissimilarity from natives shape their impact on nitrogen cycling: a meta-analysis. New Phytologist. 2017;213(1):128–139. doi:10.1111/nph.14115
MA056	Zvereva EL, Zverev V, Kruglova OY, Kozlov MV. Strategies of chemical anti-predator defences in leaf beetles: is sequestration of plant toxins less costly than de novo synthesis? Oecologia. 2017;183(1):93–106. doi:10.1007/s00442-016-3743-x
MA057	Xiao Z, Wang X, Koricheva J, Kergunteuil A, Le Bayon RC, Liu M, et al. Earthworms affect plant growth and resistance against herbivores: A meta-analysis. Functional Ecology. 2018;32(1):150–160. doi:10.1111/1365-2435.12969
MA058	Camenzind T, Hättenschwiler S, Treseder KK, Lehmann A, Rillig MC. Nutrient limitation of soil microbial processes in tropical forests. Ecological Monographs. 2018;88(1):4–21. doi:10.1002/ecm.1279
MA059	Yang H, Zhang Q, Koide RT, Hoeksema JD, Tang J, Bian X, et al. Taxonomic resolution is a determinant of biodiversity effects in arbuscular mycorrhizal fungal communities. Journal of Ecology. 2017;105(1):219–228. doi:10.1111/1365-2745.12655
MA060	Winternitz J, Abbate JL, Huchard E, Havlíček J, Garamszegi LZ. Patterns of MHC-dependent mate selection in humans and nonhuman primates: a meta-analysis. Molecular Ecology. 2017;26(2):668–688. doi:10.1111/mec.13920
MA061	Molnar JL, Diogo R, Hutchinson JR, Pierce SE. Reconstructing pectoral appendicular muscle anatomy in fossil fish and tetrapods over the fins-to-limbs transition. Biological Reviews. 2017; p. n/a–n/a. doi:10.1111/brv.12386
MA062	Grueber CE, Gray LJ, Morris KM, Simpson SJ, Senior AM. Intergenerational effects of nutrition on immunity: a systematic review and meta-analysis. Biological Reviews. 2018;93(2):1108–1124. doi:10.1111/brv.12387

ID	Study
MA063	Gruner DS, Bracken MES, Berger SA, Eriksson BK, Gamfeldt L, Matthiessen B, et al. Effects of experimental warming on biodiversity depend on ecosystem type and local species composition. Oikos. 2016;126(1):8–17. doi:10.1111/oik.03688
MA064	Dynarski KA, Houlton BZ. Nutrient limitation of terrestrial free-living nitrogen fixation. New Phytologist. 2018;217(3):1050–1061. doi:10.1111/nph.14905
MA065	Noble DWA, Stenhouse V, Schwanz LE. Developmental temperatures and phenotypic plasticity in reptiles: a systematic review and meta-analysis. Biological Reviews. 2018;93(1):72–97. doi:10.1111/brv.12333
MA066	Junker RR, Kuppler J, Amo L, Blande JD, Borges RM, van Dam NM, et al. Covariation and phenotypic integration in chemical communication displays: biosynthetic constraints and eco-evolutionary implications. New Phytologist. 2017; p. n/a–n/a. doi:10.1111/nph.14505
MA067	Risely A, Klaassen M, Hoye BJ. Migratory animals feel the cost of getting sick: A meta-analysis across species. Journal of Animal Ecology. 2017;87(1):301–314. doi:10.1111/1365-2656.12766
MA068	Ronget V, Gaillard J, Coulson T, Garratt M, Gueyffier F, Lega J, et al. Causes and consequences of variation in offspring body mass: meta-analyses in birds and mammals. Biological Reviews. 2017;93(1):1–27. doi:10.1111/brv.12329
MA069	Rohner PT, Teder T, Esperk T, Lüpold S, Blanckenhorn WU. The evolution of male-biased sexual size dimorphism is associated with increased body size plasticity in males. Functional Ecology. 2017;In press:n/a–n/a. doi:10.1111/1365-2435.13004
MA070	Goodell K, Parker IM. Invasion of a dominant floral resource: effects on the floral community and pollination of native plants. Ecology. 2017;98(1):57–69. doi:10.1002/ecy.1639
MA071	Sievers M, Hale R, Parris KM, Swearer SE. Impacts of human-induced environmental change in wetlands on aquatic animals. Biological Reviews. 2017;93(1):529–554. doi:10.1111/brv.12358
MA074	Harts AMF, Booksmythe I, Jennions MD. Mate guarding and frequent copulation in birds: A meta-analysis of their relationship to paternity and male phenotype. Evolution. 2016;70(12):2789–2808. doi:10.1111/evo.13081
MA075	Poorter H, Fiorani F, Pieruschka R, Wojciechowski T, Putten WH, Kleyer M, et al. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. New Phytologist. 2016;212(4):838–855. doi:10.1111/nph.14243
MA076	Edge CB, Houlahan JE, Jackson DA, Fortin M. The response of amphibian larvae to environmental change is both consistent and variable. Oikos. 2016;125(12):1700–1711. doi:10.1111/oik.03166
MA077	Gill AL, Finzi AC. Belowground carbon flux links biogeochemical cycles and resource-use efficiency at the global scale. Ecology Letters. 2016;19(12):1419–1428. doi:10.1111/ele.12690
MA078	Buoro M, Olden JD, Cucherousset J. Global Salmonidae introductions reveal stronger ecological effects of changing intraspecific compared to interspecific diversity. Ecology Letters. 2016;19(11):1363–1371. doi:10.1111/ele.12673

ID	Study
MA079	Besson AA, Lagisz M, Senior AM, Hector KL, Nakagawa S. Effect of maternal diet on offspring coping styles in rodents: a systematic review and meta-analysis. Biological Reviews. 2016;91(4):1065–1080. doi:10.1111/brv.12210
MA080	Flick AJ, Acevedo MA, Elderd BD. The negative effects of pathogen-infected prey on predators: a meta-analysis. Oikos. 2016;125(11):1554–1560. doi:10.1111/oik.03458
MA081	Jaffé R, Pope N, Acosta AL, Alves DA, Arias MC, De la Rúa P, et al. Beekeeping practices and geographic distance, not land use, drive gene flow across tropical bees. Molecular Ecology. 2016;25(21):5345–5358. doi:10.1111/mec.13852
MA082	Tsen EWJ, Sitzia T, Webber BL. To core, or not to core: the impact of coring on tree health and a best-practice framework for collecting dendrochronological information from living trees. Biological Reviews. 2016;91(4):899–924. doi:10.1111/brv.12200
MA090	Derroire G, Balvanera P, Castellanos-Castro C, Decocq G, Kennard DK, Lebrija-Trejos E, et al. Resilience of tropical dry forests – a meta-analysis of changes in species diversity and composition during secondary succession. Oikos. 2016;125(10):1386–1397. doi:10.1111/oik.03229
MA091	Lemoine NP, Hoffman A, Felton AJ, Baur L, Chaves F, Gray J, et al. Underappreciated problems of low replication in ecological field studies. Ecology. 2016;97(10):2554–2561. doi:10.1002/ecy.1506
MA092	Xu X, Medvigy D, Powers JS, Becknell JM, Guan K. Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. New Phytologist. 2016;212(1):80–95. doi:10.1111/nph.14009
MA093	Hidding B, Bakker ES, Hootsmans MJM, Hilt S. Synergy between shading and herbivory triggers macrophyte loss and regime shifts in aquatic systems. Oikos. 2016;125(10):1489–1495. doi:10.1111/oik.03104
MA094	Turney S, Buddle CM. Pyramids of species richness: the determinants and distribution of species diversity across trophic levels. Oikos. 2016;125(9):1224–1232. doi:10.1111/oik.03404
MA095	Gibert A, Gray EF, Westoby M, Wright IJ, Falster DS. On the link between functional traits and growth rate: meta-analysis shows effects change with plant size, as predicted. Journal of Ecology. 2016;104(5):1488–1503. doi:10.1111/1365-2745.12594
MA096	Jessop TS, Lane ML, Teasdale L, Stuart-Fox D, Wilson RS, Careau V, et al. Multiscale Evaluation of Thermal Dependence in the Glucocorticoid Response of Vertebrates. The American Naturalist. 2016;188(3):342–356. doi:10.1086/687588
MA097	Katz DSW. The effects of invertebrate herbivores on plant population growth: a meta-regression analysis. Oecologia. 2016;182(1):43–53. doi:10.1007/s00442-016-3602-9
MA098	Fuzessy LF, Cornelissen TG, Janson C, Silveira FAO. How do primates affect seed germination? A meta-analysis of gut passage effects on neotropical plants. Oikos. 2016;125(8):1069–1080. doi:10.1111/oik.02986
MA099	Cooke J, Leishman MR. Consistent alleviation of abiotic stress with silicon addition: a meta-analysis. Functional Ecology. 2016;30(8):1340–1357. doi:10.1111/1365-2435.12713

ID	Study
MA100	Yoon S, Read Q. Consequences of exotic host use: impacts on Lepidoptera and a test of the ecological trap hypothesis. Oecologia. 2016;181(4):985–996. doi:10.1007/s00442-016-3560-2
MA101	Coetzee BWT, Chown SL. A meta-analysis of human disturbance impacts on Antarctic wildlife. Biological Reviews. 2016;91(3):578–596. doi:10.1111/brv.12184
MA102	Östman Ö, Eklöf J, Eriksson BK, Olsson J, Moksnes PO, Bergström U. Top-down control as important as nutrient enrichment for eutrophication effects in North Atlantic coastal ecosystems. Journal of Applied Ecology. 2016;53(4):1138–1147. doi:10.1111/1365-2664.12654
MA103	He Q, Silliman BR. Consumer control as a common driver of coastal vegetation worldwide. Ecological Monographs. 2016;86(3):278–294. doi:10.1002/ecm.1221
MA106	Delgado-Baquerizo M, Maestre FT, Reich PB, Trivedi P, Osanai Y, Liu YR, et al. Carbon content and climate variability drive global soil bacterial diversity patterns. Ecological Monographs. 2016;86(3):373–390. doi:10.1002/ecm.1216
MA107	Thom D, Seidl R. Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests. Biological Reviews. 2016;91(3):760–781. doi:10.1111/brv.12193
MA108	Liu J, Wu N, Wang H, Sun J, Peng B, Jiang P, et al. Nitrogen addition affects chemical compositions of plant tissues, litter and soil organic matter. Ecology. 2016;97(7):1796–1806. doi:10.1890/15-1683.1
MA109	Liao H, D'Antonio CM, Chen B, Huang Q, Peng S. How much do phenotypic plasticity and local genetic variation contribute to phenotypic divergences along environmental gradients in widespread invasive plants? A meta-analysis. Oikos. 2016;125(7):905–917. doi:10.1111/oik.02372
MA110	Arceo-Gómez G, Ashman TL. Invasion status and phylogenetic relatedness predict cost of heterospecific pollen receipt: implications for native biodiversity decline. Journal of Ecology. 2016;104(4):1003–1008. doi:10.1111/1365-2745.12586
MA111	Oduor AMO, Leimu R, Kleunen Mv. Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. Journal of Ecology. 2016;104(4):957–968. doi:10.1111/1365-2745.12578
MA112	Ficetola GF, Maiorano L. Contrasting effects of temperature and precipitation change on amphibian phenology, abundance and performance. Oecologia. 2016;181(3):683–693. doi:10.1007/s00442-016-3610-9
MA115	Kriengwatana B, Spierings MJ, ten Cate C. Auditory discrimination learning in zebra finches: effects of sex, early life conditions and stimulus characteristics. Animal Behaviour. 2016;116:99–112. doi:10.1016/j.anbehav.2016.03.028
MA117	Samia DSM, Blumstein DT, Stankowich T, Cooper WE. Fifty years of chasing lizards: new insights advance optimal escape theory. Biological Reviews. 2016;91(2):349–366. doi:10.1111/brv.12173
MA118	Smilanich AM, Fincher RM, Dyer LA. Does plant apparency matter? Thirty years of data provide limited support but reveal clear patterns of the effects of plant chemistry on herbivores. New Phytologist. 2016;210(3):1044–1057. doi:10.1111/nph.13875
MA119	Boyle WA, Sandercock BK, Martin K. Patterns and drivers of intraspecific variation in avian life history along elevational gradients: a meta-analysis. Biological Reviews. 2016;91(2):469–482. doi:10.1111/brv.12180

ID	Study
MA120	Blount JD, Vitikainen EIK, Stott I, Cant MA. Oxidative shielding and the cost of reproduction. Biological Reviews. 2016;91(2):483–497. doi:10.1111/brv.12179
MA121	Wesner JS. Contrasting effects of fish predation on benthic versus emerging prey: a meta-analysis. Oecologia. 2016;180(4):1205–1211. doi:10.1007/s00442-015-3539-4
MA122	Moore MP, Riesch R, Martin RA. The predictability and magnitude of life-history divergence to ecological agents of selection: a meta-analysis in livebearing fishes. Ecology Letters. 2016;19(4):435–442. doi:10.1111/ele.12576
MA123	Rowen E, Kaplan I. Eco-evolutionary factors drive induced plant volatiles: a meta-analysis. New Phytologist. 2016;210(1):284–294. doi:10.1111/nph.13804
MA124	Thurman TJ, Barrett RDH. The genetic consequences of selection in natural populations. Molecular Ecology. 2016;25(7):1429–1448. doi:10.1111/mec.13559
MA125	Hébert MP, Beisner BE, Maranger R. A meta-analysis of zooplankton functional traits influencing ecosystem function. Ecology. 2016;97(4):1069–1080. doi:10.1890/15-1084.1
MA126	Anderson JT. Plant fitness in a rapidly changing world. New Phytologist. 2016;210(1):81–87. doi:10.1111/nph.13693
MA127	Lorenz S, Martinez-Fernández V, Alonso C, Mosselman E, García de Jalón D, González del Tánago M, et al. Fuzzy cognitive mapping for predicting hydromorphological responses to multiple pressures in rivers. Journal of Applied Ecology. 2016;53(2):559–566. doi:10.1111/1365-2664.12569
MA128	Dougherty LR, Shuker DM. Variation in pre- and post-copulatory sexual selection on male genital size in two species of lygaeid bug. Behavioral Ecology and Sociobiology. 2016;70(4):625–637. doi:10.1007/s00265-016-2082-6
MA129	Crouzeilles R, Curran M. Which landscape size best predicts the influence of forest cover on restoration success? A global meta-analysis on the scale of effect. Journal of Applied Ecology. 2016;53(2):440–448. doi:10.1111/1365-2664.12590
MA130	van Katwijk MM, Thorhaug A, Marbà N, Orth RJ, Duarte CM, Kendrick GA, et al. Global analysis of seagrass restoration: the importance of large-scale planting. Journal of Applied Ecology. 2016;53(2):567–578. doi:10.1111/1365-2664.12562
MA133	McCary MA, Mores R, Farfan MA, Wise DH. Invasive plants have different effects on trophic structure of green and brown food webs in terrestrial ecosystems: a meta-analysis. Ecology Letters. 2016;19(3):328–335. doi:10.1111/ele.12562
MA135	Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecology Letters. 2016;19(3):336–347. doi:10.1111/ele.12564
MA136	Allen DC, Wesner JS. Synthesis: comparing effects of resource and consumer fluxes into recipient food webs using meta-analysis. Ecology. 2016;97(3):594–604. doi:10.1890/15-1109.1
MA137	Zvereva EL, Kozlov MV. The costs and effectiveness of chemical defenses in herbivorous insects: a meta-analysis. Ecological Monographs. 2016;86(1):107–124. doi:10.1890/15-0911.1

ID	Study
MA138	Barton KE. Tougher and thornier: general patterns in the induction of physical defence traits. Functional Ecology. 2016;30(2):181–187. doi:10.1111/1365-2435.12495
MA140	Street SE, Cross CP, Brown GR. Exaggerated sexual swellings in female nonhuman primates are reliable signals of female fertility and body condition. Animal Behaviour. 2016;112(Supplement C):203–212. doi:10.1016/j.anbehav.2015.11.023
MA145	Moore FR, Shuker DM, Dougherty L. Stress and sexual signaling: a systematic review and meta-analysis. Behavioral Ecology. 2016;27(2):363–371. doi:10.1093/beheco/arv195
MA146	Roca IT, Desrochers L, Giacomazzo M, Bertolo A, Bolduc P, Deschesnes R, et al. Shifting song frequencies in response to anthropogenic noise: a meta-analysis on birds and anurans. Behavioral Ecology. 2016;27(5):1269–1274. doi:10.1093/beheco/arw060
MA147	Holman L. Bet hedging via multiple mating: A meta-analysis. Evolution. 2016;70(1):62–71. doi:10.1111/evo.12822
MA148	Vico G, Manzoni S, Nkurunziza L, Murphy K, Weih M. Trade-offs between seed output and life span – a quantitative comparison of traits between annual and perennial congeneric species. New Phytologist. 2016;209(1):104–114. doi:10.1111/nph.13574
MA149	Daskin JH, Pringle RM. Does primary productivity modulate the indirect effects of large herbivores? A global meta-analysis. Journal of Animal Ecology. 2016;85(4):857–868. doi:10.1111/1365-2656.12522
MA150	Gunton RM, Pöyry J. Scale-specific spatial density dependence in parasitoids: a multi-factor meta-analysis. Functional Ecology. 2016;30(9):1501–1510. doi:10.1111/1365-2435.12627
MA151	German RN, Thompson CE, Benton TG. Relationships among multiple aspects of agriculture's environmental impact and productivity: a meta-analysis to guide sustainable agriculture. Biological Reviews. 2017;92(2):716–738. doi:10.1111/brv.12251
MA152	Shantz AA, Lemoine NP, Burkepile DE. Nutrient loading alters the performance of key nutrient exchange mutualisms. Ecology Letters. 2016;19(1):20–28. doi:10.1111/ele.12538
MA153	Dillingham PW, Moore JE, Fletcher D, Cortés E, Curtis KA, James KC, et al. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. Ecological Applications. 2016;26(1):322–333. doi:10.1890/14-1990
MA155	Strader ME, Aglyamova GV, Matz MV. Red fluorescence in coral larvae is associated with a diapause-like state. Molecular Ecology. 2016;25(2):559–569. doi:10.1111/mec.13488
MA157	Wood KA, O'Hare MT, McDonald C, Searle KR, Daunt F, Stillman RA. Herbivore regulation of plant abundance in aquatic ecosystems. Biological Reviews. 2017;92(2):1128–1141. doi:10.1111/brv.12272
MA158	Deng Q, Hui D, Luo Y, Elser J, Wang YP, Loladze I, et al. Down-regulation of tissue N:P ratios in terrestrial plants by elevated CO2. Ecology. 2015;96(12):3354–3362. doi:10.1890/15-0217.1
MA159	Garamszegi LZ, Markó G, Szász E, Zsebők S, Azcárate M, Herczeg G, et al. Among-year variation in the repeatability, within- and between-individual, and phenotypic correlations of behaviors in a natural population. Behavioral Ecology and Sociobiology. 2015;69(12):2005–2017. doi:10.1007/s00265-015-2012-z

ID	Study
MA160	Lüpold S, Simmons LW, Tomkins JL, Fitzpatrick JL. No evidence for a trade-off between sperm length and male premating weaponry. Journal of Evolutionary Biology. 2015;28(12):2187–2195. doi:10.1111/jeb.12742
MA162	Moreno-Mateos D, Meli P, Vara-Rodríguez MI, Aronson J. Ecosystem response to interventions: lessons from restored and created wetland ecosystems. Journal of Applied Ecology. 2015;52(6):1528–1537. doi:10.1111/1365-2664.12518
MA163	Katano I, Doi H, Eriksson BK, Hillebrand H. A cross-system meta-analysis reveals coupled predation effects on prey biomass and diversity. Oikos. 2015;124(11):1427–1435. doi:10.1111/oik.02430
MA164	Graham S, Chapuis E, Meconcelli S, Bonel N, Sartori K, Christophe A, et al. Size-assortative mating in simultaneous hermaphrodites: an experimental test and a meta-analysis. Behavioral Ecology and Sociobiology. 2015;69(11):1867–1878. doi:10.1007/s00265-015-1999-5
MA168	Goessling JM, Kennedy H, Mendonça MT, Wilson AE. A meta-analysis of plasma corticosterone and heterophil : lymphocyte ratios – is there conservation of physiological stress responses over time? Functional Ecology. 2015;29(9):1189–1196. doi:10.1111/1365-2435.12442
MA169	Romero GQ, Gonçalves-Souza T, Vieira C, Koricheva J. Ecosystem engineering effects on species diversity across ecosystems: a meta-analysis. Biological Reviews. 2015;90(3):877–890. doi:10.1111/brv.12138
MA170	Nielsen JM, Popp BN, Winder M. Meta-analysis of amino acid stable nitrogen isotope ratios for estimating trophic position in marine organisms. Oecologia. 2015;178(3):631–642. doi:10.1007/s00442-015-3305-7
MA171	James J, Slater FM, Vaughan IP, Young KA, Cable J. Comparing the ecological impacts of native and invasive crayfish: could native species' translocation do more harm than good? Oecologia. 2015;178(1):309–316. doi:10.1007/s00442-014-3195-0
MA176	Esteban R, Barrutia O, Artetxe U, Fernández-Marín B, Hernández A, García-Plazaola JI. Internal and external factors affecting photosynthetic pigment composition in plants: a meta-analytical approach. New Phytologist. 2015;206(1):268–280. doi:10.1111/nph.13186
MA178	Becker DJ, Streicker DG, Altizer S. Linking anthropogenic resources to wildlife–pathogen dynamics: a review and meta-analysis. Ecology Letters. 2015;18(5):483–495. doi:10.1111/ele.12428
MA179	Starko S, Claman BZ, Martone PT. Biomechanical consequences of branching in flexible wave-swept macroalgae. New Phytologist. 2015;206(1):133–140. doi:10.1111/nph.13182
MA180	Hsu YH, Schroeder J, Winney I, Burke T, Nakagawa S. Are extra-pair males different from cuckolded males? A case study and a meta-analytic examination. Molecular Ecology. 2015;24(7):1558–1571. doi:10.1111/mec.13124
MA181	Mazza CA, Ballaré CL. Photoreceptors UVR8 and phytochrome B cooperate to optimize plant growth and defense in patchy canopies. New Phytologist. 2015;207(1):4–9. doi:10.1111/nph.13332
MA182	Jackson MC. Interactions among multiple invasive animals. Ecology. 2015;96(8):2035–2041. doi:10.1890/15-0171.1
MA183	Arct A, Drobniak SM, Cichoń M. Genetic similarity between mates predicts extrapair paternity—a meta-analysis of bird studies. Behavioral Ecology. 2015;26(4):959–968. doi:10.1093/beheco/arv004

ID	Study
MA184	Boudreau SA, Anderson SC, Worm B. Top-down and bottom-up forces interact at thermal range extremes on American lobster. Journal of Animal Ecology. 2015;84(3):840–850. doi:10.1111/1365-2656.12322
MA185	Yang LY, Machado CA, Dang XD, Peng YQ, Yang DR, Zhang DY, et al. The incidence and pattern of copollinator diversification in dioecious and monoecious figs. Evolution. 2015;69(2):294–304. doi:10.1111/evo.12584
MA186	Siefert A, Violle C, Chalmandrier L, Albert CH, Taudiere A, Fajardo A, et al. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. Ecology Letters. 2015;18(12):1406–1419. doi:10.1111/ele.12508
MA187	Beninde J, Veith M, Hochkirch A. Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. Ecology Letters. 2015;18(6):581–592. doi:10.1111/ele.12427
MA188	Senior AM, Nakagawa S, Lihoreau M, Simpson SJ, Raubenheimer D. An Overlooked Consequence of Dietary Mixing: A Varied Diet Reduces Interindividual Variance in Fitness. The American Naturalist. 2015;186(5):649–659. doi:10.1086/683182
MA189	Albertson LK, Allen DC. Meta-analysis: abundance, behavior, and hydraulic energy shape biotic effects on sediment transport in streams. Ecology. 2015;96(5):1329–1339. doi:10.1890/13-2138.1
MA190	Frankham R. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Molecular Ecology. 2015;24(11):2610–2618. doi:10.1111/mec.13139
MA191	Voje KL. Scaling of Morphological Characters across Trait Type, Sex, and Environment: A Meta-analysis of Static Allometries. The American Naturalist. 2015;187(1):89–98. doi:10.1086/684159
MA192	Slot M, Kitajima K. General patterns of acclimation of leaf respiration to elevated temperatures across biomes and plant types. Oecologia. 2015;177(3):885–900. doi:10.1007/s00442-014-3159-4
MA193	Jahnke M, Olsen JL, Procaccini G. A meta-analysis reveals a positive correlation between genetic diversity metrics and environmental status in the long-lived seagrass Posidonia oceanica. Molecular Ecology. 2015;24(10):2336–2348. doi:10.1111/mec.13174
MA194	Iacarella JC, Dick JTA, Alexander ME, Ricciardi A. Ecological impacts of invasive alien species along temperature gradients: testing the role of environmental matching. Ecological Applications. 2015;25(3):706–716. doi:10.1890/14-0545.1
MA195	Quesnelle PE, Lindsay KE, Fahrig L. Relative effects of landscape-scale wetland amount and landscape matrix quality on wetland vertebrates: a meta-analysis. Ecological Applications. 2015;25(3):812–825. doi:10.1890/14-0362.1
MA197	Dougherty LR, Shuker DM. The effect of experimental design on the measurement of mate choice: a meta-analysis. Behavioral Ecology. 2015;26(2):311–319. doi:10.1093/beheco/aru125
MA198	Paz-Vinas I, Loot G, Stevens VM, Blanchet S. Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. Molecular Ecology. 2015;24(18):4586–4604. doi:10.1111/mec.13345

ID	Study
MA199	Sistla SA, Appling AP, Lewandowska AM, Taylor BN, Wolf AA. Stoichiometric flexibility in response to fertilization along gradients of environmental and organismal nutrient richness. Oikos. 2015;124(7):949–959. doi:10.1111/oik.02385
MA200	Jauni M, Gripenberg S, Ramula S. Non-native plant species benefit from disturbance: a meta-analysis. Oikos. 2015;124(2):122–129. doi:10.1111/oik.01416
MA201	Bunn RA, Ramsey PW, Lekberg Y. Do native and invasive plants differ in their interactions with arbuscular mycorrhizal fungi? A meta-analysis. Journal of Ecology. 2015;103(6):1547–1556. doi:10.1111/1365-2745.12456
MA202	Mehrabi Z, Tuck SL. Relatedness is a poor predictor of negative plant–soil feedbacks. New Phytologist. 2015;205(3):1071–1075. doi:10.1111/nph.13238
MA203	Wang X, Taub DR, Jablonski LM. Reproductive allocation in plants as affected by elevated carbon dioxide and other environmental changes: a synthesis using meta-analysis and graphical vector analysis. Oecologia. 2015;177(4):1075–1087. doi:10.1007/s00442-014-3191-4
MA204	Albert A, Auffret AG, Cosyns E, Cousins SAO, D'hondt B, Eichberg C, et al. Seed dispersal by ungulates as an ecological filter: a trait-based meta-analysis. Oikos. 2015;124(9):1109–1120. doi:10.1111/oik.02512
MA205	Mijangos JL, Pacioni C, Spencer PBS, Craig MD. Contribution of genetics to ecological restoration. Molecular Ecology. 2015;24(1):22–37. doi:10.1111/mec.12995
MA206	Tamburello L, Maggi E, Benedetti-Cecchi L, Bellistri G, Rattray AJ, Ravaglioli C, et al. Variation in the impact of non-native seaweeds along gradients of habitat degradation: a meta-analysis and an experimental test. Oikos. 2015;124(9):1121–1131. doi:10.1111/oik.02197
MA207	Ferreira V, Castagneyrol B, Koricheva J, Gulis V, Chauvet E, Graça MAS. A meta-analysis of the effects of nutrient enrichment on litter decomposition in streams. Biological Reviews. 2015;90(3):669–688. doi:10.1111/brv.12125
MA208	Stephens AEA, Westoby M. Effects of insect attack to stems on plant survival, growth, reproduction and photosynthesis. Oikos. 2015;124(3):266–273. doi:10.1111/oik.01809
MA210	Anderegg WRL. Spatial and temporal variation in plant hydraulic traits and their relevance for climate change impacts on vegetation. New Phytologist. 2015;205(3):1008–1014. doi:10.1111/nph.12907
MA211	Yuan ZY, Chen HYH. Negative effects of fertilization on plant nutrient resorption. Ecology. 2015;96(2):373–380. doi:10.1890/14-0140.1
MA212	Valls A, Coll M, Christensen V. Keystone species: toward an operational concept for marine biodiversity conservation. Ecological Monographs. 2015;85(1):29–47. doi:10.1890/14-0306.1
MA213	Colautti RI, Lau JA. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. Molecular Ecology. 2015;24(9):1999–2017. doi:10.1111/mec.13162
MA214	Fuiman LA, Connelly TL, Lowerre-Barbieri SK, McClelland JW. Egg boons: central components of marine fatty acid food webs. Ecology. 2015;96(2):362–372. doi:10.1890/14-0571.1

ID	Study
MA215	Baeten L, Davies TJ, Verheyen K, Calster HV, Vellend M. Disentangling dispersal from phylogeny in the colonization capacity of forest understorey plants. Journal of Ecology. 2015;103(1):175–183. doi:10.1111/1365-2745.12333
MA217	Taranu ZE, Gregory-Eaves I, Leavitt PR, Bunting L, Buchaca T, Catalan J, et al. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. Ecology Letters. 2015;18(4):375–384. doi:10.1111/ele.12420
MA218	Pintor LM, Byers JE. Do native predators benefit from non-native prey? Ecology Letters. 2015;18(11):1174–1180. doi:10.1111/ele.12496
MA220	Harper KA, Macdonald SE, Mayerhofer MS, Biswas SR, Esseen PA, Hylander K, et al. Edge influence on vegetation at natural and anthropogenic edges of boreal forests in Canada and Fennoscandia. Journal of Ecology. 2015;103(3):550–562. doi:10.1111/1365-2745.12398
MA222	Culina A, Radersma R, Sheldon BC. Trading up: the fitness consequences of divorce in monogamous birds. Biological Reviews. 2015;90(4):1015–1034. doi:10.1111/brv.12143
MA223	Bracken MES, Hillebrand H, Borer ET, Seabloom EW, Cebrian J, Cleland EE, et al. Signatures of nutrient limitation and co-limitation: responses of autotroph internal nutrient concentrations to nitrogen and phosphorus additions. Oikos. 2015;124(2):113–121. doi:10.1111/oik.01215
MA224	Périquet S, Fritz H, Revilla E. The Lion King and the Hyaena Queen: large carnivore interactions and coexistence. Biological Reviews. 2015;90(4):1197–1214. doi:10.1111/brv.12152
MA226	Elliott KH, Hare JF, Vaillant ML, Gaston AJ, Ropert-Coudert Y, Anderson WG. Ageing gracefully: physiology but not behaviour declines with age in a diving seabird. Functional Ecology. 2015;29(2):219–228. doi:10.1111/1365-2435.12316
MA227	Lafuente A, Pérez-Palacios P, Doukkali B, Molina-Sánchez MD, Jiménez-Zurdo JI, Caviedes MA, et al. Unraveling the effect of arsenic on the model Medicago–Ensifer interaction: a transcriptomic meta-analysis. New Phytologist. 2015;205(1):255–272. doi:10.1111/nph.13009
MA229	Gamfeldt L, Lefcheck JS, Byrnes JEK, Cardinale BJ, Duffy JE, Griffin JN. Marine biodiversity and ecosystem functioning: what's known and what's next? Oikos. 2015;124(3):252–265. doi:10.1111/oik.01549

S3 REVIEW OF JOURNAL POLICIES ON DATA AND CODE SHARING

The availability of data and code for individual articles needs to be evaluated in the context of the publishing journals' policies about making data and code available at the time of publication. Due to the retrospective nature of this study, information about the journals' data and code policies contemporaneous with the articles published 2015–17 was not available. The journals' policies on data and code were inferred from other sources, including previous studies of journal policies and initiatives such as the Joint Data Archiving Policy (JDAP).

Data policies

JDAP was adopted by a number of journals in the fields of ecology and evolutionary biology in 2011 (235). JDAP introduced data archiving as a requirement for publication: the data that underlie the results of the article must be deposited in a public data repository, such as the Dryad Digital Repository (36). Four of the journals in this study adopted JDAP in 2011: *Evolution* (236), *Journal of Evolutionary Biology* (25), *Molecular Ecology* (237), and *The American Naturalist* (26). *Functional Ecology* adopted a slight variation of JDAP in 2014 (238), along with other journals published by the British Ecological Society: *Journal of Animal Ecology*, *Journal of Applied Ecology*, and *Journal of Ecology* (239). Therefore, assuming the journals' adoption of JDAP (or slight variation thereof) has persisted, we expected that these eight journals would have mandated data archiving for all studies published through 2015–17.

(33) investigated both the data and code policies for 17 of the 21 journal titles in this study. The policies checked were as of 1st June 2015, which is within the 2015–17 time period. Specifically, regarding data, (33) recorded whether journals' policies *required* data to be released as a condition of publication—that is, beyond mere encouragement to make data available. We regarded the findings in (33) as representing journals' policies on data and code at the start of 2015 (it is possible that some meta-analyses published in the first five months of 2015 were published under a different journal policy that then changed to the policy found by (33), but for simplicity we discounted this possibility). The four journals that were not reviewed in (33) are *Animal Behaviour*, *Biological Reviews*, *New Phytologist*, and *Quarterly Review of Biology*. We examined other sources to get an indication of their data and code policies.

The data policy of the journal *Animal Behaviour* was surveyed in January 2014 by (30). This survey found that the journal encouraged authors to make data available, but did not make it mandatory. In the absence of other information (which was searched for in e.g., editorials or news releases, but not found), we assume that this was the data policy of the journal during 2015–17. (When checked again in 2021, the journal was found to have the same policy of encouraging data sharing, so it seems safe to assume the policy has been consistently in place since 2014.) When checked in 2021, the journal *Biological Reviews* "encourages" authors to make data available, but does not require authors to do so, or to include data availability statements (240). In the absence of other information, we assume that this was the data policy of the journal during 2015–17. The data policy of the journal *New Phytologist* was surveyed in August/September 2013 by (241). This survey classified the policy of *New Phytologist* as *weak*, meaning that data sharing was encouraged but not required. In the absence of other information, we assume that this was the data policy of the journal during 2015–17. When checked in 2021, the instructions to authors webpage for the journal *Quarterly Review of*

Biology[1] makes no mention of data sharing, archiving, or availability. Similarly, an archived snapshot of the instructions to authors webpage as it was on 28th May 2016[2] made no mention of any data policy. In the absence of other information, we assume that not requiring data sharing was the effective data policy of the journal during 2015–17. A summary of the data-sharing policies of the journals in this study is given in Table S3.

Journal	JDAP member	Data sharing required?	Source
Animal Behaviour	N	N	(30)
Behavioral Ecology	N	N	(33)
Behavioral Ecology and Sociobiology	N	N	(33)
Biological Reviews	N	N	Journal website
Ecological Applications	N	Y	(33)
Ecological Monographs	N	Y	(33)
Ecology	N	Y	(33)
Ecology Letters	N	Y	(33)
Evolution	Y	Y	(33)
Evolutionary Ecology	N	N	(33)
Functional Ecology	Y	Y	(33)
Journal of Animal Ecology	Y	Y	(33)
Journal of Applied Ecology	Y	Y	(33)
Journal of Ecology	Y	Y	(33)
Journal of Evolutionary Biology	Y	Y	(33)
Molecular Ecology	Y	Y	(33)
New Phytologist	N	N	(241)
Oecologia	N	N	(33)
Oikos	N	Y	(33)
The American Naturalist	Y	Y	(33)
The Quarterly Review of Biology	N	N	Journal website

Table S3: Summary of whether data sharing was found to be required for each journal surveyed in this study, along with JDAP member status and source of the information. In the columns "JDAP member" and "Data sharing required?", "Y" indicates "yes" and "N" indicates "no".

Code policies

The principal source for information about journals' code policies came from (33). (34) updated the information about journals' code policies in 2020. Both studies recorded information about the code policies of 17 of the journals included in this present study. (33) recorded whether journals required the release of code as a requirement for publication as a binary yes/no variable (the same way as how journals' data policies were recorded). The updated survey in (34) distinguished between policies where code sharing was "encour-

---

[1] https://www.journals.uchicago.edu/journals/qrb/instruct

[2] https://web.archive.org/web/20160528051141/http://www.journals.uchicago.edu/journals/qrb/instruct

aged" and policies where code sharing was "mandatory" (the authors note that some journal policies were ambiguously worded such that it could not be determined whether code sharing was merely encouraged or a mandatory requirement; they designated such policies "encouraged/mandatory"). For the four journals not covered in (33), we found that *Animal Behaviour* had a policy of encouraging code sharing, but we could not find mention of polic(ies) about code in the online information for *Biological Reviews*, *New Phytologist*, and *The Quarterly Review of Biology*. For the purposes of this study, we shall regard these four journals as not having had a policy requiring code sharing during 2015–17.

Journal	2015 survey	2020 survey	2021 check
Animal Behaviour	-	-	E
Behavioral Ecology	N	N	-
Behavioral Ecology and Sociobiology	N	N	-
Biological Reviews	-	-	N.F.
Ecological Applications	Y	M	-
Ecological Monographs	Y	M	-
Ecology	Y	M	-
Ecology Letters	N	E/M	-
Evolution	N	M	-
Evolutionary Ecology	N	E	-
Functional Ecology	Y	E/M	-
Journal of Animal Ecology	Y	E/M	-
Journal of Applied Ecology	Y	E/M	-
Journal of Ecology	Y	E/M	-
Journal of Evolutionary Biology	N	M	-
Molecular Ecology	Y	E	-
New Phytologist	-	-	N.F.
Oecologia	N	N	-
Oikos	N	N	-
The American Naturalist	Y	E	-
The Quarterly Review of Biology	-	-	N.F.

Table S4: Summary of the code sharing policies found for each journal. The column "2015 survey" refers to (33), the column "2020 survey" refers to (34), and the column "2021 check" refers to our own checks made in 2021. Within the table columns, "Y" indicates "yes", "N" indicates "no", "E" indicates "encouraged", "M" indicates "mandatory", and "N.F." indicates "not found".

S4 CODING SCHEME FOR CODE AND DATA SHARING

The assessment process for each article for shared data and code was as follows: first, we inspected the end sections of each article for any mention of supplemental material, and for the existence of a data/code availability statement of any kind. In cases without an explicit data availability statement, or where data/code were not listed as supplements, we reviewed the methods and results sections for any possible in-text mention of data/code availability, first by performing a keyword search for "data". Regardless of what was mentioned in the article, we also inspected the journal webpage for each article (accessed via The University of Melbourne library) for indications and details of supplemental materials, shared data and shared code. We attempted to download and briefly inspect all files at the journal webpage that we found. Where supplemental material, data and/or code were reported as existing at other web links (e.g., an online data archive), we followed the web links and attempted to download and inspect all files we found. The coding scheme in Table S5 captures the results of this process.

This coding scheme assumes that if data and/or code were shared, there would be some positive indication of this fact somewhere in the article itself, or on the journal publisher's web page for the article (either as supplemental material, or as a link to an independent resource). An absence of any such indication was taken to mean that data/code was not shared. This approach does not account for the possibility that authors may have in fact shared the data and code associated with their article (say, by publishing it in a data repository such as Dryad) but not included any indication either in the article itself, or on the journal web page for the article. (One possible reason this might occur is when authors decide to share the data/code after the article had been published.) We decided not to attempt to check for such possibilities when assessment of an article and its journal web page found no indications of shared data or code.

In this coding scheme, items 1–4 concern supplemental material in general, items 5–12 concern shared data in particular, and items 13–20 concern shared code in particular (item 21 was used to record any additional notes). The items recording the existence (or not) of shared supplemental materials, data, and code are items 1, 5–6, and 13–14. For the data and code sharing, we separated out the nominal sharing of these from the actual sharing of these (in retrospect, we should have done the same for supplemental materials too). In this context, data and code were recorded as having been actually shared only if we were personally able to successfully download (via The University of Melbourne library) and inspect the relevant file(s).

The numbers of files shared (items 2, 7, and 15) were recorded to help keep track of downloaded files. These values were recorded only if the respective preceding items indicated that such files existed. This did lead to some anomalies: the number of files was recorded as 0 in cases where the article stated that all relevant data was made available within tables of the article itself (and not as a separate data file).

S5 RECORDING MENTIONS OF SOFTWARE USED

The review process for mentions of software in an article consisted of checking the text of each article/supplementary document for the following keywords (using a case-insensitive search):

- "CMA", referring to the software package *Comprehensive Meta-Analysis* (242);
- "MetaWin", referring to the software package *MetaWin* (243);

Index	Field	Values	Description
1	Supplements included	Y, N	Does the article include supplementary information?
2	No. supplement files	0-99	Number of discrete files or documents included as supplementary information
3	Supplements mentioned	Y, N	Does the article *mention* the existence of the supplementary information?
4	Supplements detailed	Y, N	Does the article provide details of the contents of supplementary information?
5	Datasets nominally included	Y, N	Does the article *indicate* that data has been shared, included?
6	Datasets included	Y, N	Was the data actually included (shared) and obtainable?
7	No. data files	0-99	Number of discrete data files included
8	Dataset sources	open text	Location of the datasets (e.g., repository name)
9	Dataset URL	open text	Link to data as applicable
10	Dataset info in article	Y, N	Is the availability of data referred to in the article?
11	Dataset info on website	Y, N	Is the availability of data referred to on the journal web page for the article?
12	Data format	open text	File format(s) of data files
13	Code nominally included	Y, N	Does the article *indicate* that code has been shared, included?
14	Code included	Y, N	Was the code actually included (shared) and obtainable?
15	No. code files	0-99	Number of discrete code files included
16	Code sources	open text	Location of the code (e.g., repository name)
17	Code URL	open text	Link to code as applicable
18	Code info in article	Y, N	Is the availability of code referred to in the article?
19	Code info on website	Y, N	Is the availability of code referred to on the journal web page for the article?
20	Code type	open text	Language or software package the code is associated with
21	Notes	open text	Any additional notes about the article's data and code sharing.

Table S5: The twenty-one item coding scheme used for recording data and code sharing in meta-analysis articles. In the Values column, "Y" indicates "yes" and "N" indicates "no".

- “metafor”, referring to the R package *metafor* (37);
- “mcmcglmm”, referring to the R package *mcmcglmm* (244).

In the absence of these keywords being found, the methods section/supplementary document was manually scanned for statements along the lines of “analyses were performed using [software package]”.

For each mention of software used (allowing for multiple mentions per article), the details were recorded using a ten-item coding scheme outlined in Table S6. The coding scheme was designed around an expectation of the frequent mention of R and R packages.

Items 1 and 2 record the name of the software package/platform as reported in the article and the page number of the mention respectively.

Items 3 and 4 record whether a specific version of the software was reported. Items 5, 6, and 7 are specific to the R software environment.

Item 5 is a flag indicating whether the mentioned software package was an R package or not. This required judgment beyond what was reported in the article: For most software mentioned, we were able to code this item based on our own knowledge of R and its packages; where we were not already familiar with the software package, we used contextual clues in the article (e.g., mentions of the function of the software or details from the citation if provided) and online searches of the software name to determine whether or not it was an R package.

Item 6 was only applicable to software identified as an R package: this recorded the location where the R package was hosted. We anticipated that there would be few discrete categories here: “base” referring to packages which are part of the base R installation; “Bioconductor” referring to R packages released as components of the Bioconductor project; “CRAN” referring to the Comprehensive R Archive Network, a repository for R packages; and “other” for all remaining cases.

Item 7 is applicable only to mentions of the R software environment at large: this records whether in addition to the mention of R, specific R packages mentioned as well.

Items 8 and 9 record whether and how the article cited/provided a reference for the software mentioned. Item 8 was initially “Y”/“N” (yes/no), during the coding process we decided to introduce an addition code “T” which was for instances of an “in text” reference for the software (e.g., the website for the software package in parentheses immediately following the software name) but with no corresponding details in the “References” section of the article. As a result, a value of “Y” indicates that the article includes a full reference to the software in the References section. The full reference (or in-text only citation) as reported in the article is recorded in Item 9.

Finally, Item 10 was used to record additional notes/context about the mention of the software as applicable.

Index	Field	Values	Description
1	Software details	open text	The name of the software as reported in the article.
2	Page reference	open text	Specify the page number of the mention.
3	Version specified	Y, N	Does the article specify the version of the software?
4	Version details	open text	The version details as reported in the article.
5	Is R package	Y, N	Is the software mentioned an R package?
6	R package location	base, Bioconductor, CRAN, other, N/A	If the software mentioned is an R package, where is the package located/hosted?
7	R packages mentioned	Y, N, N/A	If the software mentioned is R, are packages mentioned elsewhere in the article?
8	Software cited	Y, N, T	Does the article include a citation for the software package?
9	Citation details	open text	The full reference to the software as reported in the article.
10	Notes	open text	Any additional notes about this mention of software.

Table S6: The ten item coding scheme used for recording software mentions. In the Values column, “Y” indicates “yes”, “N” indicates “no”, “T” indicates “in-text only”, and “N/A” indicates “not applicable”.

S6 DATA AND CODE SHARING

For the 133 articles with data, we had found some kind of indication about data availability somewhere in the article itself or in the supplementary documentation for all but one article (in this one case, the information indicating that data was available was on the journal's web page for the article instead). This took the form of either an explicit data availability statement in the article, or a mention in the body of the article, as part of an in-article statement about the content of supplemental/supporting information, or in the supplemental/supporting information itself. For example, *Evolution* articles included a data availability statement in a dedicated section titled "Data Archiving" located at the end of the article, just before the references section.

Failures to obtain data and code

We failed to obtain data for five articles for three reasons: for the first three cases, a supplemental document indicated that data files were included as part of the supplemental material. However, the files referred to could not be found as part of the online supplement; it is possible that while the documentation for the data was uploaded, the actual files themselves were not. In the fourth case, the data availability statement said that data would be uploaded to Dryad upon acceptance of the article, however no link or details of how to find the data were provided (failing to update the data availability statement may have been an oversight when the article was being finalised for publication). In the final case, the article stated that data had been deposited in a research institute's database, but failed to provide any details apart from a link to the institute's main web page. The institute maintains a number of databases, and there was no clear way to identify which data in which database was relevant to the meta-analysis.

The one case where we could not obtain code is the one of the articles discussed above in reference to data availability, where files listed as being part of the supplement could not be found.

Data and Code Sharing by Journal

Figure S1 shows the rate of data sharing within each journal title. Figure S2 compares the observed rate of data sharing amongst JPAP journals to that of non-JDAP journals. Figure S3 compares the observed rate of data sharing between all journals requiring data sharing (JDAP or not) against those journals without a data sharing requirement. These figures show that in this sample, journal policies did increase the likelihood that a meta-analysis article would include accessible data, and although clearly a common practice among researchers submitting to these journals, adherence to the data sharing policy was not 100%.

Figure S4 shows the rate of code sharing within each journal title. Figure S5 compares the sharing rates for code across three categories: code required, code not required, and policy not known. This figure re-

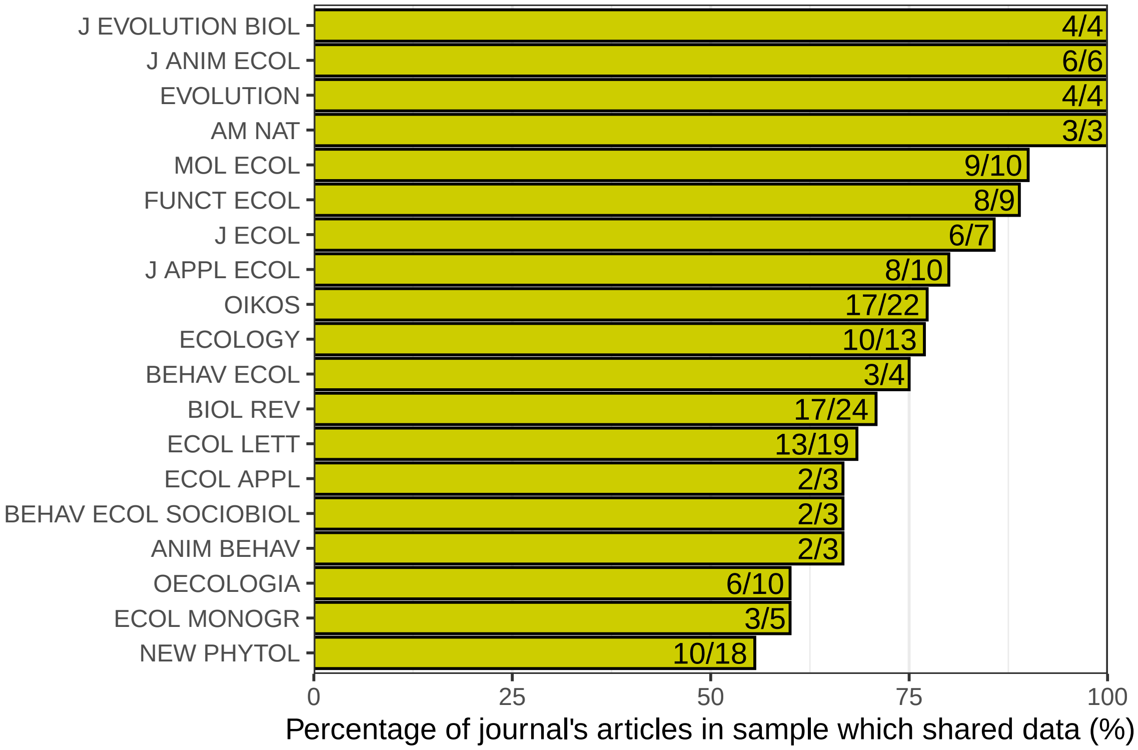

Figure S1: Comparison of data sharing rates in articles by journal.

emphasises that code sharing during this period was low, despite some journals' policies requiring code (these journals did have the highest rate of code sharing, but it was still less than 25% of articles published). During this period, code sharing lagged well behind data sharing (which was common practice in comparison).

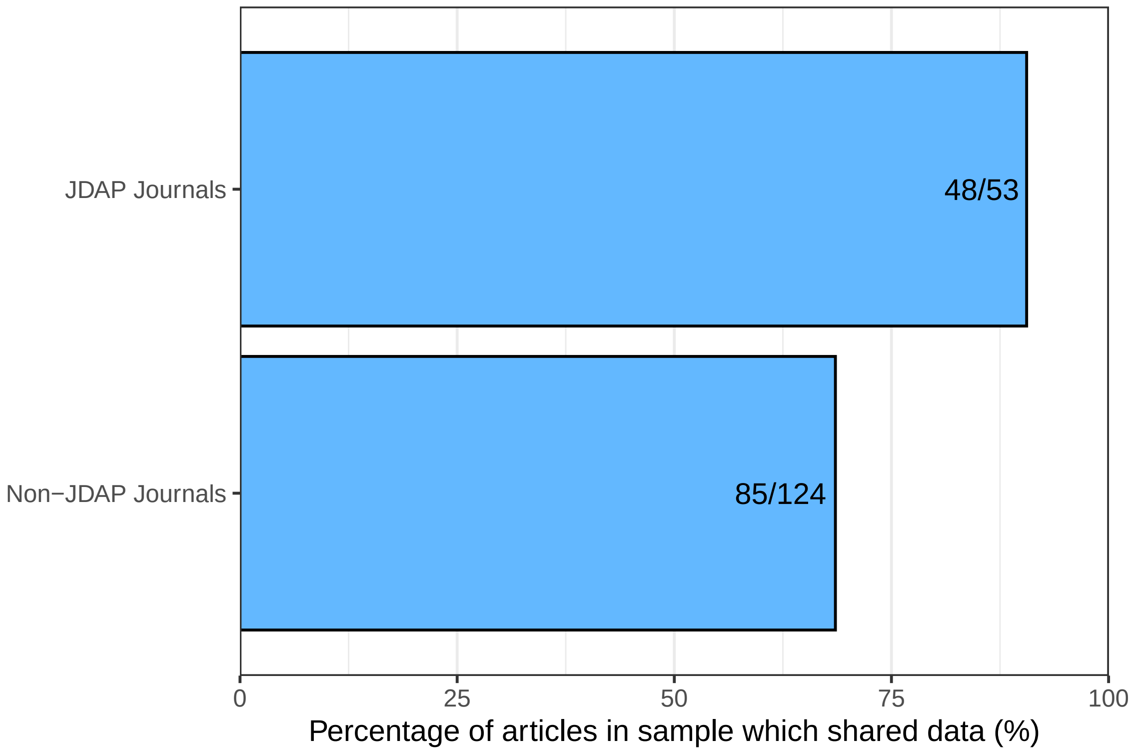

Figure S2: Comparison of data sharing rates in articles by journal JDAP membership status.

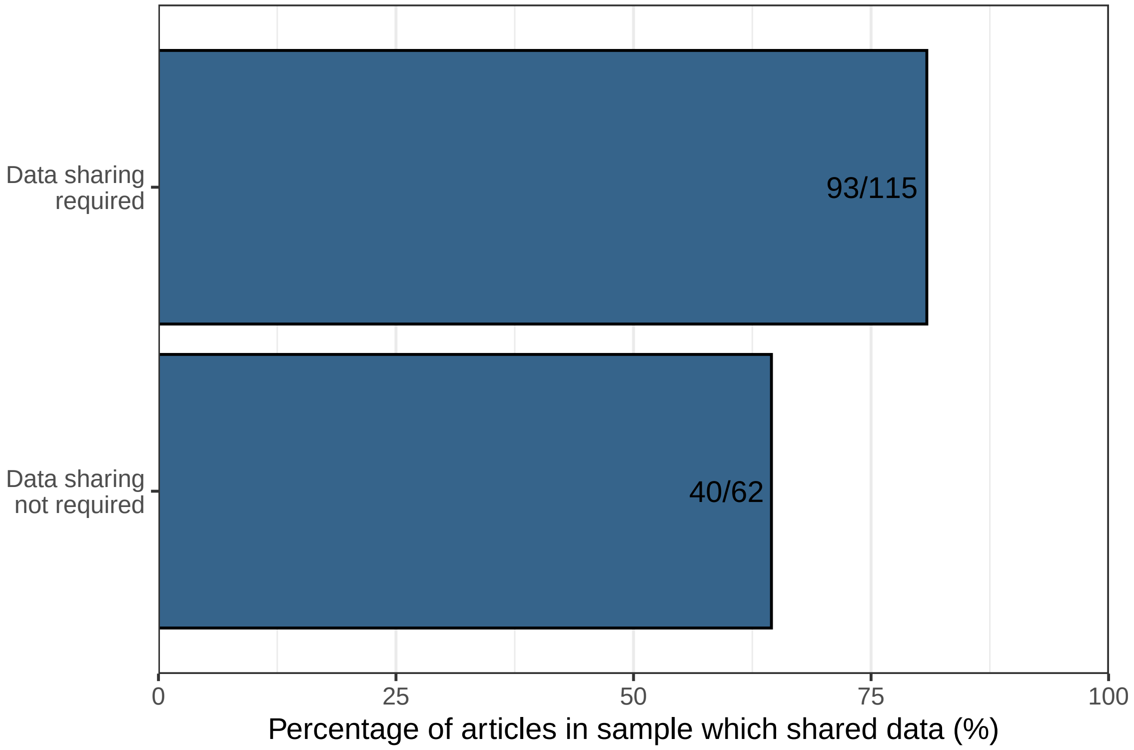

Figure S3: Comparison of data sharing rates in articles by journal data policy 2015–17.

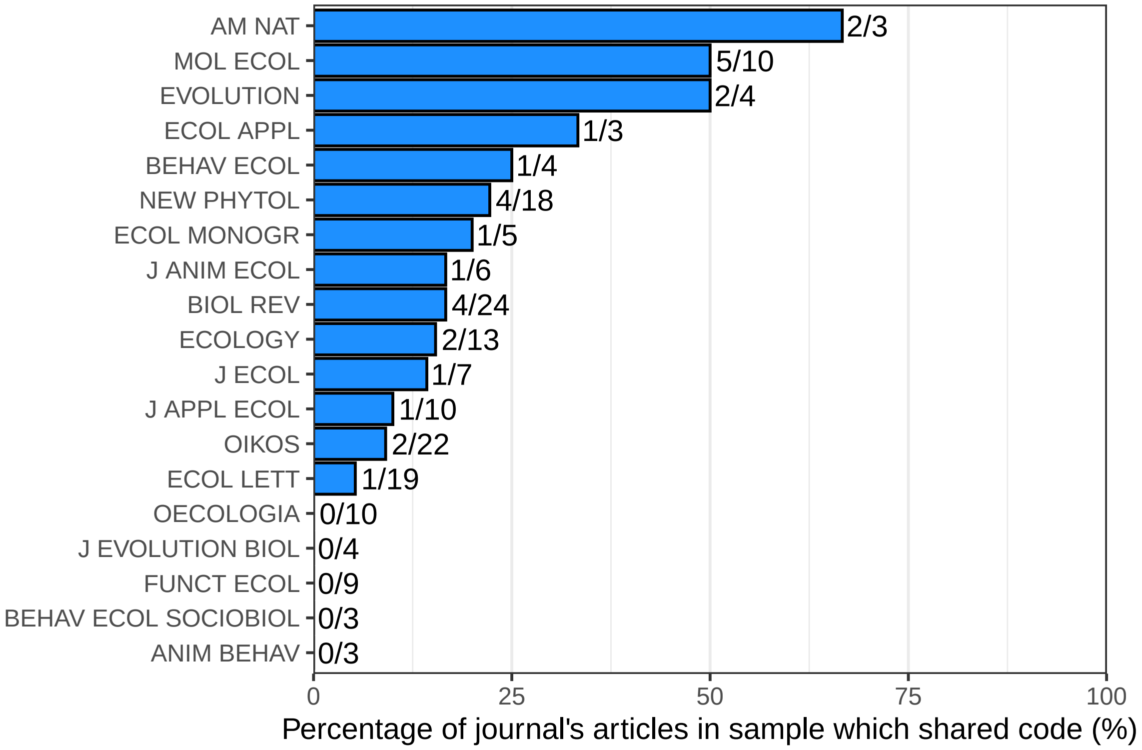

Figure S4: Comparison of code sharing rates in articles by journal.

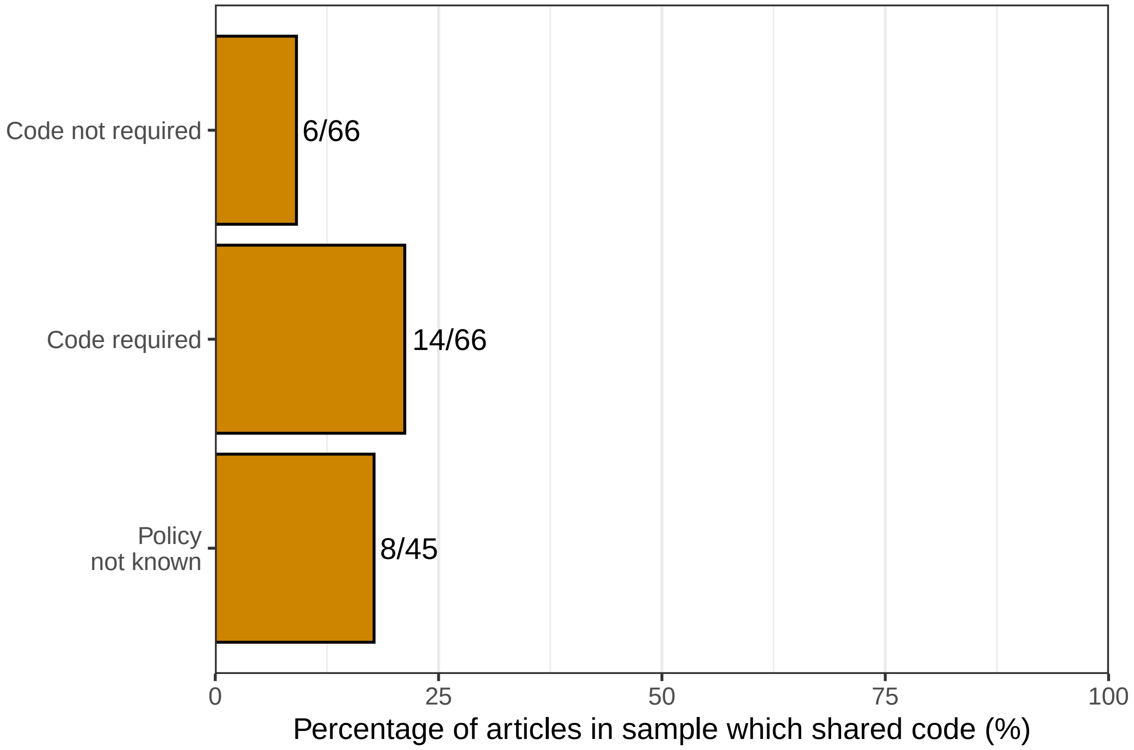

Figure S5: Comparison of code sharing rates in articles by journal code policy according to (33).

S7 SOFTWARE MENTIONED IN ARTICLES

Figure S6 shows the distribution of the number of different software packages mentioned in each article (or in its supplementary material). Here, R packages have been treated as special cases: articles mentioning multiple R packages have been treated as just mentioning the R software environment. For example, an article which mentioned the R software environment and four R packages was regarded as mentioning one software package (the R software environment) rather than five software packages.

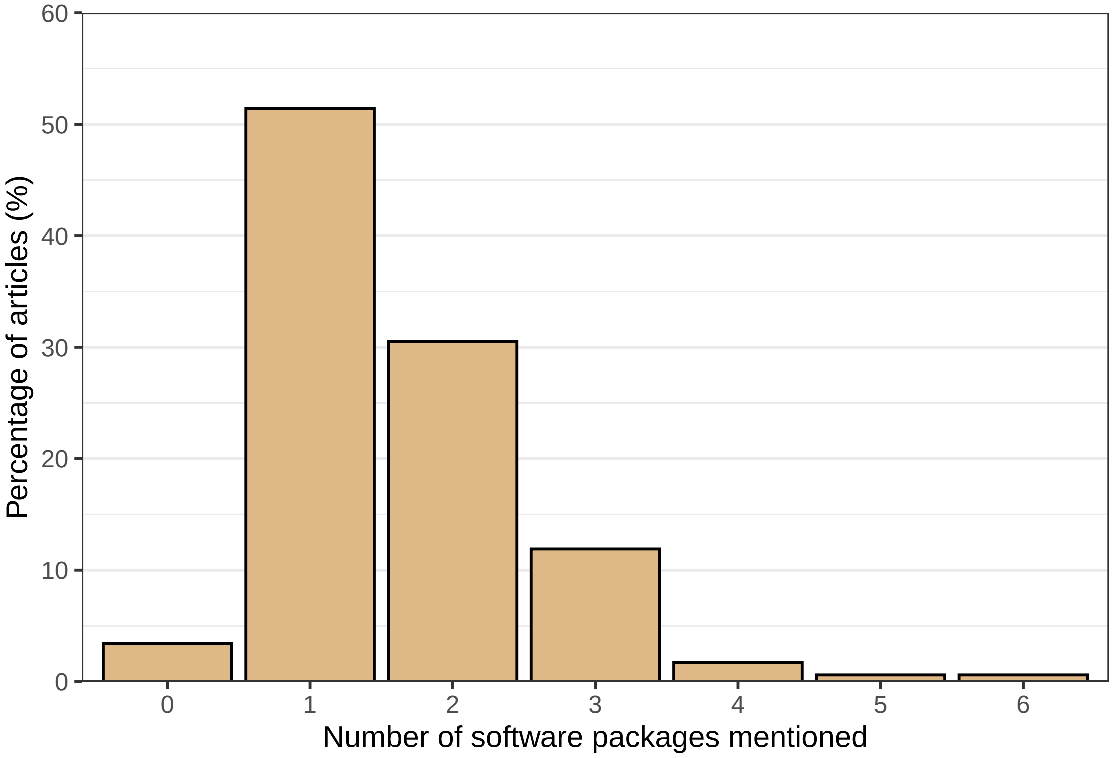

Figure S6: Distribution of the number of different software packages mentioned in each article (or its supplementary material).

Figure S7 is a version of Figure S6 which shows the distribution of the number of different software packages mentioned in each article (or in its supplementary material), *including* mentions of R packages. For example, an article which mentioned the R software environment and four R packages was regarded as mentioning five software packages rather than one software package (i.e., the R software in general).

Table S7 lists all software packages mentioned in the 177 meta-analysis articles. This table includes all mentions of the R software environment, but specifically excludes mentions of R packages, which are listed in the following table.

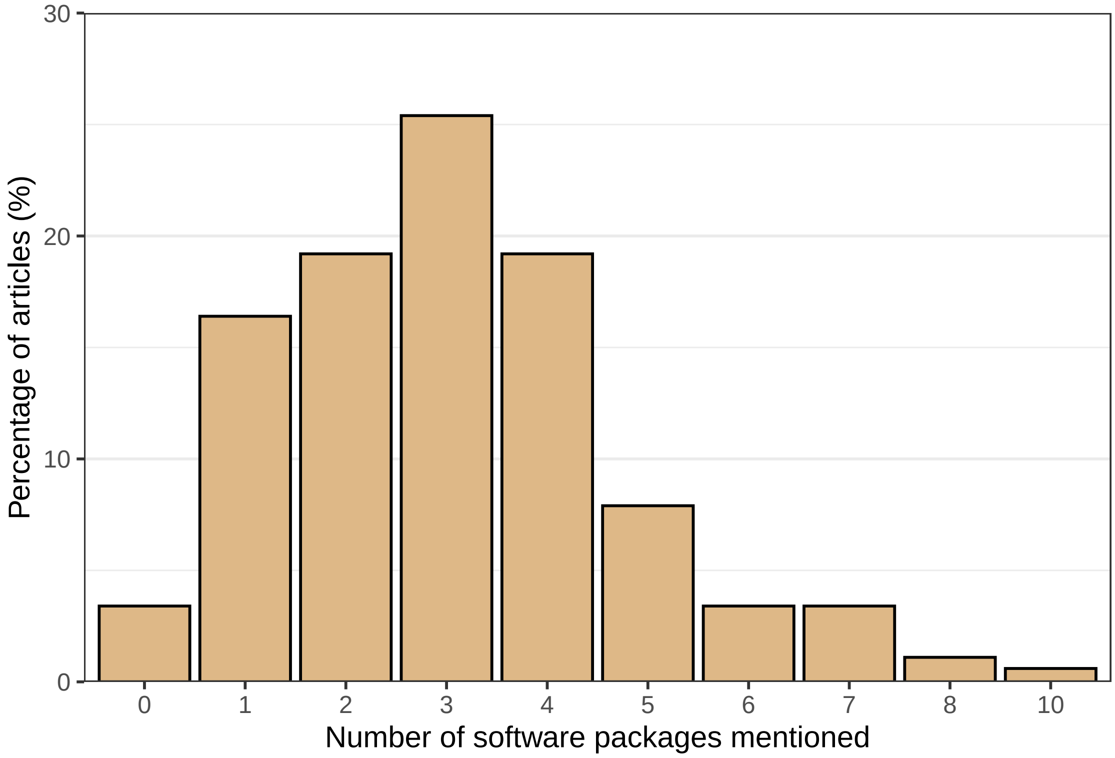

Figure S7: Distribution of the number of different software packages mentioned in each article (or its supplementary material), including mentions of R packages.

Table S7: All software packages mentioned in the 177 meta-analysis articles. Note that this table does not list individual R packages.

Name of software package	N	%
R	141	79.7
MetaWin	20	11.3
WebPlotDigitizer	10	5.6
DataThief III	9	5.1
SAS	9	5.1
ImageJ	7	4.0
GraphClick	6	3.4
PlotDigitizer	5	2.8
GetData Graph Digitizer	4	2.3
JMP	4	2.3
RStudio	4	2.3
SPSS/PASW	4	2.3
Minitab	3	1.7

Name of software package	N	%
Phylocom	3	1.7
Phylomatic	3	1.7
RAxML	3	1.7
ArcGIS	2	1.1
ArcMap	2	1.1
CMA	2	1.1
Engauge Digitizer	2	1.1
GENALEX	2	1.1
MAFFT	2	1.1
Python	2	1.1
Stan	2	1.1
AbstrackR	1	0.6
ADZE	1	0.6
AMOS	1	0.6
ARLSUMSTAT	1	0.6
ASReml-R	1	0.6
Bowtie2	1	0.6
Cervus	1	0.6
Circuitscape	1	0.6
Cytoscape	1	0.6
Digitize It 2010	1	0.6
Ecopath	1	0.6
ED2 (FORTRAN)	1	0.6
Excel	1	0.6
FigTree	1	0.6
GenClone	1	0.6
GrabIt! XP	1	0.6
GRASS GIS	1	0.6
Image Pro Plus	1	0.6
JAGS	1	0.6

Name of software package	N	%
LocARNA	1	0.6
MEGA 4	1	0.6
Mesquite	1	0.6
Modeltest	1	0.6
MrBayes	1	0.6
OpenBUGS	1	0.6
OriginPro	1	0.6
Perl	1	0.6
Photoshop	1	0.6
phyloMeta	1	0.6
PRIMER	1	0.6
QGIS	1	0.6
SigmaPlot	1	0.6
Techdig	1	0.6
xyscan	1	0.6

In total, there were 398 mentions of R and R packages across the articles: 141 mentions of the R software environment, and 257 mentions of specific R packages. Figure S8 shows the distribution of the number of packages mentioned by each R-using article. As the figure shows, it was most common for R-using articles to mention only one or two packages (68%); only 6% of R-using articles mentioned more than three R packages.

Table S8 lists all R packages mentioned in the 141 meta-analysis articles that mentioned using R. The table includes the location of each R package (whether CRAN, Bioconductor, a base R package, or from another source). Note: At the time of checking (2nd August 2022), four packages (*empiricalFDR.DESeq2*, *foodweb*, *MAc*, and *VIF*) have been removed from CRAN[3]. The vast majority (74, or 92%) of the mentioned R packages were from the Comprehensive R Archive Network (CRAN), with 3 (4%) from the Bioconductor project and 2 from other websites. One article mentioned the package *stats*, which is part of the "base" set of R packages that are an integral part of the R software.

---

[3] Package *empiricalFDR.DESeq2* was archived 13th June 2022 (`https://cran.r-project.org/package=empiricalFDR.DESeq2`); package *foodweb* was archived 21st June 2022 (`https://cran.r-project.org/package=foodweb`); package *MAc* was archived 4th March 2022 (`https://cran.r-project.org/package=MAc`); package *VIF* was archived 9th May 2022 (`https://cran.r-project.org/package=VIF`).

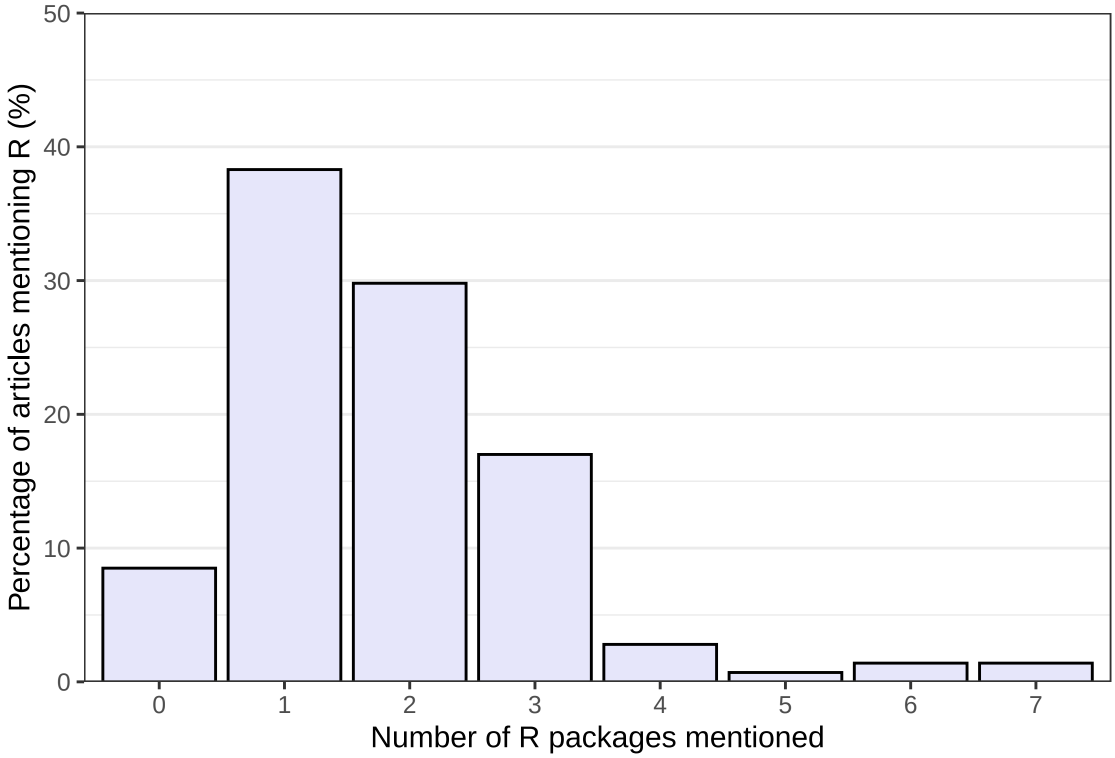

Figure S8: Distribution of the number of different R packages mentioned in each article which mentioned using R.

Table S8: All R packages mentioned in the sample of 141 meta-analysis articles which mentioned using R.

Name of R package	Package source	N	%
metafor	CRAN	75	53.2
MCMCglmm	CRAN	26	18.4
lme4	CRAN	20	14.2
ape	CRAN	13	9.2
MuMIn	CRAN	8	5.7
vegan	CRAN	7	5.0
nlme	CRAN	6	4.3
ggplot2	CRAN	5	3.5
phytools	CRAN	5	3.5
compute.es	CRAN	4	2.8
glmulti	CRAN	4	2.8
multcomp	CRAN	3	2.1
raster	CRAN	3	2.1
ade4	CRAN	2	1.4

Name of R package	Package source	N	%
boot	CRAN	2	1.4
lmerTest	CRAN	2	1.4
meta	CRAN	2	1.4
mgcv	CRAN	2	1.4
mice	CRAN	2	1.4
party	CRAN	2	1.4
picante	CRAN	2	1.4
randomForest	CRAN	2	1.4
rjags	CRAN	2	1.4
rmeta	CRAN	2	1.4
A3	CRAN	1	0.7
abc	CRAN	1	0.7
adegenet	CRAN	1	0.7
AICcmodavg	CRAN	1	0.7
arrayQualityMetrics	BioConductor	1	0.7
betareg	CRAN	1	0.7
caper	CRAN	1	0.7
coda	CRAN	1	0.7
coin	CRAN	1	0.7
DESeq2	BioConductor	1	0.7
dismo	CRAN	1	0.7
ecodist	CRAN	1	0.7
effects	CRAN	1	0.7
empiricalFDR.DESeq2	CRAN	1	0.7
foodweb	CRAN	1	0.7
gbm	CRAN	1	0.7
GENHET	other	1	0.7
Hmisc	CRAN	1	0.7
ICC	CRAN	1	0.7
igraph	CRAN	1	0.7

Name of R package	Package source	N	%
inext	CRAN	1	0.7
Kendall	CRAN	1	0.7
KOGMWU	CRAN	1	0.7
languageR	CRAN	1	0.7
leaps	CRAN	1	0.7
lmodel2	CRAN	1	0.7
lsmeans	CRAN	1	0.7
MAc	CRAN	1	0.7
maps	CRAN	1	0.7
maptools	CRAN	1	0.7
MASS	CRAN	1	0.7
merTools	CRAN	1	0.7
metahdep	BioConductor	1	0.7
MODISTools	CRAN	1	0.7
pez	CRAN	1	0.7
pheatmap	CRAN	1	0.7
plotmcmc	CRAN	1	0.7
plyr	CRAN	1	0.7
PVR	CRAN	1	0.7
R2WinBUGS	CRAN	1	0.7
rfPermute	CRAN	1	0.7
rgdal	CRAN	1	0.7
RInSp	CRAN	1	0.7
rms	CRAN	1	0.7
rotl	CRAN	1	0.7
rstan	CRAN	1	0.7
rvest	CRAN	1	0.7
segmented	CRAN	1	0.7
shape	CRAN	1	0.7
smatr	CRAN	1	0.7

Name of R package	Package source	N	%
STANDARICH	other	1	0.7
stats	base	1	0.7
vif	CRAN	1	0.7
visreg	CRAN	1	0.7
weights	CRAN	1	0.7
zoo	CRAN	1	0.7

Table S9 shows all R versions mentioned in the articles, as they originally appeared in the articles. This includes one article where the authors mention using two different versions of R for their study (v2.14.1 and v3.0.0), a study which included the R version twice, first in the body of the text and second as part of the citation in the references section, but where the versions differed (v3.1.0 and v3.0.1, which might be due to a typing error), and six articles where the version information provided was not complete (v2.12, v2.13, v2.14, v2.15, v3.1, v3.2). In the case of the six incomplete R version statements, it is possible that the authors were intending to refer to the "0" versions, i.e., 2.12.0, 2.13.0, etc.

R version	N	%
2.12	1	0.7
2.13	1	0.7
2.14	1	0.7
2.14.1	4	2.8
2.15	1	0.7
2.15.2	3	2.1
3.0.0/2.14.1	1	0.7
3.0.1	8	5.7
3.0.2	14	9.9
3.0.3	3	2.1
3.1	2	1.4
3.1.0	2	1.4
3.1.0/3.0.1	1	0.7
3.1.1	4	2.8
3.1.2	12	8.5
3.1.3	3	2.1
3.2	1	0.7
3.2.0	1	0.7
3.2.1	6	4.3
3.2.2	6	4.3
3.2.3	6	4.3
3.2.4	1	0.7
3.3.0	2	1.4
3.3.1	1	0.7
3.3.2	1	0.7
3.4.0	1	0.7
3.4.1	1	0.7
(No version mentioned)	53	37.6

Table S9: All R versions as originally mentioned in the sample of 141 meta-analysis articles which mentioned using R.

Variable	Value
ID	MA092
Study	(60)
Result source	in text and from Table 1 (pp.84-85)
Result type	Regression model results for all data
Regression result	$\text{TLP} = -4.67 + 0.725 \times \log(\text{SLA}) - 0.937 \times \log(\text{WD})$
N	68
R^2_{adj}	0.32 (p-value < 0.001)
RMSE	0.55

Table S10: Details of the target result for article MA092, (60). For context, TLP – turgor loss point, SLA – specific leaf area, WD – wood density, RMSE – root mean square error.

Variable	Value
ID	MA094
Study	(61)
Result source	in text (p.1227)
Result type	Ordination analysis result
N	n.s.
R^2	0.494 ($p < 0.0001$)

Table S11: Details of the target result for article MA094, (61). n.s. – not stated.

Variable	Value
ID	MA129
Study	(62)
Result source	Table 1 (A) (p.444)
Result type	Comparison of models by AICc
Result values	See Table S13

Table S12: Summary of the target result for article MA129, (62).

Buffer	Δ_i	w_i	R^2
Null	0.00	0.20	
50	0.96	0.12	
25	0.98	0.12	
75	1.17	0.11	
200	1.34	0.10	
150	1.56	0.09	
10	1.56	0.09	
100	1.61	0.09	
5	1.70	0.08	

Table S13: Details of the target result for article MA129, (62). The table headings and values are taken directly from Table 1 (A), p.444. Note that blank/missing values in the R^2 column are as per the original table. Here, Buffer is radius in km, Δ_i is AICc$_i$ − minimum AICc (where AICc is the corrected Akaike information criterion), w_i is Akaike weight, R^2 is coefficient of determination, omitted in this table.

Variable	Value
ID	MA212
Study	(63)
Result source	Table 2 (p.38)
Result type	Counts of matches
Result values	See Table S15

Table S14: Summary of the target result for article MA212, (63).

KS index	Match impact	Match biomass	No match	Overall match
KS_1	91	10	0	match impact
KS_2	5	81	15	match biomass
KS_3	50	28	23	
KS_4	25	54	22	match biomass
KS_5	86	12	3	match impact
KS_6	0	94	7	match biomass
KS_7	32	35	34	
KS_8	11	70	20	match biomass
KS_9	91	10	0	match impact
KS_{10}	25	54	22	match biomass
KS_{11}	71	20	10	match impact
KS_{12}	46	39	16	

Table S15: Details of the target result for article MA212, (63). The table headings and values are taken directly from Table 2, p.38. Note that blank/missing values in rows 3, 7, and 12 of column "Overall match" are as per the original table.

S9 REPRODUCIBILITY REPORTS

Reproducibility report design

We decided to record all the steps of each reproduction attempt in a document integrating the running of analysis code with explanatory prose to contextualise the attempt and its outcome. This an attempt to follow the literate programming approach (245), which emphasises that computer programs ought to be human-readable and understandable. It's been recognised that this integration of analysis code and word processing facilitates reproducibility (246; 247), especially when the document is packaged with the data files required for the analysis into a compendium (248). As will be described in section 3 of the results, nearly all code that was shared was code for the R language (249). We mention this here because this fact determined the specifics of our technical approach to constructing the reproducibility reports.

We wrote a reproducibility report template using RMarkdown, a format for reproducible documents in the R language. An RMarkdown file can be compiled to produce a formatted, human-readable output document (such as an HTML or PDF document), which reports the results of running all included R code. The R source code in the RMarkdown document is re-run each time the document is compiled. We structured the template similarly to the RMarkdown reproducibility reports used in (17) to reproduce results from articles published in *Psychological Science*. Each report was structured as follows:

- A reference to the article and numerical details of the target result to be reproduced;
- Details of the shared data and code files;
- As assessment of the applicability of the shared data and code files;
- Set up of the R environment as required for the analyses;
- Importing and cleaning of data;
- Running the analysis code to reproduce the target result;
- Comparison of the original and reproduced target result value(s);
- A summary of information about the R computational environment used.

Within the RMarkdown source file, each report section consists of a combination of text marked up for appropriate formatting and "chunks" of R code which, when executed, perform in order the relevant tasks for the analysis (e.g., importing data from a file).

We set up each reproducibility report to run within its own Docker container. A container is a structured package of software designed to run a particular application in a virtual computing environment. The advantage of this approach is that applications can run on different computers without users needing to deal with software or system dependencies or settings. Docker is a tool for creating and running containers (250; 251). In particular, Docker allows users to build upon existing containers in an easy way. We created a container for each reproducibility report by starting with a pre-built container running R maintained by the Rocker project (252). The Rocker container already included all elements required to run an R session in an isolated computational environment. On top of this pre-built "layer" we built containers which installed all additional R packages required for the analyses in the reproducibility reports, including custom functions written by us to facilitate comparison of the original and reproduced values. We controlled the versions of both R and all R packages: the Rocker project maintains multiple containers with different versions of R; we selected version 3.5.0. We installed R packages from a snapshot of the Comprehensive R Archive Network (CRAN) frozen at 2 July 2018, to ensure compatibility with R 3.5.0[4] The final layer of the container for each reproducibility report incorporated the specific data and code files required for data analysis. The result of this work was a small, self-contained application with everything required to compile the reproducibility report for each of the articles with shared data and code. This is a variation on the "research compendium" (253): a research compendium is usually envisioned as being created by the original authors of a research project, to facilitate the reproducibility of their own results, rather than being created by a third party after the fact.

Running code

The core of each reproducibility report was the section which conducted the data analysis and calculated the target result. Because each reproducibility report is fundamentally an assessment of the shared data and code, we envisioned that each report would by default only execute lines of code taken directly from the shared code file(s) except where unavoidable. Importing data files was the principle situation where we anticipated we would need to modify lines of code and/or write new code.[5] In order to differentiate original lines of code from additional lines of code written by us, we wrote a function to specify which particular lines of an external code file to execute. This way, the original shared code could be run by a call to a function, rather than needing to be manually inserted into the source of the RMarkdown report. All additional, custom

---

[4] The date 2 July 2018 is the last day before the release of the succeeding version of R. We used the Microsoft mirror of CRAN: https://cran.microsoft.com/snapshot/2018-07-02/.

[5] We anticipated that most if not all shared code concerning operations involving external files would require modification. This was due to the fact that at the very least, file paths to data files, etc. would need to be changed to match the file system structure set up within each Docker container.

code would be written directly into the RMarkdown source file. Running individual lines of code from the original files in this way also had the advantage that only the code that was required to calculate the target result could be run, rather than the entire code file. For analyses that involved random number generation, we set an arbitrary random seed so that the specific set of numbers calculated would be reproduced over successive compilations of the report.

Table S16: The original and reproduced values of all target results.

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA016	(38)	correlation	Pearson's *r*	point est.	N	-0.83	-0.83	0.00	E
MA016	(38)	correlation	Pearson's *r*	*p*-value	N	< 0.001	< 0.001		E
MA016	(38)	correlation	Pearson's *r*	*N*	N	49	49	0.00	E
MA060	(39)	mean	Fisher *z*-transformation	point est.	N	0.044	0.043	2.27	R
MA060	(39)	mean	Fisher *z*-transformation	HPDI lower	N	-0.174	-0.194	11.49	10%+
MA060	(39)	mean	Fisher *z*-transformation	HPDI upper	N	0.289	0.268	7.27	< 10%
MA060	(39)	mean	Fisher *z*-transformation	*N*	N	37	37	0.00	E
MA062	(40)	mean	Hedges' *d*	point est.	N	-0.205	-0.204	0.49	R
MA062	(40)	mean	Hedges' *d*	CI lower	N	-0.444	-0.446	0.45	< 10%
MA062	(40)	mean	Hedges' *d*	CI upper	N	0.035	0.039	11.43	10%+
MA062	(40)	mean	Hedges' *d*	*N*	N	37	37	0.00	E
MA065	(41)	mean	Hedges' *g*	point est.	N	-8.42	-8.87	5.34	< 10%
MA065	(41)	mean	Hedges' *g*	CI lower	N	-10.73	-10.85	1.12	< 10%
MA065	(41)	mean	Hedges' *g*	CI upper	N	-6.63	-6.68	0.75	< 10%
MA065	(41)	mean	Hedges' *g*	*N*	N	703	703	0.00	E
MA067	(42)	mean	Hedges' *g*	point est.	N	-0.21	-0.21	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA067	(42)	mean	Hedges' g	SE	N	0.07	0.07	0.00	E
MA067	(42)	mean	Hedges' g	z-score	N	-2.7	-2.8	3.70	R
MA067	(42)	mean	Hedges' g	p-value	N	0.006	0.005	16.67	R
MA067	(42)	mean	Hedges' g	N	N	52	52	0.00	E
MA068	(43)	mean	odds ratio	point est.	N	1.82			F
MA068	(43)	mean	odds ratio	HPDI lower	N	1.37			F
MA068	(43)	mean	odds ratio	HPDI upper	N	2.41			F
MA068	(43)	mean	odds ratio	N	N	75			F
MA071	(44)	mean	response ratio	point est.	N	-0.26	-0.27	3.85	R
MA071	(44)	mean	response ratio	CI lower	N	-1.02	-1.03	0.98	R
MA071	(44)	mean	response ratio	CI upper	N	0.51	0.49	3.92	< 10%
MA071	(44)	mean	response ratio	N	N	50	50	0.00	E
MA074	(45)	correlation	Pearson's r	point est.	N	0.183	0.185	1.09	< 10%
MA074	(45)	correlation	Pearson's r	CI lower	N	0.089	0.089	0.00	E
MA074	(45)	correlation	Pearson's r	CI upper	N	0.274	0.281	2.55	< 10%
MA074	(45)	correlation	Pearson's r	N	N	43	43	0.00	E
MA081	(46)	mean	slope parameter	point est.	N	1.30	1.30	0.00	E
MA081	(46)	mean	slope parameter	CI lower	N	0.95			F
MA081	(46)	mean	slope parameter	CI upper	N	1.66			F
MA081	(46)	mean	slope parameter	N	N	1296	1296	0.00	E
MA091	(47)	mean	Cohen's d	point est.	N	0.56	0.56	0.00	E
MA091	(47)	mean	Cohen's d	CI lower	N	0.42	0.42	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA091	(47)	mean	Cohen's *d*	CI upper	N	0.69	0.69	0.00	E
MA091	(47)	mean	Cohen's *d*	N	N	65	65	0.00	E
MA092	(60)	model output	n.a.	R^2_{adj}	N	0.32	0.33	3.13	R
MA092	(60)	model output	n.a.	RMSE	N	0.55	0.55	0.00	E
MA092	(60)	model output	n.a.	intercept	N	-4.67	-4.18	10.49	10%+
MA092	(60)	model output	n.a.	log(SLA) coeff.	N	0.725	0.730	0.69	< 10%
MA092	(60)	model output	n.a.	log(WD) coeff.	N	-0.937	-0.980	4.59	< 10%
MA092	(60)	model output	n.a.	N	N	68	68	0.00	E
MA094	(61)	model output	n.a.	R^2	N	0.494			F
MA094	(61)	model output	n.a.	p-value	N	< 0.0001			F
MA095	(48)	mean	Fisher *z*-transformation	point est.	N	0.76	0.76	0.00	E
MA095	(48)	mean	Fisher *z*-transformation	CI lower	N	0.61	0.61	0.00	E
MA095	(48)	mean	Fisher *z*-transformation	CI upper	N	0.91	0.91	0.00	E
MA095	(48)	mean	Fisher *z*-transformation	N	N	25	25	0.00	E
MA126	(49)	mean	log odds ratio	point est.	N	-1.11	-1.11	0.00	E
MA126	(49)	mean	log odds ratio	SE	N	0.49	0.49	0.00	E
MA126	(49)	mean	log odds ratio	CI lower	N	-2.06	-2.06	0.00	E
MA126	(49)	mean	log odds ratio	CI upper	N	-0.15	-0.15	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA126	(49)	mean	log odds ratio	z-score	N	-2.28	-2.28	0.00	E
MA126	(49)	mean	log odds ratio	p-value	N	0.023	0.023	0.00	E
MA129	(62)	Table 1A, rank1	n.a.	Δ_i	N	0.00	0.00		E
MA129	(62)	Table 1A, rank1	n.a.	w_i	N	0.20	0.20	0.00	E
MA129	(62)	Table 1A, rank1	n.a.	buffer (km radius)	C	Null	Null		E
MA129	(62)	Table 1A, rank2	n.a.	Δ_i	N	0.96	0.96	0.00	E
MA129	(62)	Table 1A, rank2	n.a.	w_i	N	0.12	0.12	0.00	E
MA129	(62)	Table 1A, rank2	n.a.	buffer (km radius)	C	50	50		E
MA129	(62)	Table 1A, rank3	n.a.	Δ_i	N	0.98	0.98	0.00	E
MA129	(62)	Table 1A, rank3	n.a.	w_i	N	0.12	0.12	0.00	E
MA129	(62)	Table 1A, rank3	n.a.	buffer (km radius)	C	25	25		E
MA129	(62)	Table 1A, rank4	n.a.	Δ_i	N	1.17	1.17	0.00	E
MA129	(62)	Table 1A, rank4	n.a.	w_i	N	0.11	0.11	0.00	E
MA129	(62)	Table 1A, rank4	n.a.	buffer (km radius)	C	75	75		E
MA129	(62)	Table 1A, rank5	n.a.	Δ_i	N	1.34	1.34	0.00	E
MA129	(62)	Table 1A, rank5	n.a.	w_i	N	0.10	0.10	0.00	E
MA129	(62)	Table 1A, rank5	n.a.	buffer (km radius)	C	200	200		E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA129	(62)	Table 1A, rank6	n.a.	Δ_i	N	1.56	1.56	0.00	E
MA129	(62)	Table 1A, rank6	n.a.	w_i	N	0.09	0.09	0.00	E
MA129	(62)	Table 1A, rank6	n.a.	buffer (km radius)	C	150	150		E
MA129	(62)	Table 1A, rank7	n.a.	Δ_i	N	1.56	1.56	0.00	E
MA129	(62)	Table 1A, rank7	n.a.	w_i	N	0.09	0.09	0.00	E
MA129	(62)	Table 1A, rank7	n.a.	buffer (km radius)	C	10	10		E
MA129	(62)	Table 1A, rank8	n.a.	Δ_i	N	1.61	1.61	0.00	E
MA129	(62)	Table 1A, rank8	n.a.	w_i	N	0.09	0.09	0.00	E
MA129	(62)	Table 1A, rank8	n.a.	buffer (km radius)	C	100	100		E
MA129	(62)	Table 1A, rank9	n.a.	Δ_i	N	1.70	1.70	0.00	E
MA129	(62)	Table 1A, rank9	n.a.	w_i	N	0.08	0.08	0.00	E
MA129	(62)	Table 1A, rank9	n.a.	buffer (km radius)	C	5	5		E
MA145	(50)	mean	Fisher z-transformation	point est.	N	-0.08	-0.08	0.00	E
MA145	(50)	mean	Fisher z-transformation	HPDI lower	N	-0.22	-0.21	4.55	R
MA145	(50)	mean	Fisher z-transformation	HPDI upper	N	0.03	0.05	66.67	10%+

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA145	(50)	mean	Fisher *z*-transformation	N	N	118	118	0.00	E
MA145	(50)	mean	Fisher *z*-transformation	N_{studies}	N	38	38	0.00	E
MA145	(50)	mean	Fisher *z*-transformation	N_{species}	N	25	25	0.00	E
MA147	(51)	mean	percentage	point est.	N	0.13	0.13	0.00	E
MA147	(51)	mean	percentage	SE	N	0.03	0.03	0.00	E
MA147	(51)	mean	percentage	CI lower	N	0.074	0.074	0.00	E
MA147	(51)	mean	percentage	CI upper	N	0.19	0.19	0.00	E
MA147	(51)	mean	percentage	N	N	49	49	0.00	E
MA155	(52)	correlation	Pearson's r	point est.	N	0.51	0.51	0.00	E
MA155	(52)	correlation	Pearson's r	p-value	N	0.01	0.01	0.00	E
MA188	(53)	mean	Log response ratio	point est.	N	-0.363	-0.363	0.00	E
MA188	(53)	mean	Log response ratio	CI lower	N	-0.408	-0.408	0.00	E
MA188	(53)	mean	Log response ratio	CI upper	N	-0.318	-0.318	0.00	E
MA188	(53)	mean	Log response ratio	N	N	818	818	0.00	E
MA191	(54)	mean	allometric slope parameter	point est.	N	0.86	0.85	1.16	R
MA191	(54)	mean	allometric slope parameter	CI lower	N	0.77	0.77	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA191	(54)	mean	allometric slope parameter	CI upper	N	0.94	0.94	0.00	E
MA191	(54)	mean	allometric slope parameter	*N*	N	553	553	0.00	E
MA198	(55)	mean	Fisher *z*-transformation	point est.	N	-0.41	-0.42	2.44	R
MA198	(55)	mean	Fisher *z*-transformation	CI lower	N	-0.55	-0.55	0.00	E
MA198	(55)	mean	Fisher *z*-transformation	CI upper	N	-0.27	-0.28	3.70	R
MA198	(55)	mean	Fisher *z*-transformation	*N*	N	79	80	1.27	< 10%
MA202	(56)	mean	Hedges' *d*	point est.	N	-0.330	-0.340	3.03	< 10%
MA202	(56)	mean	Hedges' *d*	CI lower	N	-0.503	-0.521	3.58	< 10%
MA202	(56)	mean	Hedges' *d*	CI upper	N	-0.156	-0.159	1.92	< 10%
MA202	(56)	mean	Hedges' *d*	*N*	N	329	329	0.00	E
MA211	(57)	mean	log response ratio	point est.	N	0.24			F
MA211	(57)	mean	log response ratio	CI lower	N	0.23			F
MA211	(57)	mean	log response ratio	CI upper	N	0.25			F
MA211	(57)	mean	log response ratio	*N*	N	3298			F
MA212	(63)	Table 2, Match biomass	n.a.	KS1	N	10	10	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA212	(63)	Table 2, Match biomass	n.a.	KS2	N	81	81	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS3	N	28	28	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS4	N	54	54	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS5	N	12	12	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS6	N	94	94	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS7	N	35	35	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS8	N	70	70	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS9	N	10	10	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS10	N	54	54	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS11	N	20	20	0.00	E
MA212	(63)	Table 2, Match biomass	n.a.	KS12	N	39	39	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA212	(63)	Table 2, Match impact	n.a.	KS1	N	91	91	0.00	E
MA212	(63)	Table 2, Match impact	n.a.	KS2	N	5	5	0.00	E
MA212	(63)	Table 2, Match impact	n.a.	KS3	N	50	51	2.00	< 10%
MA212	(63)	Table 2, Match impact	n.a.	KS4	N	25	25	0.00	E
MA212	(63)	Table 2, Match impact	n.a.	KS5	N	86	87	1.16	< 10%
MA212	(63)	Table 2, Match impact	n.a.	KS6	N	0	0		E
MA212	(63)	Table 2, Match impact	n.a.	KS7	N	32	33	3.12	< 10%
MA212	(63)	Table 2, Match impact	n.a.	KS8	N	11	11	0.00	E
MA212	(63)	Table 2, Match impact	n.a.	KS9	N	91	91	0.00	E
MA212	(63)	Table 2, Match impact	n.a.	KS10	N	25	25	0.00	E
MA212	(63)	Table 2, Match impact	n.a.	KS11	N	71	72	1.41	< 10%

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA212	(63)	Table 2, Match impact	n.a.	KS12	N	46	47	2.17	< 10%
MA212	(63)	Table 2, No match	n.a.	KS1	N	0	0		E
MA212	(63)	Table 2, No match	n.a.	KS2	N	15	15	0.00	E
MA212	(63)	Table 2, No match	n.a.	KS3	N	23	22	4.35	< 10%
MA212	(63)	Table 2, No match	n.a.	KS4	N	22	22	0.00	E
MA212	(63)	Table 2, No match	n.a.	KS5	N	3	2	33.33	10%+
MA212	(63)	Table 2, No match	n.a.	KS6	N	7	7	0.00	E
MA212	(63)	Table 2, No match	n.a.	KS7	N	34	33	2.94	< 10%
MA212	(63)	Table 2, No match	n.a.	KS8	N	20	20	0.00	E
MA212	(63)	Table 2, No match	n.a.	KS9	N	0	0		E
MA212	(63)	Table 2, No match	n.a.	KS10	N	22	22	0.00	E
MA212	(63)	Table 2, No match	n.a.	KS11	N	10	9	10.00	10%+
MA212	(63)	Table 2, No match	n.a.	KS12	N	16	15	6.25	< 10%
MA212	(63)	Table 2, Overall match	n.a.	KS1	C	match impact	match impact		E
MA212	(63)	Table 2, Overall match	n.a.	KS2	C	match biomass	match biomass		E
MA212	(63)	Table 2, Overall match	n.a.	KS3	C	(none)	match impact		NC
MA212	(63)	Table 2, Overall match	n.a.	KS4	C	match biomass	match biomass		E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA212	(63)	Table 2, Overall match	n.a.	KS5	C	match impact	match impact		E
MA212	(63)	Table 2, Overall match	n.a.	KS6	C	match biomass	match biomass		E
MA212	(63)	Table 2, Overall match	n.a.	KS7	C	(none)	(none)		E
MA212	(63)	Table 2, Overall match	n.a.	KS8	C	match biomass	match biomass		E
MA212	(63)	Table 2, Overall match	n.a.	KS9	C	match impact	match impact		E
MA212	(63)	Table 2, Overall match	n.a.	KS10	C	match biomass	match biomass		E
MA212	(63)	Table 2, Overall match	n.a.	KS11	C	match impact	match impact		E
MA212	(63)	Table 2, Overall match	n.a.	KS12	C	(none)	(none)		E
MA213	(58)	mean	difference in means	point est.	N	-0.07	-0.07	0.00	E
MA213	(58)	mean	difference in means	*p*-value	N	0.362	0.362	0.00	E
MA213	(58)	mean	difference in means	*N*	N	654	654	0.00	E
MA229	(59)	mean	log response ratio	point est.	N	0.40	0.39	2.50	R
MA229	(59)	mean	log response ratio	CI lower	N	0.24	0.26	8.33	< 10%
MA229	(59)	mean	log response ratio	CI upper	N	0.53	0.53	0.00	E

ID	Study	Result type	Effect size type	Target result	Value type	Original	Reproduced	Percent error (%)	Status
MA229	(59)	mean	log response ratio	*N*	N	57	57	0.00	E

Examining dependency between reproduced values within articles

Table 5 lists 19 articles where (i) the result type is a summary effect and (ii) the code is relevant. For these 19 articles, the set of target result values are broadly similar in type: there is a point estimate, a sample size, and some kind of measure of uncertainty (e.g., the upper and lower bounds of a confidence interval). To gauge the level of dependency between the reproductions of these different types of values within articles, Table S17 breaks down the results for each article by target value type, specifying how closely the target result value was reproduced (using the same categories reported in Table 6). For this summary, measures of uncertainty other than confidence interval bounds (e.g., standard errors) were ignored.

ID	N	Point est.	CI lower	CI upper
MA091	Exact	Exact	Exact	Exact
MA095	Exact	Exact	Exact	Exact
MA147	Exact	Exact	Exact	Exact
MA188	Exact	Exact	Exact	Exact
MA145	Exact	Exact	Rounding	At Least 10%
MA081	Exact	Exact	Failure	Failure
MA067	Exact	Exact	n.a.	n.a.
MA213	Exact	Exact	n.a.	n.a.
MA191	Exact	Rounding	Exact	Exact
MA071	Exact	Rounding	Rounding	Within 10%
MA229	Exact	Rounding	Within 10%	Exact
MA062	Exact	Rounding	Within 10%	At Least 10%
MA060	Exact	Rounding	At Least 10%	Within 10%
MA074	Exact	Within 10%	Exact	Within 10%
MA065	Exact	Within 10%	Within 10%	Within 10%
MA202	Exact	Within 10%	Within 10%	Within 10%
MA198	Within 10%	Rounding	Exact	Rounding
MA211	Failure	Failure	Failure	Failure
MA126	n.a.	Exact	Exact	Exact

Table S17: A breakdown of how closely target result values were reproduced for each article with relevant code and a summary effect result type. The target result value types are sample size N, point estimate, and confidence interval bounds (CI lower and CI upper). Values of "n.a." indicate that that particular target result value type was not reported for that article.

Table S17 considers how closely the sample size, point estimate, lower confidence interval bound, and upper confidence interval bound could be reproduced for each article. The closeness of the reproduced values were considered progressively, from left to right. The table shows that most values of sample size could be reproduced exactly, but the closeness of the reproduced values dropped off considerably after that for the point estimate, etc. There are two identifiable clusters: a cluster of four articles (MA091, MA095, MA147, MA188) where all target values could be reproduced exactly, and a cluster of three articles (MA065, MA071, MA202) where the sample size was reproduced exactly, and the remaining values were within 10%. This clustering may indicate that there is some dependency between values from the same article regarding how closely they will be reproduced. However, the sample is small and the categories of reproduction closeness are relatively coarse.

S10 REPRODUCING TARGET RESULTS WHEN CODE NOT RELEVANT

Table S18 details the circumstances of the six cases where shared code was judged not relevant to reproducing the target result.

ID	Study	Code language	Description
MA016	(38)	Python	Not relevant. The code shared is for simulations of leaf longevity, reported separately from the meta-analysis.
MA068	(43)	R	Partially relevant. The code shared regards the extraction of effect sizes from primary studies used in the meta-analysis. The code does not conduct the meta-analysis itself.
MA092	(60)	Fortran	Not relevant. The code shared is the source code for a modified version of the Ecosystem Demography Biosphere Model, ED2 (254). Simulations using this model were reported separately from the meta-analysis.
MA094	(61)	R	Partially relevant. The code shared is for generating null food web models. Although necessary, the code is not sufficient to reproduce the chosen result. Further, there was a "missing" code file: in the Oikos online appendix, one listed code file was actually missing (`hierarchy_measure.R`), while the other listed code file (`null_models.R`) was duplicated, resulting in two code files in the appendix with the same contents.
MA155	(52)	R	Not relevant. The code shared is for conducting Gene Ontology analyses, and for producing article Figure 1D. These are separate results from the meta-analysis.
MA212	(63)	R	Partially relevant. The code shared runs Spearman rank correlation tests, relevant to meta-analysis results presented in Table 3 of article. The code is not relevant to the selected meta-analysis result.

Table S18: The articles with shared code which was either not relevant or only partially relevant to reproducing the chosen meta-analysis results.

In the cases of MA016, MA092, and MA155, the shared code had nothing to do with the reported meta-analysis results. In the case of MA212, the shared code was partially relevant, but was practically unusable for the purposes of reproducing the specific results in the article. (Specifically, the shared code for MA212, written to calculate Spearman's rank correlation coefficient for multiple sets of data and summarise the corresponding p-values, seemed to be an extract from a larger code base; the code assumed a specific data structure that was not defined anywhere in the shared materials, nor did the data structure implied by the code correspond to any of the shared data files. Lacking contextual information on the setup required for the code to work, we decided that the code as provided for MA212 was unusable.) For these four cases we attempted to reproduce the originally selected target results detailed in Tables 4 and S10 by writing entirely new R code.

There were 59 target result values across MA016, MA092, MA155, and MA212. This set of target result values included 12 non-numeric values: these were entries from the table in article MA212, see Table S15. Percent error was not applicable to these non-numeric values, and so the reproduced values were assessed as being either exact text string matches with the original or non-matches. The details of the individual reproduction attempts for all these values are reported in Table S16.

In the cases of MA068 and MA094, the code was relevant to other parts of the meta-analysis described in the articles. We selected alternative target results from these articles that were directly relevant to the shared code. For MA068, the shared code performed simulations of logistic regression slopes and standard errors; these simulations were performed to supplement under-reported results from two primary studies included in the meta-analysis. The target results were the values of the simulations as reported in the article. For MA094, the shared code simulated species richness in food webs using different food web models; the simulation results were compared with the results from a sample of published food webs in a figure. The target results were the widths of bars in that figure, which represented the average proportion of species richness at different trophic levels for the published and simulated food webs. There were 3 target result values for MA068, and 21 target result values for MA094. The details of the alternative target results are reported in Tables S19-S21, and the results of comparing the reproduced values with the original values of these target results are in Table S22.

Variable	Value
ID	MA068
Study	(43)
Result source	Table 1 (pp.7-8)
Result type	Simulated slope parameters and standard errors to supplement incompletely reported primary study results
Standard error 1	0.001
Mean slope parameter 2	-0.001
Standard error 2	0.113

Table S19: Detail of the alternative target results selected for article MA068, (43). These alternative target results were selected due to being relevant to the shared code. Standard error 1 is simulated to supplement a result from (255), mean slope parameter 2 and standard error 2 are simulated to supplement a result from (256).

We used the shared code for these two articles to successfully calculate values for all 24 alternative target values. This perfect success rate is perhaps to be expected, since the alternative target results were specifically selected on the basis of being relevant to the shared code. Seven values out of the 24 (29%) were reproduced exactly (to the same precision as reported), another seven reproduced values (29%) were within 10% of the original value, and the remaining ten (42%) reproduced values were 10% or more from the original value. All

Variable	Value
ID	MA094
Study	(61)
Result source	in text (p.1227)
Result type	Descriptive statistics of species richness (i.e., N_{species}) in a sample of published food webs
$N_{\text{food webs}}$	72
Mean	90.21
Standard deviation	31.27
Minimum	50
Maximum	209

Table S20: Detail of the first set of alternative target results selected for article MA094, (61). These alternative target results were selected due to being relevant to the shared code.

ten reproduced values with substantial percent errors (10% or more) compared to the original were target results from simulations, which use pseudo-random number generation, and neither R script set a random seed which would have facilitated the exact reproduction of the simulations. For one target result in MA068, the mean slope parameter for a logistic regression, the reproduced value was 0.001, compared with an original value of -0.001. This is the only case in this study of a reproduced target result not being in the same direction as the original target result value. However, by using different random seeds, repeated simulations of this target result could yield different results, which might more closely agree with the original value.

Variable	Value
ID	MA094
Study	(61)
Result source	Figure 1 (p.1227)
Result type	Bar widths (in pixels) representing average proportions of species richness at different trophic levels for different food web types
Published food webs, top trophic level	215
Published food webs, intermediate trophic level	475
Published food webs, herbivore trophic level	430
Published food webs, basal trophic level	549
Random food webs, top trophic level	589
Random food webs, intermediate trophic level	521
Random food webs, herbivore trophic level	51
Random food webs, basal trophic level	108
Cascade food webs, top trophic level	79
Cascade food webs, intermediate trophic level	934
Cascade food webs, herbivore trophic level	158
Cascade food webs, basal trophic level	221
Niche food webs, top trophic level	441
Niche food webs, intermediate trophic level	408
Niche food webs, herbivore trophic level	102
Niche food webs, basal trophic level	385

Table S21: Detail of the second set of alternative target results selected for article MA094, (61). These alternative target results were selected due to being relevant to the shared code.

Table S22: The original and reproduced values of all alternative target results for MA068 and MA094. All target result values are numeric.

ID	Study	Result type	Effect size type	Target result	Original	Reproduced	Percent error (%)	Status
MA068	(43)	logistic regression model (Rödel)	slope parameter	SE	0.001	0.001	0.00	E
MA068	(43)	logistic regression model (Barber-Meyer)	slope parameter	point est.	-0.001	0.001	200.00	10%+
MA068	(43)	logistic regression model (Barber-Meyer)	slope parameter	SE	0.113	0.113	0.00	E
MA094	(61)	mean	species richness	point est.	90.21	90.21	0.00	E
MA094	(61)	mean	species richness	SD	31.27	31.27	0.00	E
MA094	(61)	mean	species richness	minimum	50	50	0.00	E
MA094	(61)	mean	species richness	maximum	209	209	0.00	E
MA094	(61)	mean	species richness	N	72	72	0.00	E
MA094	(61)	average proportion of species present	pixel width	published food webs, top trophic level	215	212	1.40	$<10\%$
MA094	(61)	average proportion of species present	pixel width	published food webs, intermediate trophic level	475	476	0.21	$<10\%$

ID	Study	Result type	Effect size type	Target result	Original	Reproduced	Percent error (%)	Status
MA094	(61)	average proportion of species present	pixel width	published food webs, herbivore trophic level	430	434	0.93	< 10%
MA094	(61)	average proportion of species present	pixel width	published food webs, basal trophic level	549	550	0.18	< 10%
MA094	(61)	average proportion of species present	pixel width	random food webs, top trophic level	589	96	83.70	10%+
MA094	(61)	average proportion of species present	pixel width	random food webs, intermediate trophic level	521	1078	106.91	10%+
MA094	(61)	average proportion of species present	pixel width	random food webs, herbivore trophic level	51	27	47.06	10%+
MA094	(61)	average proportion of species present	pixel width	random food webs, basal trophic level	108	64	40.74	10%+
MA094	(61)	average proportion of species present	pixel width	cascade food webs, top trophic level	79	252	218.99	10%+
MA094	(61)	average proportion of species present	pixel width	cascade food webs, intermediate trophic level	934	759	18.74	10%+
MA094	(61)	average proportion of species present	pixel width	cascade food webs, herbivore trophic level	158	160	1.27	< 10%
MA094	(61)	average proportion of species present	pixel width	cascade food webs, basal trophic level	221	227	2.71	< 10%
MA094	(61)	average proportion of species present	pixel width	niche food webs, top trophic level	441	359	18.59	10%+

ID	Study	Result type	Effect size type	Target result	Original	Reproduced	Percent error (%)	Status
MA094	(61)	average proportion of species present	pixel width	niche food webs, intermediate trophic level	408	441	8.09	< 10%
MA094	(61)	average proportion of species present	pixel width	niche food webs, herbivore trophic level	102	129	26.47	10%+
MA094	(61)	average proportion of species present	pixel width	niche food webs, basal trophic level	385	439	14.03	10%+

S11 REVISITING THE DEFINITION OF REPRODUCIBILITY

In this section, we return to the definition of reproducibility provided in the introduction, "reproducibility is obtaining consistent results using the same input data; computational steps, methods, and code; and conditions of analysis" (15, p.43) and consider each component of this definition in turn, in the context of the results of this study.

Consistent As is noted in the NAS report, there can be different standards for what is considered "consistent". In some scenarios, bitwise consistency may be required. In others, obtaining results in the same direction as the original might be considered good enough. The reproduced results in this study were compared to their original counterparts by looking at the percentage error. Looking at Table 6, relaxing standards for consistency from exact matches only to matches differing only by rounding precision and matches within 10% of the original increased the percentage of target results considered "consistent" from 43% to 57%. In the context of meta-analysis, what might be considered sufficient consistency will likely depend on the purposes that the results are put to use, and the sensitivity of those purposes to variation in the inputs. Meta-analysis in particular is an interesting case because meta-analyses can be updated with additional primary studies, and is complicated by differences of judgment over which primary studies ought to be included and excluded, etc. Given this, there may be an expectation that meta-analytic summary effects are already subject to variation beyond formal statistical error. In this context, there may be a tolerance for a certain amount of inconsistency in any asserted summary effect, such that small discrepancies of up to 10% in value when reproduced are not fatal (albeit perhaps still worthy of rigorous checking).

Results In the context of reproducing a numerical result, "results" are those numbers printed in the published article. Ordinarily, we take them as they are presented. However, in this study, we have the example of a result reported in article MA062 (40) which contains a typo (a missing minus sign). Here, the code and data produce the "correct" result, with a value less than zero. Here, interpretation and judgement is required: a reader can see that there is supposed to be a minus sign in front of the reported effect size, since that would then agree with the stated confidence interval. This example is particularly straightforward and obvious.

This is important to note because a lot of the challenge of evaluating computational reproducibility of results is in getting the data and code to "work"; understandably, that's where a lot of the focus is. But this perhaps takes for granted that the target value in the published article that is being reproduced is valid, and has not been corrupted during rounds of revision, copy editing, type setting, etc. (This is of course the chief motivation behind reproducible reports/documents, where analysis and text are contained within the one document, and so issues such as transcription error, etc. are mitigated.)

Another issue is the coverage/completeness of results. Do all "results" (e.g., all numerical values reported in text, all tables and figures) in an article need to be reproducible? For tractability, this study selected a single target result for reproduction across a number of articles, with the goal of selecting the first mentioned summary effect where possible. Even though this "bare minimum" attempt for each article covered only a tiny proportion of all results reported, the successes and failures were still informative.

Same input data Data sharing policies and advocacy perhaps may take for granted that the data file(s) that get shared are the same as the data file(s) that were actually used for the calculations reported in the article. But, this may not necessarily be the case: (i) Authors may "clean up" their data files in preparation for them to be shared. This may involve recoding of data values, or renaming of variables to make them more explicable to outside readers. This could introduce changes to how the data needs to be pre-processed or recoded for analysis. (ii) Some data files may be updated or edited over time, especially if used in projects which span more than a single article. It may become a non-trivial task to identify a single version of the data file(s) that applies to all results reported in an article.

In one meta-analysis (57), the data file shared was not the data file used with the provided code, and according to the content of the article itself, could not have been the data file used to calculate all results reported. This was due to a missing variable in the data set.

There is one meta-analysis where the authors explicitly provide two different versions of their data: the first which is the one actually used in the meta-analysis (and so is the one to be used to "reproduce" the results in the article), and the second which is a corrected version of the first, and which the authors recommend be used for further analysis. This example is particularly striking because of the transparency of the authors and the delineation they provide between "original" results and what might be called "correct" results.

The point of this is to say that when we say "the same data", we might not necessarily mean or intend to refer to "the specific original file(s) used by the authors in the calculation of the results". What we mean is a set of data that has the same substantive content as the original data, regardless of whether that version of the data was used by the authors to calculate the results or not.

Same computational steps, methods, and code This has multiple components: first, the sharing of code files has the same issues as the sharing of data files, as detailed above. Code files may be "cleaned up" for public release, or comments may be added, or code might be passed through a tool to format the code for easier reading. The point of this is to say that when we say "the same code", we may not necessarily mean the "specific original file(s) used by the authors".

The "same computational steps" also requires some nuance: for example, it's taken as given in studies evaluating reproducibility that things like the file system paths of input files do not really count as meaningful barriers to computational reproducibility. It does not seem "fair" to declare a result as unable to be reproduced purely because the code as written assumes a different file folder structure than exists on the reproducer's computer system.

The above example of file paths seems unambiguous enough (and is very common), but "alterations to the code as supplied" exists on a spectrum: if we agree that altering the computational steps to enable files to be read is at the end of "insubstantial changes", at what point do changes to the code as supplied become substantial, and we agree that we are no longer taking the same computational steps?

Examples from this study include typos in code that once corrected produce matching results. Correcting the (perhaps obvious) typo is making an act of interpretation: we're intuiting what the original authors intended, even though it is not literally what they have written in code.

This particular example also feeds back to the "same code" issue: if an analysis script contains a typo/syntax error that does not produce the results reported in the paper, can it be "the same" code run by the authors? If it was, they would have obtained an error message instead of a result, and so could not have reported that result in the paper. This makes it clear that the shared code file is not literally the code that was run to obtain the original results.

Beneath all this, there is some notion that when we refer to computational steps, we are referring to the computational steps that "really matter" to the calculation of the result.

Same conditions of analysis Should we take this to mean the same computing and software environment? In the context of this study and its results, it seems that the conditions of analysis mostly concerns software packages and perhaps their versions—the fact that meta-analysis results were typically only reported to the third decimal place at most, as well as the observed success rate at the target result level indicates that information about computer hardware (e.g., architecture, processors) is wholly unnecessary, as is information about operating and file systems. This would not be universal across research disciplines, but it seems reasonable for meta-analysis.

As found in this study, study authors often reported the software tools they used for analysis, even when code was not shared. Not all mentions of software were accompanied by information about versions, but this study shows that a lack of version information was not fatal to reproduction attempts, although software version differences may be contributing to the discrepancies between original and reproduced values.

One condition of analysis that clearly stood out as an issue was the lack of specification of random seeds for procedures which involved pseudo-random number generation. Without random seeds, such procedures become an inescapable source of discrepancy between original and reproduced values. Specifying a random seed can nullify this problem, however.