Insect Conservation and Diversity

Insect Conservation and Diversity (2018) **11,** 1–12 doi: 10.1111/icad.12284

FORUM & POLICY

Validating taxonomic identifications in entomological research

LAURENCE PACKER, SPENCER K. MONCKTON, THOMAS M. ONUFERKO and RAFAEL R. FERRARI Department of Biology, York University, Toronto, ON, Canada

Abstract. 1. We surveyed the treatment of taxonomic information in 567 papers published in nine entomological journals in 2016.

2. The proportion of papers that provide taxonomic data in sufficient detail to permit precise validation of taxonomic identifications is vanishingly small: most did not cite identification methods, most did not state whether identified material had been vouchered, and taxon concepts were almost universally absent in non-taxonomic papers. Overall, the combination of all three factors was provided less than 2% of the time and almost two-thirds of all papers provided none of the three.

3. We suggest that journals should modify the templates used by editors and reviewers by overtly including the following questions:

i. Are Order and Family named in the title, abstract or keywords?

ii. Are the methods used for identification of all studied taxa stated clearly?

iii. Is it clear who did the identifications, are they named and is their contact information and/or institutional affiliation provided?

iv. Is the literature whereupon these identifications are based cited appropriately? This would include some reference to as thorough a revisional taxon concept statement as possible, preferably from recent revisions if available.

v. Are exemplars of all focal species (or all sampled individuals) vouchered in a named repository (ideally with contact person name and accession numbers or other means of ready detection)?

4. Accurate and replicable taxonomic identification is the cornerstone of biology, without which entomological research risks becoming irreproducible and thus not scientific.

Key words. Identification keys, instructions to authors, revisional taxon concept, vouchers.

Introduction

The way in which taxonomic research output is cited by the scientific community has been the subject of some debate, especially given the increasing importance of citation metrics as measures of scientific impact. With these metrics, taxonomic papers fare poorly (Krell, 2000, 2002) despite that the vast majority of publications in biology concern organisms that have been identified at some point. This has led some authors to suggest that whenever a Latin binomial is noted in a research article, the paper in which the species was originally described should be included in the literature cited (e.g. Van der Velde, 2001; Seifert *et al.*, 2008; Sundberg & Strand, 2009; Wägele *et al.*, 2011; but see Dubois, 2008). This suggestion has motivated some journals to require that papers that mention a Latin binomial cite the original description (Bininda-Emonds, 2011). However, this is not always a sensible approach, as pointed out in a recent opinion piece by Meier (2017) who gave Meigen's (1830) two-line description of *Drosophila melanogaster* Meigen (Diptera: Drosophilidae) as an example of why original citations

Correspondence: Laurence Packer, Department of Biology, York University, 4700 Keele St., Toronto, ON M3J 1P3, Canada. E-mail: xeromelissa@gmail.com

are often misleading. Meier (2017) provided a rational argument for best practices in citing taxonomic publications. This included advice against citing original species descriptions when i) they are not adequate to permit identification, ii) more recent, rigorous identification tools are available, or iii) the original species delimitation is now known to be inaccurate (all three issues apply in the case of *D. melanogaster*). Meier (2017) argued that the methods sections of papers should include statements about identification techniques and taxon concepts (*sensu* Franz, 2005; but see below) and suggested that the actual identifications should be presented in the results section. For papers with large numbers of species identified, he recommended that a table should be provided that includes identification methods and relevant literature, species boundary (i.e. taxon concept) information, and voucher depository. His paper is already stimulating the editorial boards of several groups of journals to reconsider their guidelines for authors so that taxonomic research can be appropriately cited.

Defining the 'taxon concept' is central to ensuring best practice in citation of the appropriate taxonomic literature. Meier (2017) cited Franz (2005) for the term, but we find Franz *et al*.'s (2008: p. 9) definition somewhat clearer, where they state that 'a taxonomic concept is the underlying meaning, or referential extension, of a scientific name as stated by a particular author in a particular publication'. It is essential to clarify that taxon concepts are not the same as species concepts. Each taxon has its own taxon concept and this is likely to change over time as our understanding of that particular taxon changes. Species concepts are hypotheses about what species are in general (e.g. according to the biological species concept they are reproductively isolated, or according to the phylogenetic species concept they are diagnosably different with the differences being heritable). Thus, one might apply a single species concept to all insects, yet every named species and all of the supraspecific lineages in the hierarchy would each have their own taxon concept.

Franz and Peet (2009) provided a more detailed outline of the meanings associated with different types of taxon concept (their Table 1). Therein the 'original concept' is listed, which refers to the original species description alone, and is precisely the approach that Meier (2017) criticized. The taxon concept from Franz and Peet (2009) that seems closest to what is required for citing species boundaries, in our opinion, is the 'revisional concept': that which 'appears in a comprehensive revision of an existing taxonomic name and lineage, for example, a monograph, a species webpage, or a field manual with descriptions and illustrations' (Franz & Peet, 2009: p.7). We would exclude the latter from being considered revisional because so few entomological field manuals provide the taxonomic history or depth of detail required for precise species delimitation. Sigovini *et al*. (2016) also seem to suggest that keys, and even catalogues, might serve as taxon concepts. But for the same reason, we believe that formal taxonomic revisions are the best source of taxon concepts, though we note that recent species descriptions may also suffice (i.e. at least as recent as 1985, following the date of adoption of the third edition of the International Code of Zoological Nomenclature which recommended that a full diagnosis be provided for each new taxon; ICZN, 1985). Arguably, papers that do not cite valid revisional taxon concepts are not replicable and therefore less scientific because readers cannot know how the authors actually applied the scientific name.

Clearly, citing the identification methods used in a study and not just the original species description is another way to ensure the replicability of research, and also another way that taxonomic research could be fairly acknowledged. This will most often be done through the use and citation of identification keys (although as noted above, these would not necessarily constitute an adequate taxon concept in their own right). Bortolus (2008) found that in 62.5% of the 80 articles he surveyed in ecological journals, the authors said nothing about how the names of their study organisms were obtained. The situation was even worse in a survey of two biological control journals surveyed over 4 years (Frewin, 2015) in which identification methods were mentioned in only 21.1% of 601 papers. Seifert *et al*. (2008) noted the concern that citing all means of identification would result in a massive increase in the length of literature cited sections of research articles. However, Vink *et al*. (2012) found that for four evolutionary biology journals surveyed, citation of identification methods would result in an increase in at most one half of a printed page. In addition, if the number of species requiring such citation is enormous, the table and associated references would likely be placed in supplementary materials anyway. These authors note that the failure to mention how research organisms were identified makes the resulting papers 'unscientific', in the sense that they do not permit replication.

An additional way in which much research remains irreproducible comes from the lack of vouchering of specimens to permit the veracity of the identifications to be checked by others (Huber, 1998; Schilthuizen *et al*., 2015). Bortolus (2008) found that only 2.5% of the 80 papers he assessed had actually stated whether vouchers had been deposited. For biological control agents, only 9% of papers mentioned vouchering of material (Frewin, 2015). Turney *et al*. (2015) found that <25% of 281 papers dealing with arthropods stated that vouchers had been deposited, albeit with an increase over time (from 3.5% in 1994 to 35% in 2014).

Here, we provide baseline data for a wide range of variables related to the treatment of taxonomic information by surveying all papers published in nine entomological journals in 2016. These variables can be roughly grouped as relating to the following four overarching questions:

A) Are basic taxonomic details provided (e.g. classification, authority, what proportion of species level units are provided with a Latin binomial)?

Table 1. Basic taxonomic information. Numbers of papers for each response are given; in brackets are relative frequencies for each response, expressed as a proportion of all cases/papers to which the variable was applicable (cases in which a variable was deemed not applicable are detailed in footnotes below). These variables were scored for all papers we assessed (n = 567).

	Orders & families indicated?					Species authorities given?			Authorities given for all or not all focal species?			
Journal	Y	N	Not consistently	Family(ies) only	Order(s) only	Y	N	N/A*	All	Not all	Authorities not given	N/A*
AESA	87 (0.93)	2 (0.02)	4 (0.04)	1 (0.01)	0	78 (0.95)	4 (0.05)	12	72 (0.88)	6 (0.07)	4 (0.05)	12
AFE	35 (0.74)	10 (0.21)	1 (0.02)	1 (0.02)	0	43 (0.91)	4 (0.09)	0	39 (0.83)	4 (0.09)	4 (0.09)	0
CE	68 (0.99)	0	1 (0.01)	0	0	68 (1.00)	0	1	68 (1.00)	0	0	1
EE	42 (0.52)	23 (0.28)	2 (0.02)	9 (0.11)	5 (0.06)	61 (0.86)	10 (0.14)	10	56 (0.79)	5 (0.07)	10 (0.14)	10
ICAD	38 (0.69)	11 (0.20)	3 (0.05)	2 (0.04)	1 (0.02)	31 (0.62)	19 (0.38)	5	27 (0.54)	4 (0.08)	19 (0.38)	5
IMB	18 (0.26)	47 (0.67)	0	2 (0.03)	3 (0.04)	24 (0.36)	43 (0.64)	3	24 (0.36)	0	43 (0.64)	3
MVE	55 (1.00)	0	0	0	0	38 (0.76)	12 (0.24)	5	35 (0.70)	3 (0.06)	12 (0.24)	5
PE	23 (0.53)	19 (0.44)	0	0	1 (0.02)	38 (0.90)	4 (0.10)	1	38 (0.90)	0	4 (0.10)	1
SE	51 (0.96)	0	0	2 (0.04)	0	43 (0.90)	5 (0.10)	5	30 (0.63)	13 (0.27)	5 (0.10)	5
Overall	417 (0.74)	112 (0.20)	11 (0.02)	17 (0.03)	10 (0.02)	424 (0.81)	101 (0.19)	42	389 (0.74)	35 (0.07)	101 (0.19)	42

	If authority given, is year given?				Original description(s) cited?					If non-focal arthropods mentioned, same guidelines followed as for focal species?				
	Y	N	Not consistently	N/A†	Y	N	Not consistently	Auth/year not given	N/A‡	Y	N	Not consistently	Guidelines undefined§	N/A¶
AESA	25 (0.32)	52 (0.67)	1 (0.01)	16	17 (0.21)	5 (0.06)	4 (0.05)	56 (0.68)	12	37 (0.51)	24 (0.33)	2 (0.03)	10 (0.14)	21
AFE	8 (0.19)	32 (0.74)	3 (0.07)	4	1 (0.02)	10 (0.21)	0	36 (0.77)	0	22 (0.67)	8 (0.24)	1 (0.03)	2 (0.06)	14
CE	5 (0.07)	63 (0.93)	0	1	2 (0.03)	2 (0.03)	1 (0.01)	63 (0.93)	1	54 (0.95)	3 (0.05)	0	0	12
EE	11 (0.18)	49 (0.80)	1 (0.02)	20	3 (0.04)	9 (0.13)	0	59 (0.83)	10	24 (0.55)	10 (0.23)	4 (0.09)	6 (0.14)	37
ICAD	15 (0.48)	16 (0.52)	0	24	0	15 (0.30)	0	35 (0.70)	5	13 (0.62)	5 (0.24)	0	3 (0.14)	34
IMB	5 (0.21)	19 (0.79)	0	46	2 (0.03)	3 (0.04)	0	62 (0.93)	3	5 (0.07)	17 (0.25)	0	46 (0.68)	2
MVE	13 (0.34)	25 (0.66)	0	17	1 (0.02)	12 (0.24)	0	37 (0.74)	5	18 (0.41)	14 (0.32)	1 (0.02)	11 (0.25)	11
PE	0	38 (1.00)	0	5	0	0	0	42 (1.00)	1	3 (0.08)	29 (0.78)	1 (0.03)	4 (0.11)	6
SE	10 (0.23)	28 (0.65)	5 (0.12)	10	2 (0.04)	10 (0.21)	3 (0.06)	33 (0.69)	5	12 (0.44)	3 (0.11)	1 (0.04)	11 (0.41)	26
Overall	92 (0.22)	322 (0.76)	10 (0.02)	143	28 (0.05)	66 (0.13)	8 (0.02)	423 (0.81)	42	188 (0.47)	113 (0.28)	10 (0.02)	93 (0.23)	163

*These papers mention no species names or treat only new species.
†These papers do not provide authorities, or are N/A as in*.
‡These papers do not provide year of description (and therefore would not constitute proper taxonomic citations, or are N/A as in*.
§These papers do not follow consistent guidelines with respect to authorities for focal species.
¶These papers do not mention any non-focal arthropods.

B) Are taxonomic methods provided (e.g. how were the names derived, what tools were used, who performed the identifications)?
C) Are the taxonomic decisions verifiable (e.g. minimally through vouchering, maximally through stating where all the studied individuals are housed; stating the taxon concept applied to each named species; or, for cultured organisms, whether the sources are fully listed)? For taxonomic papers, an analogue of this would be whether the species concept was overtly stated.
D) How is the work of taxonomists recognised (e.g. citation of works in the references, are they overtly named)?

Materials and methods

We looked at all papers published in 2016 in the following entomological journals: *Annals of the Entomological Society of America* (AESA), *Canadian Entomologist* (CE) and all the journals of the Royal Entomological Society of London: *Agricultural & Forest Entomology* (AFE), *Ecological Entomology* (EE), *Insect Conservation and Diversity* (ICAD), *Insect Molecular Biology* (IMB*), Medical and Veterinary Entomology* (MVE), *Physiological Entomology* (PE), and *Systematic Entomology* (SE). We removed from consideration corrigenda, errata, editorials, tributes, or front & back matter, as well as special issue 148 (S1) of the Canadian Entomologist. For each remaining paper we obtained data on the 24 variables listed in Table S1, where we also state precisely how we evaluated them. Our analyses relate only to species level taxa unless otherwise stated.

As different journals were assessed by different co-authors, we went through three rounds of cross-validation. First, five papers from each journal were re-read by two authors and the data entries compared. This resulted in a refining of the initial criteria evaluated and more careful wording of the questions. Second, the senior author read five papers from each journal assessed by each of the other co-authors and any differences in interpretation were discussed. This resulted in further refinement of the questions asked and the decision to enter detailed comments for ambiguous cases where it was difficult to decide upon a straightforward answer to the question. Lastly, the penultimate version of the raw data was carefully inspected and all coauthors discussed how to deal with any ambiguities until consensus was achieved.

In addition to looking through titles, abstracts and the main text of the papers, we also checked the acknowledgements (e.g. taxonomic experts were often only named there) and the supplementary information (e.g. species lists were often only provided there).

We did not assess the quality of taxonomic information provided for non-arthropod organisms, such as plant hosts of herbivorous arthropods or the mammalian hosts of ectoparasites, as these groups are well beyond our own taxonomic expertise. Lastly, we looked at the 'guidelines to authors' sections on journal websites to assess what information was requested from authors. This enabled us to investigate the extent to which expressly stated expectations were met in practice.

Chi-Square Tests of Independence were carried out using the chisq.test function in the R package 'stats' (R Core Team, 2016) to look for differences between systematics or taxonomy papers and other kinds of entomological research papers in how they followed a particular practice (such as giving species authorities or listing vouchered material). Mann–Whitney U tests were carried out using the formulae and statistical tables in Sokal and Rohlf (1981) and Rohlf and Sokal (1981), respectively, to see if papers published in journals that mandated a particular practice (see Table S2) (giving the species authority, stating the Order and Family, and listing vouchered material) followed that practice more often than papers in journals that did not request that information.

Results

In total, we assessed taxonomic information from 567 papers with the largest number from AESA (94) and the smallest (43) in PE. Thirty-eight papers did not involve direct examination of insects (review articles, meta-analyses, opinion pieces, methods papers, letters to the editor) and were therefore excluded from questions pertaining to identification methods, specimen vouchering, and so on.

Basic taxonomic information

The number of focal taxa ranged from 1 to 1186, although it was not possible to determine this variable for 20 papers, because, for example, the total numbers of taxa were provided per habitat and the extent of overlap among habitats was impossible to deduce from all available information presented. Summary data in terms of average, SD, median and modal numbers of species treated in each journal, and overall, are provided in Table S3. Over half (51%) of all papers dealt with a single species, and this figure was larger (68%) in the overtly experimental journals (AFE, IMB, MVE, PE) than the rest (41%).

Species authority (although the ICZN mentions authorship, we use authority to avoid confusion with authorship of papers we are discussing) was provided for at least one focal taxon in 81% of all papers (ranging from 36% in IMB to 100% in CE; Table 1), and for all focal taxa in 74% of papers (ranging from 36% in IMB to 100% in CE; in other words, all papers in IMB and CE either provided species authorities for all focal species or provided no authorities at all). The species description date was given in only 22% of the papers that had provided species authority information for at least one species (ranging from 0% in PE to 48% in ICAD; Table 1). The instructions to authors in two of the target journals (AFE and IMB; Table S2) said nothing about whether species authorities should be provided, yet this information was given in 59% of cases (Table 1). Of the journals that required species authorities to be provided (the remaining seven), 87% of papers complied. The two groups did not differ significantly (Mann–Whitney U test, $U = 5$, $n_1 = 2$, $n_2 = 7$, $P > 0.1$, one-tailed). A total of 28 papers provided accurate taxonomic citations for all focal taxa for which authorities were given, while another eight papers did so for some. Thus, of the 525 papers with named species, only 5% cited species descriptions appropriately, ranging from 0% in ICAD and PE to 21% in AESA. Systematic or taxonomic papers included authorities for focal species 91% of the time whereas the proportion of other kinds of research papers that did the same was significantly lower, at 80% (analysed as raw counts) χ^2_1 ($n = 499$) = 4.6, $P = 0.031$.

Seventy-four percent of all papers provided Order and Family for all focal taxa (Table 1), an additional 2% did so for some, 2% gave Order only and 3% gave Family only. As above, AFE and IMB do not explicitly state that Order and Family should be provided yet, in combination, 45% of their papers provided this information for all focal taxa. For the journals for which stating Order and Family was required, the compliance rate was 81%. The difference was not significant (Mann–Whitney U test, $U = 3$, $n_1 = 2$, $n_2 = 7$, $P > 0.1$, one tailed). A citation to the classificatory scheme the authors applied to their focal taxa was provided in 19% of all papers that mentioned species level taxa, varying from 0% (PE) to 96% (SE).

Of the 404 papers in which non-focal arthropod taxa were named, only 47% provided author, or author and date, information in a consistent manner as for focal taxa. This ranged from 7% (IMB) to 95% (CE). In at least three instances, authorities for species names were provided for non-focal taxa but not for focal taxa.

Identification method

With new species description papers removed from the sample, only 29% of articles (Table 2) provided some information on the methods used for identification of the focal organisms, with a range of from 0% (PE) to 57% (ICAD). Lower values (20%, range 0–43%) were generally found in journals in which studies using model or cultured organisms predominated (AFE, IMB, MVE, PE) compared to the others (35%, range 22–57%). Of the 148 papers that cited an identification method, 86.5% used morphology, 6.1% used only molecular methods, 2.0% overtly used integrative approaches, and 5.4% did not involve study of the actual insects themselves, but were based upon other features (galls, leaf mines, or damage to plant hosts). Only eleven papers (7%) that did not involve the description of new species mentioned that types were consulted as part of the identification procedure.

Keys and/or other taxonomic literature were mentioned as being used to identify taxa in 57% of cases in which morphology-based identifications were performed. However, we found examples in which the keys were not all listed or cited, or were for different geographic regions than the provenance of the identified material and therefore inapplicable to at least some of the focal taxa.

Two papers presented results based upon identification of individuals entirely performed in the field without the possibility of verification and another five retained only the more difficult species for confirmation in the laboratory (in total 5%).

Of the 525 papers that had named species, 108 (20%) were partly or entirely based upon cultured organisms obtained from pre-existing stocks (Table 2; in an additional 32 cases, 6% of the total, it was not possible to determine whether the cultures were pre-existing or newly developed). Of these 108, 19% did not state the source of origin of the stocks used and 64% did not say anything about whether the stocks were being maintained (and were therefore potentially verifiable in the future). Only six papers based upon stock cultures stated that material was vouchered, and of these only three also stated that the stocks were being maintained. AESA is the only journal with instructions to authors which state how to cite cultures: ‘when possible, please provide as much genetic and/or colony information available is [*sic*] useful (for example, “Rockefeller colony of *Aedes aegypti* (L.)”). Including geographic origin and generations in culture is also useful, but may not always be known and is not required’. Yet detailed culture origin data were provided in this journal no more often than it was in others.

Thirty-nine papers (8%) that were not taxonomic in nature provided some information on the taxon concept that was applied to the identified material (Table 2). In the 48 overtly taxonomic papers, not a single one cited the species concept applied to their decisions, although two did discuss the issue to some extent (Table 2).

Vouchers

Overall, only 24% of the papers clearly stated that any studied specimens had been vouchered, ranging from 0% (PE) to 89% (SE) (Table 3). Only three journals provide some guidance as to whether vouchering should be performed (Table S2): it is required for AESA and CE, while SE states that it should be done when ‘[N]ew distributional and other noteworthy records’ are presented. When vouchering is seemingly required by a journal’s instructions to authors, only 50% of the papers comply with the recommendation; when nothing is stated about vouchers, only 7% of papers clearly stated that material was vouchered. These two groups differed significantly (Mann–Whitney U test, $U = 18$, $n_1 = 3$, $n_2 = 6$, $P < 0.025$ one-tailed). Almost half (48%) of all studies that vouchered any material made all specimens potentially available for additional study. However, only 87% of the

Table 2. Reporting of identification methods and related information. Numbers of papers for each response are given; in brackets are relative frequencies for each response, expressed as a proportion of all cases/papers to which the variable was applicable. Only papers involving direct study of material were assessed (n = 529).

	Method of identification mentioned?				Were focal taxa obtained from a pre-existing stock culture?				If yes, source given?			If yes, is stock maintained?		
Journal	Y	N	In part	N/A*	Y	In part	N	Not clear	Y	N	N/A†	Y	Not indicated	N/A†
AESA	28 (0.37)	47 (0.63)	0	15	15 (0.17)	1 (0.01)	73 (0.81)	1 (0.01)	15 (0.94)	1 (0.06)	74	9 (0.56)	7 (0.44)	74
AFE	14 (0.30)	30 (0.64)	3 (0.06)	0	11 (0.23)	0	36 (0.77)	0	11 (1.00)	0	36	1 (0.09)	10 (0.91)	36
CE	18 (0.30)	43 (0.70)	0	6	4 (0.06)	3 (0.04)	58 (0.87)	2 (0.03)	3 (0.43)	4 (0.57)	60	1 (0.14)	6 (0.86)	60
EE	14 (0.22)	51 (0.78)	0	2	4 (0.06)	2 (0.03)	55 (0.82)	6 (0.09)	3 (0.50)	3 (0.50)	61	0	6 (1.00)	61
ICAD	28 (0.57)	18 (0.37)	3 (0.06)	0	1 (0.02)	0	48 (0.98)	0	0	1 (1.00)	48	0	1 (1.00)	48
IMB	5 (0.07)	63 (0.93)	0	0	29 (0.43)	2 (0.03)	19 (0.28)	18 (0.26)	22 (0.71)	9 (0.29)	37	17 (0.55)	14 (0.45)	37
MVE	23 (0.43)	31 (0.57)	0	1	12 (0.22)	0	41 (0.75)	2 (0.04)	12 (1.00)	0	43	11 (0.92)	1 (0.08)	43
PE	0	39 (1.00)	0	0	23 (0.59)	0	13 (0.33)	3 (0.08)	20 (0.87)	3 (0.13)	16	0	23 (1.00)	16
SE	12 (0.33)	24 (0.67)	0	11	0	1 (0.02)	46 (0.98)	0	1 (1.00)	0	46	0	1 (1.00)	46
Overall	142 (0.29)	346 (0.70)	6 (0.01)	35	99 (0.19)	9 (0.02)	389 (0.74)	32 (0.06)	87 (0.81)	21 (0.19)	421	39 (0.36)	69 (0.64)	421

	Existing classification system cited?			Is/are species-level taxon concept(s) given?				If taxonomy paper, species concept given?			
	Y	N	Not for all‡	Y	N	Somewhat§	N/A¶	Y	N	Nearly‖	N/A**
AESA	17 (0.19)	72 (0.80)	1 (0.01)	9 (0.13)	58 (0.87)	0	23	0	21 (0.95)	1 (0.05)	68
AFE	3 (0.06)	44 (0.94)	0	0	47 (1.00)	0	0	0	0	0	47
CE	20 (0.30)	47 (0.70)	0	8 (0.13)	56 (0.88)	0	3	0	10 (1.00)	0	57
EE	4 (0.06)	63 (0.94)	0	7 (0.11)	58 (0.89)	0	2	0	0	0	67
ICAD	5 (0.10)	41 (0.84)	3 (0.06)	2 (0.04)	46 (0.94)	1 (0.02)	0	0	0	0	49
IMB	2 (0.03)	66 (0.97)	0	1 (0.01)	67 (0.99)	0	0	0	0	0	68
MVE	4 (0.07)	51 (0.93)	0	6 (0.12)	45 (0.87)	1 (0.02)	3	0	5 (1.00)	0	50
PE	0	39 (1.00)	0	0	39 (1.00)	0	0	0	0	0	39
SE	45 (0.96)	2 (0.04)	0	6 (0.67)	3 (0.33)	0	38	0	10 (0.91)	1 (0.09)	36
Overall	100 (0.19)	425 (0.80)	4 (0.01)	39 (0.08)	419 (0.91)	2 (0.00)	69	0	46 (0.96)	2 (0.04)	481

*These papers treat only new species.
†These papers do not use cultured material or treat only new species.
‡These papers cite a classification system for only a subset of all focal taxa.
§These papers discuss the matter in some detail, but do not cite a formal concept.
¶These papers do not analyse species-level data or are taxonomic revisions.
‖These papers fall short of explicitly stating a species concept, but do outline their approach to delineating species.
**These papers do not treat species-level taxonomy.

Table 3. Vouchers and acknowledgement of identifying experts. Numbers of papers for each response are given; in brackets are relative frequencies for each response, expressed as a proportion of all cases/papers to which the variable was applicable. Only papers involving direct study of material were assessed (n = 529).

	Are specimens accessible/vouchered? If so, all or not all specimens?				If vouchered, repository(ies) named?					If vouchered, can specimens be easily located?			
Journal	Y, all	Y, not all	Y, indet.*	N	Y	N	In part	Not directly†	N/A	Y	N	Not for all‡	N/A
AESA	21 (0.23)	16 (0.18)	3 (0.03)	50 (0.56)	38 (0.95)	2 (0.05)	0	0	50	16 (0.40)	21 (0.53)	3 (0.08)	50
AFE	0	5 (0.11)	0	42 (0.89)	4 (0.80)	0	1 (0.20)	0	42	0	5 (1.00)	0	42
CE	8 (0.12)	9 (0.13)	2 (0.03)	48 (0.72)	18 (0.95)	0	0	1 (0.05)	48	7 (0.37)	12 (0.63)	0	48
EE	2 (0.03)	0	2 (0.03)	63 (0.94)	3 (0.75)	1 (0.25)	0	0	63	0	4 (1.00)	0	63
ICAD	3 (0.06)	7 (0.14)	0	39 (0.80)	10 (1.00)	0	0	0	39	0	10 (1.00)	0	39
IMB	0	1 (0.01)	0	67 (0.99)	1 (1.00)	0	0	0	67	0	1 (1.00)	0	67
MVE	3 (0.05)	0	1 (0.02)	51 (0.93)	4 (1.00)	0	0	0	51	1 (0.25)	3 (0.75)	0	51
PE	0	0	0	39 (1.00)	0	0	0	0	39	0	0	0	39
SE	23 (0.49)	9 (0.19)	10 (0.21)	5 (0.11)	31 (0.74)	5 (0.12)	6 (0.14)	0	5	21 (0.50)	19 (0.45)	2 (0.05)	5
Overall	60 (0.11)	47 (0.09)	18 (0.03)	404 (0.76)	109 (0.87)	8 (0.06)	7 (0.06)	1 (0.01)	404	45 (0.36)	75 (0.60)	5 (0.04)	404

	If DNA used for ID, is it vouchered?			If expert-identified, experts named?			If expert-identified, institution given?		
Journal	Y	N/not stated	N/A§	Y	N	N/A¶	Y	N	N/A¶
AESA	5 (1.00)	0	85	10 (1.00)	0	80	7 (0.70)	3 (0.30)	80
AFE	4 (0.80)	1 (0.20)	62	8 (1.00)	0	39	6 (0.75)	2 (0.25)	39
CE	1 (0.33)	2 (0.67)	64	4 (1.00)	0	63	4 (1.00)	0	63
EE	2 (0.40)	3 (0.60)	63	1 (0.50)	1 (0.50)	65	0	2 (1.00)	65
ICAD	3 (0.60)	2 (0.40)	50	7 (0.70)	3 (0.30)	39	2 (0.20)	8 (0.80)	39
IMB	1 (0.50)	1 (0.50)	45	0	0	68	0	0	68
MVE	2 (1.00)	0	47	2 (0.67)	1 (0.33)	52	2 (0.67)	1 (0.33)	52
PE	1 (0.33)	2 (0.67)	44	0	0	39	0	0	39
SE	0	0	39	2 (1.00)	0	45	2 (1.00)	0	45
Overall	19 (0.63)	11 (0.37)	499	34 (0.87)	5 (0.13)	490	23 (0.59)	16 (0.41)	490

*For these papers it is impossible to determine whether or not all specimens are vouchered.
†Repositories are listed in an online database to which the paper refers.
‡These papers list only some specimens with sufficient detail to facilitate their discovery.
§These papers do not use DNA for identification or treat only new species.
¶These papers do not rely on assistance from experts or treat only new species.

papers that stated material had been vouchered clearly indicated where the specimens were housed (in some cases accession numbers were not associated with a named depository) and in only approximately one-third (36%) of cases was vouchered material likely to be easily discovered within the depository (through accession numbers or, for taxonomic papers, type labels). Differences were also apparent in comparing papers that were taxonomic in nature (in which 83% listed vouchered material) compared to those that were not (12%). This difference was significant χ^2_1 (n = 529) = 206.1, $P < 0.001$.

An average of 63% of papers using DNA data for identification clearly stated which markers were used and that the sequences were vouchered (Table 3).

Acknowledgement of taxonomists

We often found it impossible to decide whether the authors of a paper were those who performed the identifications ('were identified' being the only information provided – ICAD is the only journal with instructions to authors that states that the passive voice should be used). Third party 'experts' or 'specialists' were overtly stated as having assisted with identifications in 39 (30%) of papers with morphology-based identifications (Table 3). However, the identifying personnel were not overtly named in five instances. Of the 39 studies, 16 (41%) did not provide institutional affiliations or addresses for the associated expert(s). In only two instances (5%) were the methods used by third party experts stated.

Discussion

Repeatability is at the core of the scientific method. For biological research, repeatability requires that the methods used to identify the study organisms be overtly stated and that the material (or at least examples of it) be available for future verification (Bortolus, 2008; Vink *et al.*, 2012).

Most entomological research published in our focal journals in 2016 failed on one or both criteria and, therefore, does not pass the test of replicable science. Only 28.7% of papers overtly stated the identification methods used and only 23.6% vouchered identified specimens. The combination of both verifiable identification methods and voucher specimen deposition was detailed in only 10.7% of all applicable papers, while the trifecta of identification methods, vouchers, and cited taxon concepts (i.e. all three of Meier, 2017 recommendations) was given in a paltry eight papers: just 1.8% (Fig. 1) of the total (excluding those that described new species). Even if we add to this number those papers in which species identifications might be expected to be uncontroversial (i.e. maintained cultures) our data suggest that only 10.1% of entomological research papers followed the standard expectation of replicable science by stating the identification method and ensuring the availability of studied material.

Perhaps some might suggest that we are exaggerating the need for the incorporation of taxonomic information to all of entomology: surely organisms in culture, or other easily identified species, do not require taxonomic verification? This might be expected to be the case especially with model organisms. However, there are numerous counter-arguments. First, some model organisms occur as species complexes for which molecular data are required to confirm identification. *Anopheles gambiae* Giles (Diptera: Culicidae) is a prime example, with no fewer than eight morphologically indistinguishable but genetically differentiated species (Coetzee *et al*., 2013). Even species one might expect to approach the status of model organism have sometimes gone misidentified for generations (of scientists). For example, the medicinal leech currently used in medical practice is not *the* medicinal leech – *Hirudo medicinalis* L. (Annelida: Hirudinida: Hirudinidae) (Siddall *et al*., 2007). Second, the sources of some cultures of readily identified organisms may be suspect. For example, *Tenebrio molitor* L. (Coleoptera: Tenebrionidae) is commonly obtained locally for experimental purposes (pet stores keep them as reptile food). The *Tenebrio* identification key (http://entnemdept.ifas.ufl.edu/teneb/Tenebrionінae_subfamily.pdf) lists cuticle dullness and darkness as the characters to separate *T. molitor* from its darker congener *T. obscurus* F., yet the figures provided to separate them (figs. 26 and 27 *loc. cit*.) show *T. obscurus* to be no darker than *T. molitor* that are raised under crowded conditions (Silva *et al*., 2016; Fig. 1). Pet food taxonomy is unlikely to be reliable in providing the necessary taxonomic accuracy. An even more telling example is that of (Frewin, 2015) who found, using DNA barcode data, that three of 24 (12.5%) biological control agents purchased from commercial suppliers in Canada seemingly contained multiple taxa. In one instance distinct barcode clusters were found for samples from different suppliers, in the other two examples multiple putative species were provided by the same suppliers (*Wolbachia* was ruled out as a cause of the deep divergences). Third, the taxa that are considered 'readily identifiable' and for which identifications are generally 'uncontroversial' will depend on the level of expertise possessed by the identifier. Unless this information is provided, the reader might consider the identifications dubious. In addition, most complex taxa have species that are known to be easy to identify in comparison to most relatives, but Packer and Taylor (1997) demonstrated that this sometimes results from the 'easily identified' species not having received the careful attention that has been applied to their more 'difficult' relatives. Perhaps the most readily identified organisms are those that are pets. Yet Ruedas *et al*. (2000) decried the lack of vouchering of samples for well-known mammals, even the domestic dog.

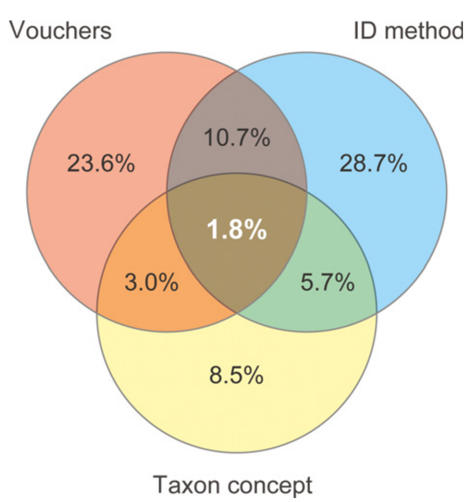

Fig. 1. Venn diagram showing the proportion of entomological research articles published in nine journals in 2016 that provided information on one or more of the three practices (how was material identied, were specimens vouchered, were taxon concepts provided for focal species) that relate to replicability of taxonomic information. [Colour figure can be viewed at wileyonlinelibrary.com].

Even information that is considered basic (and usually required according to journal guidelines to authors), such as Order and Family of the studied organism(s) or species authority when Latin binomials are used, were often lacking or treated inconsistently within single journal issues and even within single papers.

When identifications were performed by third party experts, these individuals were not always named, and even when they were their institutional affiliation was often not provided. The methods used by third party taxonomists were almost never mentioned. This is equivalent to stating that 'statistical analyses were performed by a statistician ($P < 0.05$)' with no mention of who the statistician was, what test they performed or what computer program they used to come up with the result. Obviously, it would be impossible to publish a paper if the statistical analyses were not explained in detail. The overwhelming impression we get from assessing entomological research published in 2016 is that taxonomic methodology is presented in a sloppy fashion such as would not be tolerated for any other part of the methods section in a scientific paper.

We assessed papers at face value and did not check whether the details provided (such as dates of original species descriptions) were accurate. However, we detected some clear mistakes, such as one paper which suggested two different authorities for the same species, and the single issue of a journal which had adjacent papers based upon both *D. melanogaster* Meigen and *D. melanogaster* L. [*sic*]. We also did not verify whether vouchered material had actually been deposited, although we know of at least one example (not in our focal journals or focal year) in which vouchers *with accession numbers* were not deposited, and those identifications proved fictitious (it included, for example, large numbers of individuals identified in the field which require molecular data for separation).

Meier (2017) suggested that data on species should be presented with information on: i) identification methods and literature; ii) species boundaries or taxon concepts and iii) the voucher depository. We would go a little further than this and make the following recommendations.

i) Authors should overtly state what identification methods were used, who performed them and cite the literature upon which they were based. Furthermore, the original sources of taxonomic information should always be cited: 'identifications provided as described in (our previous paper)' should not be an option (we did find examples when such citations proved not to cite identification methods either). This information should be provided even if the organism has been maintained in laboratory culture for generations. When was the last time independent taxonomic evaluation of the culture was made, and by which methods? If expert identifications were performed, even if the name and address of the relevant experts are given, the materials (keys, type specimen comparisons) the expert used to identify the specimens should still be mentioned. This is likely the best way to ensure replicability if the expert becomes inactive.
Similarly, the dates (years will likely suffice) over which the identifications were performed should be stated (as few as two focal papers provided this overtly). This is in part because even the global expert (s) on a taxonomic group will have temporal variation in their accuracy of identification for some taxa and their understanding of many will change over time, and also because the accepted taxonomy of that group changes over time. As an example, one genus included in Onuferko *et al.* (2018) was treated differently by the relevant expert early during this long-term study than toward the end.

ii) Authors should explicitly provide a valid taxon concept for each species-level taxon identified. The 'revisional concept' of Franz and Peet (2009) as modified above (i.e. based upon the most recent revisions or thorough descriptions with diagnoses) is arguably most appropriate as it incorporates the current and likely the most accurate limits of a particular species-level taxon (see also Meier, 2017). Taxon concepts should provide the Latin binomial, species authority and be followed by '*sec'* (meaning *secundum* – according to) followed by the source of that particular concept (Franz & Peet, 2009; Sigovini *et al.*, 2016). For example, *Chilicola neffi* Toro & Moldenke, *sec*. Monckton (2016) (Hymenoptera: Colletidae). In this instance the original description and identification key was found to encompass an additional species that was subsequently described by Monckton (2016) who thereby provided an updated taxon concept for *C. neffi*. Also, it would be useful if the authors state how confident they are in the identifications provided in their paper (a vanishingly small number of papers stated anything on this). Have any of the identifiers of the material treated in the paper examined type material of the species concerned (we found only two examples of a non-taxonomic paper in which type material was mentioned), did they have access to specimens that had been compared to the type(s), did they consult the entire species description and associated figures, did they have both sexes available? Sigovini *et al.* (2016) provide guidelines on how to cite different levels of taxonomic uncertainty.

iii) While vouchering of specimens has long been recommended as standard practice (e.g. Huber, 1998), whenever possible, all specimens included in larger scale survey research should be housed in a named repository. The obvious reason for this is that misidentifications are likely to occur not only with the vouchered exemplar chosen to represent the identity of that named species, but also in the non-vouchered material: species X may turn out to be a mixture of species X, Y, and Z. When taxon concepts change as a result of more recently uncovered taxonomic complexity, reinterpretations of earlier data will be required and this will not be possible without all material being available for future research.
We have heard it said that institutional repositories would not have the facilities to maintain the specimens that would result if this recommendation were implemented. But this seems to us as a circular argument – the more these facilities are used the greater their perceived value which should result in increased allocation of resources. Furthermore, arguments for maintaining, or increasing the resources available to such specimen housing institutions would be strengthened by statements such as 'without properly maintained specimen collections, 98% of research papers in entomology become unreplicable and therefore unscientific'. Scientists are increasingly encouraged to make their data available to posterity through archiving (e.g. Whitlock *et al.*, 2010). All this effort will count for little if there is no way of validating the names of the taxa whose data are stored. The fact that it is not

possible to validate the results of more than 90% of biological control studies because specimens were not vouchered (Frewin, 2015) is alarming given the contrasting results often found with 'the same' biological control agent and the societal and environmental impacts associated with use of the alternatives.

iv) Some literature searches are done by named taxonomic group. Academic search engines generally target titles and abstracts, and some tools like Google Scholar find keywords mentioned throughout the body of the document as well, but with a higher weighting on words used in the title (Beel *et al*., 2010). Order and Family names should therefore be stated at least in the title and abstract. Alas, one of the journals we evaluated (MVE) has guidelines that preclude their mention in titles, in which case, stating them in the abstract should be considered mandatory.

Those researchers using laboratory colonies, stocks or other long-term cultures, model organisms or other seemingly 'easily identified' taxa may baulk at the idea of having to add taxonomic information to their manuscripts. However, as noted above, there are examples of inaccurate taxonomy even with these. We would also note that research published in journals using cultured arthropods is often obviously 'applied' in nature and therefore errors are more likely to be problematic for society at large in the short term, suggesting that increased attention to taxonomic identifications are indeed required even in these instances. Interestingly, two of the more experimental journals (IMB and MVE) state in their instructions to authors that the methods should be presented in sufficient detail to permit the experiment to be replicated. We suggest that this is not possible unless identification methods and taxon concepts are overtly included and material is vouchered in a named and accessible repository.

Some may argue that as many arthropods are well known enough to have common names, their taxonomic details do not require citing. Indeed, one of our target journals (MVE) states that insect common names should be listed in the title but that Order and Family should *not* (italics ours). This is problematic for two reasons. First different jurisdictions use different common names for the same species (e.g. *Noctua pronuba* L. (Lepidoptera: Noctuidae) is listed as the large yellow underwing in the list used by CE, but the winter cutworm in the one used for AESA). Second, terms such as 'worms' occur in many insect Orders (not to mention various animal Phyla) and it is not always easy for even seasoned entomologists to know to which Order a particular entomological 'worm' belongs.

Our recommendations will undoubtedly cause concern among those using 'easily identified' study organisms, but there are two ways to simplify things. First, the identity of 'easily identified' species could be verified through sending vouchers to an acknowledged taxonomic expert on the group or through DNA barcoding (Ratnasingham & Hebert, 2013; Frewin, 2015). Routine independent verification of the identifications of plant pests (European and Mediterranean Plant Protection Organization, 2016) and cell line cultures (American National Standards Institute, 2015) are now expected. In one case a cell line culture was found to be incorrectly identified at the level of Kingdom (Lee *et al*., 2011). There should be nothing preventing entomologists using stock cultures from verifying their identifications as part of standard quality control procedures. Second, collaborations between taxonomists and those using 'easily identified' organisms could result in online databases where meaningful revisional taxon concepts can be maintained, updated whenever necessary and cited. Authors can then reasonably easily verify whether they have used the current concept. This would have the added benefit of increasing the citation metrics for taxonomic research. It has been suggested that taxonomic authority files (TAFs, Vanden-Berghe *et al*., 2015; Tessarolo *et al*., 2017) should be checked to ensure that the most recent name and classification is used (we found examples of the same species being stated to belong to different genera in different papers because updated classifications had not been checked). An online extension of something like this, that is updated every time a revisional taxon concept is changed, would be most usefully constructed for the most commonly studied insects.

The recommendations above are guidelines to what may be considered best practices for the treatment of taxonomic information (65% of papers followed 'worst practices' by adhering to none of them). These are largely no more than elaborations of ideas that have been around for a long time. It is somewhat sobering to discover how rarely published papers adhere to any of them (let alone all of them), whether doing so is stated in journal guidelines or not. It has been suggested that ensuring these changes take place is the responsibility of journal editors and reviewers (e.g. Vink *et al*., 2012). However, this places the onus on individuals who are already volunteering a lot of their time to the research community they serve and making their job more difficult is likely to be counterproductive. Most journals now use formal templates for assessing research papers. These commonly include questions such as 'have the statistical analyses been performed correctly, are ethical guidelines followed etc.?'. We suggest that the easiest way to ensure that taxonomic research is cited appropriately and for studies using named species to be made replicable (and therefore more truly scientific) is through changes to the templates used by editors and reviewers. Thus, our main recommendations are that these templates include the following questions.

1 Are Order and Family named in the title, abstract or keywords?

2 Are the methods used for identification of all studied taxa stated clearly?

3 Is it clear who did the identifications, are they named and is their contact information and/or institutional affiliation provided?

4 Is the literature whereupon these identifications are based cited appropriately? This would include some reference to as thorough a revisional taxon concept statement as possible, preferably from recent revisions if available.

5 Are exemplars of all focal species (or all sampled individuals) vouchered in a named repository (ideally with contact person name and accession numbers or other means of ready detection)?

If any of these questions elicit a 'no' or an 'unclear' response, then at least minor revisions should be required. These suggestions are more than just taxonomic niceties: they are essential to permit validation and replication in all areas of biology.

Acknowledgements

We are grateful to Raphael Didham for encouragement and for detailed comments on an earlier version of this manuscript and to Robert Hanner for discussion and for suggesting some references. SKM is funded by a Natural Sciences and Engineering Research Council postgraduate award, supplemented by a York University Graduate Fellowship and bursaries from the Faculty of Graduate Studies. TMO is supported by a Susan Mann Dissertation Scholarship. RRF thanks Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for his PhD fellowship (#11875/13-5).

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: doi: 10.1111/icad.12284:

Table S1. List of variables assessed, possible responses, and notes on how they were evaluated.

Table S2. Summary of taxonomically relevant instructions to authors for the nine focal entomological journals.

Table S3. Summary data in terms of number of focal species and proportion of species named (as opposed to morphospecies), for each journal and overall. Only papers involving direct study of material were assessed, and only those for which the number of focal species, or the proportion of those named, could actually be determined are included here. Sample sizes for these two metrics are not equal, because in some cases one could be determined but not the other.

References

American National Standards Institute (2015) Species-level identification of animal cells through mitochondrial cytochrome c oxidase subunit 1 (CO1) DNA Barcodes. ANSI/ATCC ASN-0003-2015.

Beel, J., Gipp, B. & Wilde, E. (2010) Academic search engine optimization (ASEO): optimizing scholarly literature for Google Scholar and Co. *Journal of Scholarly Publishing*, **41**, 176–190.

Bininda-Emonds, O. (2011) Supporting species in ODE: explaining and citing. *Organismal Diversity and Evolution*, **11**, 1–2.

Bortolus, A. (2008) Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *AMBIO: A Journal of the Human Environment*, **37**, 114–118.

Coetzee, M., Hunt, R.H., Wilkerson, R., Della Torre, A., Coulibaly, M.B. & Besansky, N.J. (2013) *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. *Zootaxa*, **3619**, 246–274.

Dubois, A. (2008) A partial but radical solution to the problem of nomenclature and taxonomic inflation and synonymy load. *Biological Journal of the Linnean Society*, **93**, 857–863.

European and Mediterranean Plant Protection Organization (2016) PM 7/129 (1) DNA barcoding as an identification tool for a number of regulated pests. *Bulletin OEPP/EPPO Bulletin*, **46**, 501–537.

Franz, N.M. (2005) On the lack of good scientific reasons for the growing phylogeny/classification gap. *Cladistics*, **21**, 495–500.

Franz, N.M. & Peet, R.K. (2009) Perspectives: towards a language for mapping relationships among taxonomic concepts. *Systematics and Biodiversity*, **7**, 5–20.

Franz, N.M., Peet, R.K. & Weakley, A.S. (2008) On the use of taxonomic concepts in support of biodiversity research and taxonomy. *Systematics Association Special*, **76**, 63–86.

Frewin, A.J. (2015) The application of DNA barcoding to enhance integrated pest management. Unpublished PhD thesis, University of Guelph, Guelph, ON, Canada.

Huber, J.T. (1998) The importance of voucher specimens with practical guidelines for preserving specimens of the major invertebrate phyla for identification. *Journal of Natural History*, **32**, 367–387.

ICZN (1985) *International Code of Zoological Nomenclature*, 3rd edn, pp. 338. The International Trust for Zoological Nomenclature, London, UK.

Krell, F.-T. (2000) Impact factors aren't relevant to taxonomy. *Nature*, **405**, 507–508.

Krell, F.-T. (2002) Why impact factors don't work in taxonomy. *Nature*, **415**, 957.

Lee, L.E.J., Bufalino, M.R., Christie, A.E., Frischer, M.E., Soin, T., Tsui, C.K.M., Hanner, R.H. & Smagghe, G. (2011) Misidentification of OLGA-PH-J/92, believed to be the only crustacean cell line. *In Vitro Cellular & Developmental Biology-Animal*, **47**, 665–674.

Meier, R. (2017) Citation of taxonomic publications: the why, when, what and what not. *Systematic Entomology*, **42**, 301–304.

Monckton, S.K. (2016) A revision of *Chilicola* (*Heteroediscelis*) a subgenus of xeromelissine bees (Hymenoptera: Colletidae) endemic to Chile. *ZooKeys*, **55**, 1–144.

Onuferko, T.M., Skandalis, D.A., Cordero, R.L. & Richards, M.H. (2018) Rapid initial recovery and long-term persistence of a bee community in a former landfill. *Insect Conservation and Diversity*, **11**, 88–99.

Packer, L. & Taylor, J.S. (1997) How many cryptic species are there? An application of the phylogenetic species concept to genetic data for some comparatively well known bee species. *Canadian Entomologist*, **129**, 587–594.

R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> 13th December 2016.

Ratnasingham, S. & Hebert, P.D.N. (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE*, **8**, e66213.

Rohlf, F.J. & Sokal, R.R. (1981) *Statistical tables*, 2nd edn. W.H. Freeman and Company, New York, NY.

Ruedas, L.A., Salazar-Brown, J., Dragoo, J.W. & Yates, T.L. (2000) The importance of being earnest: what, if anything, constitutes a "specimen examined?". *Molecular Phylogenetics and Evolution*, **17**, 129–132.

Schilthuizen, M., Vairappan, C.S., Slade, E.M., Mann, D.M. & Miller, J.A. (2015) Specimens as primary data: museums and 'open science'. *Trends in Ecology and Evolution*, **30**, 237–238.

Seifert, K.A., Crous, P.W. & Frisvad, J.C. (2008) Correcting the impact factors of taxonomic journal by appropriate citation of taxonomy (ACT). *Persoonia*, **20**, 103–107.

Siddall, M.E., Trontelj, P., Utevsky, S.Y., Nkamany, M. & Macdonald, K.S. III (2007) Diverse molecular data demonstrate that commercially available medicinal leeches are not *Hirudo medicinalis*. *Proceedings of the Royal Society B: Bilological Sciences*, **274**, 1481–1487.

Sigovini, M., Keppel, E. & Tagliapietra, D. (2016) Open nomenclature in the biodiversity era. *Methods in Ecology & Evolution*, **7**, 1217–1225.

Silva, F.W.S., Araujo, L.S., Azevedo, D.O., Serrão, J.S. & Elliot, S.L. (2016) Physical and chemical properties of primary defences in *Tenebrio molitor*. *Physiological Entomology*, **41**, 121–126.

Sokal, R.R. & Rohlf, F.J. (1981) *Biometry: the principles and practice of statistics in biological research*, 2nd edn. W.H. Freeman and Company, New York, NY.

Sundberg, P. & Strand, M. (2009) Taxonomic inflation or taxonomist deflation? A comment on Dubois. *Biological Journal of the Linnean Society*, **96**, 712–714.

Tessarolo, G., Ladle, R., Rangel, T. & Hortal, J. (2017) Temporal degradation of data limits biodiversity research. *Ecology and Evolution*, **7**, 6863–6870.

Turney, S., Cameron, E.R., Cloutier, C.A. & Buddle, C.M. (2015) Non-repeatable science: assessing the frequency of voucher specimen deposition reveals that most arthropod research cannot be verified. *PeerJ*, **3**, e1168.

Van der Velde, G. (2001) Taxonomists make a name for themselves. *Nature*, **414**, 148.

Vanden-Berghe, E., Coro, G., Bailly, N., Fiorellato, F., Aldemita, C., Ellenbroek, A. & Pagano, P. (2015) Retrieving taxa names from large biodiversity data collections using a flexible matching workflow. *Ecological Informatics*, **28**, 29–41.

Vink, C.J., Paquin, P. & Cruickshank, R.H. (2012) Taxonomy and irreproducible biological science. *BioScience*, **62**, 451–452.

Wägele, H., Klussmann-Kolb, A., Kuhlmann, M., Haszprunar, G., Lindberg, D., Koch, A. & Wägele, J.W. (2011) The taxonomist – an endangered race. A practical proposal for its survival. *Frontiers in Zoology*, **8**, 25.

Whitlock, M.C., McPeek, M.A., Rausher, M.D., Rieseberg, L. & Moore, A.J. (2010) Data archiving. *The American Naturalist*, **175**, 145–146.

Accepted 21 November 2017

Editor/associate editor: Raphael Didham