NAMING DINOSAUR SPECIES: THE PERFORMANCE OF PROLIFIC AUTHORS

MICHAEL J. BENTON

Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, U.K., mike.benton@bristol.ac.uk

ABSTRACT—Many current debates about biodiversity and large-scale evolution have identified the need for comprehensive species inventories. Such species lists may be incomplete because more collecting is needed, or because of errors by systematists. Empirical studies show that error rates are high, as much as 30–50% for many living and fossil groups. A clear requirement is skilled systematists, and more of them; but who does the best work? An empirical investigation of the 321 authors who named all 1400 species of dinosaurs since 1824 shows that prolific authors do worse than authors who name only a few dinosaurs, and the key difference is between the 147 authors who named only one species, and the 174 who named two or more. The most prolific author was Othniel Marsh, who named 98 species of dinosaurs (including 80 non-avian dinosaurs and 18 Mesozoic birds), but only 35 (36%) of his names are currently regarded as valid. The poor showing by prolific authors is not a result of their working at different times over the last two centuries, nor on dinosaurs of a particular age, body size, or quality class, nor that their work has been over-revised, but most likely because many prolific authors of dinosaur species names have been too interested in splitting species. Current tougher refereeing standards and international communication should continue to improve standards in naming new taxa.

INTRODUCTION

Correct species lists are crucial for good studies of evolutionary biology, ecology, paleontology, and biodiversity conservation (Gaston and Mound, 1993; Purvis and Hector, 2000; Wilson, 2003; Isaac and Purvis, 2004). The problem of error in such lists has long been recognized (May, 1990; Wilson, 1992; Gaston and Mound, 1993), and most of that error has been ascribed to synonymy, the act of naming a species (or other taxon) that has already been named. Such errors are unavoidable in view of the vast diversity of life, variability between genders, growth stages, and geographic populations, and the diversity of languages and countries in which taxonomists work. Errors are generally identified when synoptic revisions are done at a later date.

Species lists depend on the accumulated labors of many workers over the decades, and yet certain individual systematists have tended to dominate each group of organisms. These 'prolific authors' are researchers who are and were motivated to explore a particular taxonomic group thoroughly. In some cases, their chosen group is so obscure that they may labor alone in the world, without a living systematist with whom to discuss their work; for other organismal groups, there may be merely a handful of experts scattered widely around the world. Biologists and paleontologists in the great museums of the world have had the resources to collect and curate material and to produce hefty monographs in which the systematics of major groups are revised, and many new species may be named. Less prolific authors range from those who published very little of any kind, and yet named one or two species, to those with distinguished research careers in biology or paleontology, but who did not include naming species as a major activity.

Gaston and Mound (1993) noted that almost half the 6112 species of Thysanoptera (thrips) named from 1901 to 1993 were authored by four people, and one of these was responsible for 1065 names. The same is true for caecilians (limbless amphibians), whose species list up to 1970 was dominated by one researcher, Edward H. Taylor, who described 40 (25%) of the then-known 157 species (Nussbaum and Wilkinson, 1989).

For fossil brachiopods, Grant (1980) noted that Gustav Arthur Cooper of the Smithsonian Institution had named, or co-named, 442 (13%) of the then-known 3300 genera. In the case of dinosaurs, three researchers, Othniel C. Marsh, Friedrich von Huene, and Edward D. Cope, dominated the field for a long time; they were responsible respectively for 98, 71, and 64 species names (Fig. 1). By 1900, Marsh and Cope together had named 169 (47%) of the 362 dinosaur species names then in existence. Now, with a further century of study, the proportion of the total represented by their names has fallen to 12% (Benton, 2008b).

Such prolific authors would be expected to have a lifetime of expertise, and so ought to have an exemplary record in decision-making. Having seen more material, often around the world, and devoting themselves solely to the task, they should have a sure touch in recognizing synonymy and in determining when a new find is truly a new species. Other authors, who may name just one or two new species in their careers, might be expected to make mistakes more often. Indeed, for Thysanoptera, Gaston and Mound (1993) found slightly mixed evidence, but the most prolific author had an invalidation rate of only 8.5%, and they made a policy recommendation, "that it would be more cost effective to increase the level of support to particularly productive individuals rather than increase the total number of taxonomists" (Gaston and Mound, 1993: 142).

The purpose of this study is to scrutinize the work of prolific dinosaurian systematists, and to determine whether they worked in different ways from their less active colleagues, and whether their work was better as a result of accumulated expertise, or whether some might have named many species because they were motivated in other ways. The measures of 'quality' are the survival of their names to the present day, and the quality of the specimens they used, ranging from isolated scraps to whole skeletons. Of course, it may be that dinosaurian systematists are a special breed, subject to different pressures and temptations than systematists of other groups; further, working on the systematics of a fossil group may lead to different behavior or different outcomes from working on extant taxa.
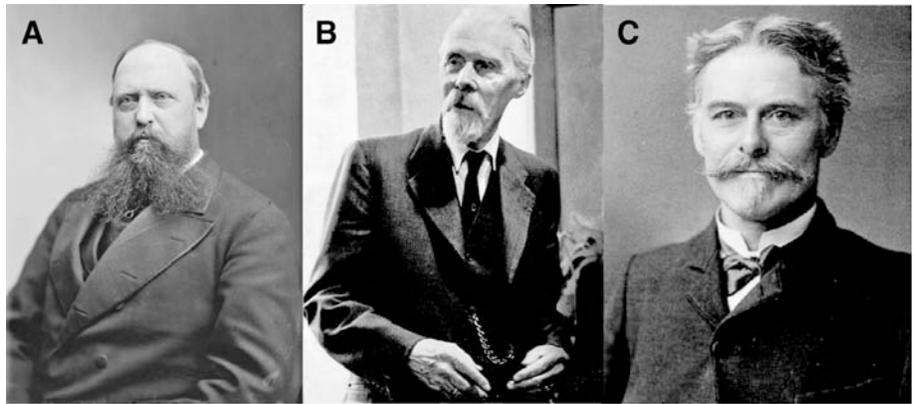
FIGURE 1. Portraits of the three most prolific authors of new dinosaur (and Mesozoic bird) names: **A**, Othniel C. Marsh (1831–1899), who named 98 species; **B**, Friedrich von Huene (1875–1969), who named 71 species; and **C**, Edward D. Cope (1840–1897), who named 64 species. (Marsh and Cope portraits from copyright-free sources; Huene portrait from Collections and Museum, Institute of Geosciences, University of Tübingen.)

SYSTEMATIC PRACTICE AND ERROR

Species lists may grow and shrink for many reasons. Hardest to document is the constant tension between 'splitters' and 'lumpers,' those systematists who divide species finely or coarsely. Splitters may recognize each population group or geographic variant as a species, whereas lumpers might call such forms varieties or subspecies of a larger species. There are no rules for determining the level at which varieties, subspecies, species, and genera ought to be discriminated, and such debates are even trickier in dealing with fossil species, where generally the only evidence is characters of hard tissues of skeletons or shells.

The current view of validity of species names probably depends as much on the quality of the earlier work as on the quality of revision. If a lumper reviews a splitter's work, many species will be moved to synonymy; in the reverse case, it is likely that most older species names will survive, and more will be added. In cases where interest is sustained and broad, such as in dinosaur systematics, revisions happen every few years or decades, and the risk of bias by a single active author or reviser is reduced. Species may survive scrutiny throughout, others may be identified as erroneous early on and remain in that state, whereas yet others may be identified as erroneous by one reviser, and then may be resuscitated by a later reviser, and so come in and out of validity many times. Alroy (2002, 2003) described a 'flux ratio' method to document the fluid nature of species lists through time. He noted that, despite the constant flux of names, most remain in the state (valid or invalid) determined by the first revision.

In early considerations of species lists and global biodiversity (e.g., May, 1990; Wilson, 1992), the error rate was assumed to be about 20–25%. Subsequent empirical studies have suggested rather higher rates, such as 7–80% (mean 31%) among insect groups (Gaston and Mound, 1993), 22% for thysanopteran insects (Solow et al., 1995), 31% for fossil mammals (Alroy, 2002), 33–88% (mean 66%) for seed plants (Wortley and Scotland, 2004), and 48% for dinosaurs (Benton, 2008b). Solow et al. (1995) and Alroy (2002) highlighted the need to consider the current synonymy total as fluid and itself subject to future revision; so current empirical synonymy rates are almost certainly underestimates because they do not include future revision of currently accepted and unrevised species. Following recalculation, they suggested a near doubling of current values, from 22% to 39% for thysanopteran insects (Solow et al., 1995), and from 31% to 50% for fossil mammals (Alroy, 2002).

Error in species lists has generally been termed 'synonymy' (May, 1990; Wilson, 1992; Gaston and Mound, 1993), and yet there are other forms of error, especially among fossil groups where material may be incomplete. Among dinosaurian species (Benton, 2008b), 675 of the 1400 named species are invalid, and yet only 230 of these (16.4%) are synonyms, 340 (24.3%) are designated nomen dubium, 47 (3.4%) are designated nomen nudum, and 58 (4.1%) are not dinosaurs. [Nomina dubia are based on undiagnostic or incomplete type materials and nomina nuda have no designated type material.] So, of invalid dinosaurian species, twice as many have been invalidated for reasons other than synonymy.

METHODS

The database (http://paleo.gly.bris.ac.uk/dinosaur/dinospecies.xls) is a complete listing of all 1400 dinosaurian species names compiled in order of publication, from the 1820s to the end of 2004. The primary source was Weishampel et al. (2004), a compilation of systematic chapters on all dinosaurian groups, but the listing was enhanced by adding papers to the end of 2004, but importantly by tracking each name through time to document invalidations and revalidations: this required a thorough reading of nearly every taxonomic paper on dinosaurs, from 1824 to 2004 (Benton, 2008a, 2008b). Data were also recorded on the date of naming, authors, geographic location (country, continent), geological age (period, epoch, stage, geological formation), size (total body length, arbitrarily divided into small [up to 2 m long], medium [2.01–5 m long], large [5.01–10 m long], very large [more than 10.01 m long]), validity (whether currently valid or not), and quality of the type specimen (isolated remains, a skull, several skulls, a skeleton, several skeleton).

In a previous paper, Benton (2008b) identified 30 prolific authors of non-avian dinosaurs, and he included species names to which they had been subsidiary authors. In the present study, all 1400 species of non-avian dinosaurs and Mesozoic birds are included, the latter representing some 100 of the total. Further, the 1400 dinosaur species names are assigned once only to the sole or first author of each name, to avoid double counting of authors. Most dinosaurs were named by a single author, and this used to be commonly the case: 359 of all 362 (99%) named up to 1900, 282 of all 303 (93%) named from 1900 to 1949, but 1032 of all 1400 (74%) named altogether up to 2005. The addition of Mesozoic birds to the database has altered the names and rankings of prolific authors of names from the list in Benton (2008a).

In this work, many calculations were done at both specific and generic levels, but results did not differ, so only species-level data are presented, as in previous publications in this series (Benton, 2008a, 2008b). In fact, 90% of valid dinosaurian genera are monospecific, with the mean number of species per genus being 1.2 (Wang and Dodson, 2006), similar to figures in my database (1401 species and 1036 genera; ratio 1.35, but this includes all valid and invalid taxa). These species-per-genus ratios are unusually low, especially when compared to living vertebrates;

for example, the species-per-genus ratio for nonpasserine birds is 3.983, for passerines 4.776, and for all birds 4.411 (Bock and Farrand, 1980). This opens a wider question of what really is a species of dinosaur, and whether paleontologists currently recognize too many or too few species. Probably the latter is the case, in that paleontologists must demonstrate morphological distinctions between taxa, and these are perhaps generally large enough to identify distinct genera, but sample sizes of individual genera are usually insufficient to differentiate the sources of smaller skeletal variations, whether ontogenetic, sexually dimorphic, geographic, or species-specific. In that many species of modern birds are distinguished from congeners on non-skeletal features such as feather colors and patterns, it may be hard to achieve a confident understanding of the true range of species within each dinosaur genus—although recent results in identifying colors and color patterns in the dinosaurs *Sinosauropteryx* (Zhang et al. 2010) and *Anchiornis* (Li et al., 2010) may point the way to a possible solution.

It is important to note that there is no natural cut-off point between prolific and less prolific authors: a rough balance in numbers is achieved when those who named 12 or more species (total, 669 species) are compared to those who named 10 or fewer (731 species). All other possible splits between prolific and non-prolific authors are considered according to a sliding scale, which considers a 'prolific' author as one who has named anything from 2 to 40 or more new species, and all numerical steps in between. The date of activity of an author is taken as the midpoint between the first and last published dinosaur name by that author.

Standard statistical approaches were used, including Spearman rank correlation for most plots of data against research time (Spearman's rho values, r_S, are quoted, and $P < 0.05$ is assumed for a significant result). Where error bars are included, these are based on calculation of minimum and maximum binomial confidence limits (Raup, 1991).

RESULTS AND DISCUSSION

Numbers

From 1824 to the end of 2004, 1400 species of purported dinosaurs (including Mesozoic birds) have been named (Benton, 2008a, 2008b). Of these, 355 are classed as small (mainly small theropods and birds), 346 as medium (mainly theropods and ornithischians), 472 as large (mainly larger theropods and sauropodomorphs), and 227 as very large (almost entirely sauropods, with a small number of theropods). The quality of type materials has improved, shifting from mostly incomplete material before 1960, to predominantly complete skulls and skeletons after that date (Benton, 2008a), and probably the quality of species discrimination has improved as a result of the adoption of cladistic methods from the 1980s and for other reasons, but that requires further investigation.

It would perhaps be premature to try to establish the significance of cladistic method on taxon-naming decisions at this stage, for three reasons: (1) the current cohort of namers of dinosaurs includes a broad mix of approaches, from strict through lax application of cladistics (e.g., whether differential diagnoses of new species are purely cladistic or not) to noncladistic, and it would be hard to divide this array into cladists and non-cladists; (2) the problem of insufficient time for revisers to consider the new taxa and make mature decisions about validity and invalidity could be very short and so would bias any comparisons with the work of earlier cohorts of workers; and, (3) conclusions might cause considerable offense to living and active systematists. Nonetheless, it is clear that there are 'good' and 'bad' practitioners among extant dinosaur systematists, and the poorer scientists are those who name incomplete material, bucking the longer-term trend to improve the quality of type specimens (Benton, 2008b), those

who do not make adequate comparisons among already-named materials and so fail to demonstrate that their new name is really necessary, and those who fail to apply strictly cladistic differential diagnostic methods. A reasonable prediction would be that their species are much more likely to fall to subsequent invalidation when the next systematic overviews are undertaken.

In all, 321 people have named one or more dinosaurian species, and the numbers named per author plotted against the midpoint of their active years (Fig. 2A) shows a slightly negative, but non-significant, slope ($r_S = -0.120$, $P = 0.031$). Authors are few in the earlier years, and build up to substantial numbers towards the present day, and those who named particularly large numbers of new species fall in the middle of the time range. The number of species named per author is highly right-skewed (Fig. 2B), with 147 authors having named 1 species each, 65 having named 2, down to one author each with 98, 71, 64, and 38 species to their credit. The pattern is similar if one tracks numbers of authors per name (Fig. 2C): most names have been authored by a single author, a very few by two, and it took until 1972 before three authors named a dinosaurian species. Since then, the majority of new dinosaurian species have been multi-authored, with totals of up to 13 authors per name. These extreme hollow-curve distributions have been noted for other groups, including insects (Gaston and Mound, 1993), decapods (De Grave, 2003), trilobites (Tarver et al., 2007), and mammals (Alroy, 2003).

The 'validity ratio' (= 'success rate'; i.e., number of currently valid species divided by total number of species named by each author) plotted against the midpoint of their active years (see online supplementary data, Fig. 3S) shows no relationship, whether with or without one-species authors. When grouped by decades (Fig. 3A), the validity ratio has improved more or less steadily from 1840 to 2000 ($r_S = 0.860$, $P = 0$), excluding the 1830 value (n = 2, and an edge effect—the first species named cannot be a synonym).

When grouped by decade (Fig. 3A), the number of authors rises exponentially through the last few decades. Numbers of species named is a highly irregular curve because it is bunched at the midpoints of the careers of authors—the two peaks represent the work of Marsh and Cope (1880) and Huene (1920), with a third and continuing peak in 1990 and 2000. When plotted by the decades in which they were named, the number of new species of dinosaurs has risen sporadically, as here, but strongly in the last three decades (Benton, 2008b).

The maximum number of valid species is clearly related to the number of species named by an author (Fig. 2SA), but the overall spread of values shows no relationship. Importantly, there is no relationship between number of species named and validity ratio when plotted by authors per year (Fig. 2SB) or per decade (Fig. 3B): the apparent slightly positive relationship is not significant ($r_S = 0.316$, $P = 0.214$).

The Performance of Prolific Authors

Twenty-three prolific namers (Table 1) were each responsible for more than ten dinosaur species names. These 23 active namers are responsible for 665 of the 1400 named species, so just under half (47.5%). Globally, 274 of their 665 new species (41%) are still regarded as valid, compared to 444 of the 735 (60%) new species described by the remaining workers. This contrast in validity ratio (Fig. 4A) is striking, and the difference is highly significant.

A sliding division was applied to see whether the more prolific authors really had less success than the others. The controlling factor appears to be the large class of authors of single taxa: they have achieved a much higher validity ratio than any other category, with 88 (60%) of their 147 species still regarded as valid. When the validity ratio for prolific and non-prolific authors is tracked (Fig. 4B) as the cut-off point between 'prolific'
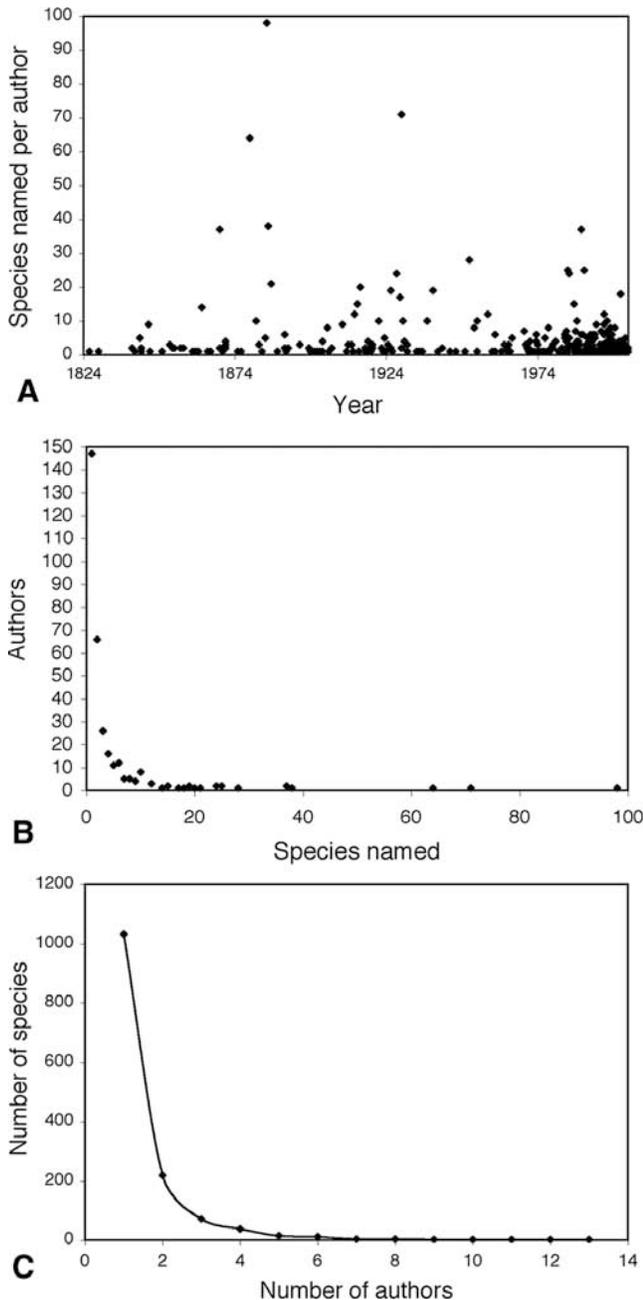
FIGURE 2. Key measures for the 321 authors who named one or more of the 1400 new species of dinosaurs. **A**, there is no correlation in the distribution of prolific authors through research time. **B**, the number of dinosaur species named per author is an extremely right-skewed curve, with very many authors of one species, and very few authors of more than 15 species. The best-fitting curve is a power-law curve ($y = 70.547x^{-1.1729}$; $r^2 = 0.8218$). **C**, the great majority of dinosaurs were named by a single author, with diminishing numbers of species named by two or more. The best-fitting line is a power-law curve ($y = 1475x^{-2.9384}$; $r^2 = 0.9669$), a far closer fit than the best exponential curve ($y = 426.34e^{-0.5509x}$; $r^2 = 0.8798$).

FIGURE 3. The quality of dinosaur species names as measured by the validity ratio. **A**, when the data are grouped by decades, number of authors follows an exponential curve (the best-fitting curve is $y = 2.7407e^{0.146x}$; $r^2 = 0.74$), whereas number of species named per author does not show a clear pattern, but validity ratio has improved to the present, barring the anomalous 1830 value. **B**, relationships between number of species named by all 321 authors of dinosaur names and validity ratio by decades. Validity ratio does not depend on number of species named, whether plotted by study year (Fig. 2SB) or by decade (here).

and 'non-prolific' slides down from 40+ to 2 species named, the values remain resolutely separate to the lowest limit. The difference in the validity ratio between prolific and non-prolific authors is not a result of the chance division of the data set. Interestingly, there is a high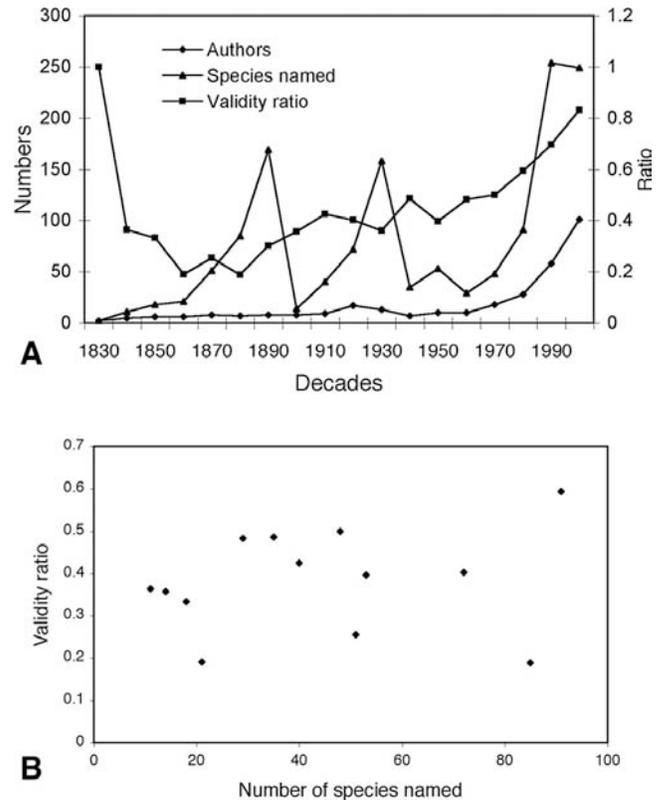ly significant negative relationship between mini-mum number of species described and validity ratio ($r_S = -0.997$, $P = 0.000$).

The three most prolific authors are Marsh, von Huene, and Cope (Fig. 1; Table 1), and there is a positive relationship between their validity ratios and number of species described, but the relationship is far from significant ($r_S = 1$, $P = 0.333$). For all 23 prolific authors, number of species named is significantly positively correlated with numbers still regarded as valid ($r_S = 0.468$, $P = 0.024$), and negatively with validity ratio, but not significantly ($r_S = -0.362$, $P = 0.089$). This confirms the findings for the whole group of authors, that the more species an author names, the more are still regarded as valid, but the lower the validity ratio.

A key criticism of this study could be the suggestion that prolific authors fare poorly because most of them operated 100 years ago, and most one-species authors are active today. In fact this is not true (Table 1): a plot of the ratio of prolific to non-prolific authors through research time (Fig. 5S) shows a non-significant horizontal-line relationship, and there is also no relationship between the numbers in each category per decade (Fig. 6S). Further, if the six living prolific namers of dinosaurs and/or Mesozoic birds (Dong Zhiming, Hou Lian-hai, José Bonaparte, Xu Xing, Peter Galton, Paul Sereno) are excluded, the positive relationship between number of species named and number still regarded as valid is even more significant ($r_S = 0.664$, $P = 0.004$), but the

TABLE 1. The authors who named most dinosaur species (including non-avian dinosaurs and Mesozoic birds).

	Author (active in last 50 years)	No. of dinosaur and bird species named	No. of Mesozoic birds included	No. of species still valid	Validity ratio (no. valid/total)	Years over which new names published
1	Othniel Marsh	98	18	35	0.36	1870–1899
2	Friedrich von Huene	71	0	18	0.25	1902–1956
3	Edward Cope	64	0	9	0.14	1866–1892
4	Harry Seeley	38	2	7	0.18	1869–1901
5	Dong Zhiming[*]	37	1	22	0.59	1973–2004
6	Richard Owen	37	1	7	0.19	1841–1897
7	Yang Zhong-jian[**]	28	0	11	0.39	1931–1972
8	Hou Lian-hai[*]	25	21	16	0.64	1975–2004
9	Alexander Nesov	25	14	11	0.44	1983–1995
10	José Bonaparte[*]	24	2	23	0.96	1969–2000
11	Charles Gilmore	24	0	13	0.54	1909–1946
12	Richard Lydekker	21	0	8	0.38	1877–1895
13	Lawrence Lambe	20	0	11	0.55	1902–1929
14	Barnum Brown	19	0	9	0.47	1908–1943
15	Charles M. Sternberg	19	0	8	0.42	1926–1953
16	Xu Xing[*]	18	0	18	1.00	1999–2004
17	William Parks	17	0	5	0.29	1922–1935
18	Peter Galton[*]	15	0	12	0.8	1971–2001
19	Franz von Nopcsa	15	0	5	0.33	1900–1929
20	Joseph Leidy	14	0	3	0.21	1854–1872
21	Albert de Lapparent	12	0	4	0.33	1955–1960
22	Henry F. Osborn	12	0	8	0.67	1903–1924
23	Paul Sereno[*]	12	1	11	0.92	1988–2004

Only those names for which the worker was sole, or first, author, are included. This reduces the life totals for some active living authors, such as Dong Zhiming, Peter Galton, Hou Lian-hai, Paul Sereno, and Xu Xing, compared to the list of all names to which active authors had contributed, given by Benton (2008b). Three over-active paleontologists, Othniel Marsh and Edward Cope in the 19th century and Friedrich von Huene in the 20th lead the table for most dinosaur species named, each with more than 60 names to their credit. The list then includes all other authors who named, or co-named, 14 or more dinosaur species. The 'no. of species still valid' column is based on current assessments (primarily Weishampel et al., 2004), and the 'validity ratio' is the ratio of names still regarded as valid to the total numbers named. The number of species named per year is based on the number of active 'naming' years for each paleontologist, dating from the first to last published paper with a dinosaur species name, and persons who are still active are marked with an asterisk (*). Active workers terminate in 2004 for the purposes of this study.

[**]Yang Zhong-jian's name was often westernized to C.-C. Young.

negative relationship between number of species named and validity ratio is less so ($r_S = -0.347$, $P = 0.173$). The same holds as prolific workers active after 1950, and other cohorts are successively pruned backwards in research time.

Comparisons of various research-time cohorts (Table 2) confirms that there is rarely a correlation between number of taxa named and validity ratio, only for the 1820–1899 and 1900–2004 cohorts of all authors, but not for other partitions of the data. These correlations are remarkable, in that they show a reversal in slope at 1900: in other words, the 1820–1899 cohort shows positive correlation between number of taxa named and validity ratio, the 1900–2004 cohort a significant negative slope. So, although the mean validity ratio doubles from the 19th century to the 20th (0.267 to 0.599), the relationship between number of taxa named by an author and that author's validity ratio switches from positive to negative. Most tellingly, the mean validity ratio for species named by prolific authors in the 1820–1899, 1900–1949, and 1950–2004 time spans are always some 5% worse than the values for non-prolific authors, although differences are not discriminated when 95% binomial confidence limits are added. In all cases, except for the nine prolific authors active from 1950 to 2004, there is a significant or highly significant correlation between number of species named by an author and number still deemed valid.

The finding was a real surprise, that for Victorian times, six prolific authors named far more dinosaur species (n = 272) than their 45 less prolific colleagues (n = 91), some 75% of the total, their overall performance was markedly poorer, and this poor performance of prolific authors persists through all divisions of research time. The overall very poor performance of prolific authors (Fig. 4) is not an effect of swamping the prolific Victorian namers by large numbers of one-species authors active today.

The Behavior of Prolific Authors

Why should this remarkable difference of some 19% exist between the validity ratios of prolific and non-prolific authors? There is no difference in the distributions of quality categories between the two author sets (Fig. 5A; $r_S = 1$, $P = 0.017$). The fossil size categories selected by both groups track each other less closely (Fig. 5B; $r_S = 0.8$, $P = 0.333$), but the differences are insufficient to explain the difference in validity ratios; in other words, prolific authors are not focusing more on large or small dinosaurs than other authors. Further, the two classes of authors have sampled more or less equally from each of the seven geological epochs from Upper Triassic to Upper Cretaceous (Fig. 4C; $r_S = 0.847$, $P = 0.016$). Further study of the database (Benton, in prep.) shows that specimens from the Late Triassic and Lower Jurassic tend to be of poorer quality than those from other epochs, and yet prolific authors actually under-sample from those poorer time intervals, and over-sample from the Upper Jurassic and Early Cretaceous, where quality and validity values are higher.

The poor performance of prolific authors cannot be explained by measurable factors of quality, size, or age of the fossils they used as their type specimens. Nor is there a clear relationship between the time when such authors operated and the current view of the validity of their species. Further, there does not seem to be a geographic reason, whether by the source of specimens or the
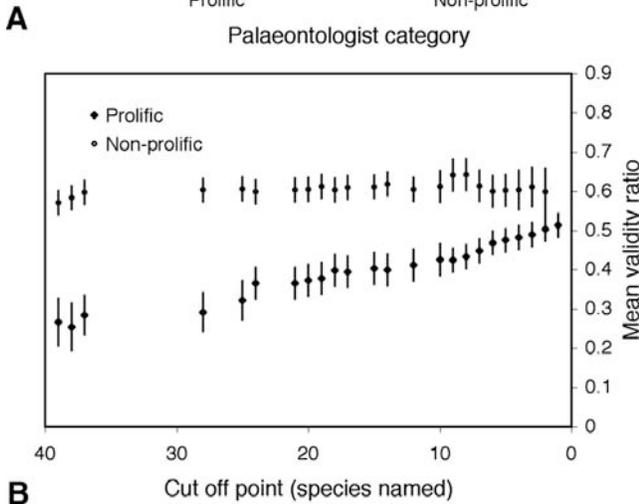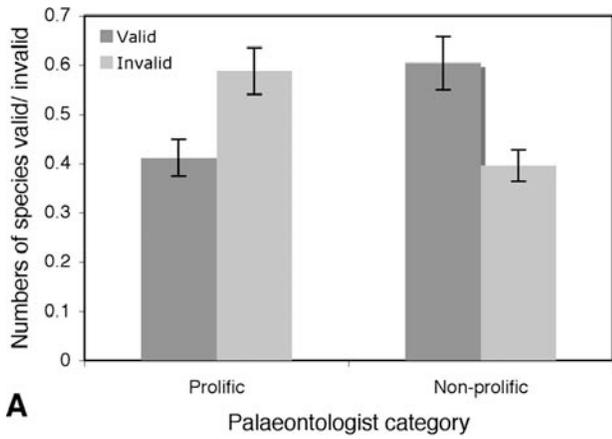
FIGURE 4. Comparison of the relative success of prolific and non-prolific authors, measured by validity ratios. **A**, prolific authors have highly significantly lower validity ratios than non-prolific authors (error bars are binomial confidence limits at $P = 0.05$, giving ranges from of 0.375 to 0.450 and 0.569 to 0.639). **B**, the poorer validity ratios of prolific authors are not an outcome of the chance division of the cohort of authors; here, the split between 'prolific' and 'non-prolific' slides from greater than 40 species named down to two or more, and the validity ratios for prolific authors remain significantly lower, according to binomial confidence intervals at $P < 0.05$, from those for less prolific authors.

home base of the systematist: of the top ten prolific namers, some worked on North American materials (Othniel Marsh, Edward Cope), others on Chinese (Dong Zhiming, Yang Zong-jian, Hou Lian-hai), British (Richard Owen, Harry Seeley), Central Asiatic (Nesov), or South American (José Bonaparte) specimens, and Friedrich von Huene worked on materials from Europe, South America, Africa, and other parts of the world, but largely not North America.

This leaves the choice between the quality of their work or their approach to systematics. It is most likely that prolific namers of dinosaur species tended to be splitters and perhaps also tended to be optimists. Their motivation was to discover new species, whereas that of other workers was focused on other themes of paleobiology or paleoecology. The strong motivation to name species stems from the essential need for correct and accurate species inventories, but might lead to splitting of species on small differential characters, and a willingness to see diagnostic characters where perhaps they do not exist. Note that it would be wrong simply to label prolific namers of taxa as splitters for two reasons: (1) many, particularly those still alive, have had extraordinary success in identifying apparently truly novel taxa and so they
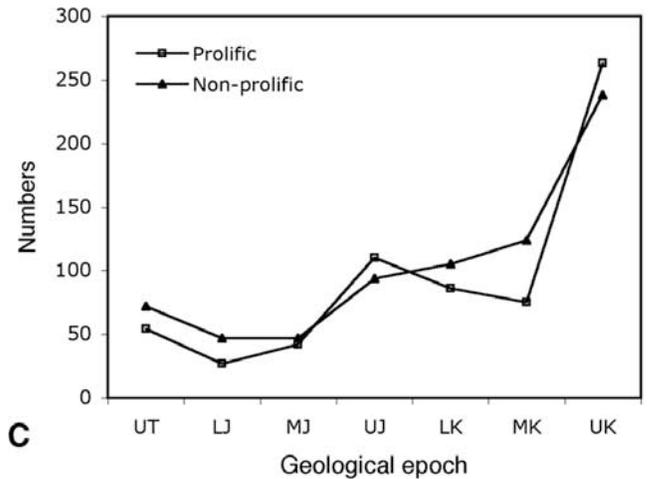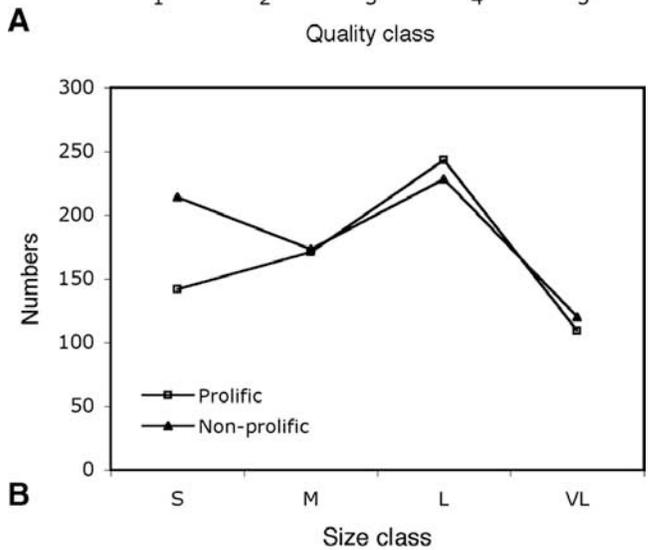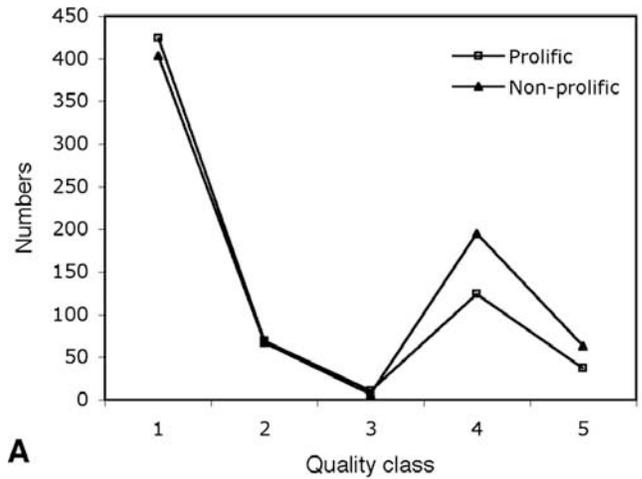

FIGURE 5. Prolific authors do not appear to differ from non-prolific authors in the distributions of (**A**) quality classes, (**B**) size classes, or (**C**) geological epochs from which their material derived.

are unlikely to see many of these invalidated; and (2) the majority of invalidity decisions, in the past at least, have concerned nomina nuda and other inadequately described or identified material (84%), rather than synonymy (16%, see above), and so their failing has perhaps been optimism rather than over-splitting.

TABLE 2. Performance of all authors and prolific authors by epoch of work.

Research epoch	Number of authors	Correlation of no. of species with:				
		Number still valid	P value	Validity ratio	P value	Validity ratio
1820–1899	45A	0.743**	0.000	0.315*	0.035	0.267
1820–1899	6P	0.812*	0.035	−0.314	0.564	0.254
1820–1899	39NP	0.582**	0.000	0.287	0.077	0.308
1900–1949	54A	0.766**	0.000	0.020	0.883	0.396
1900–1949	8P	0.879*	0.004	−0.228	0.588	0.391
1900–1949	46NP	0.600**	0.000	0.039	0.799	0.403
1950–2004	231A	0.767**	0.000	0.123	0.067	0.654
1950–2004	9P	0.402	0.284	0.252	0.513	0.653
1950–2004	222NP	0.733**	0.000	0.097	0.158	0.701
1820–1949	99A	0.742**	0.000	0.124	0.220	0.327
1820–1949	14P	0.640*	0.014	−0.419	0.146	0.311
1820–1949	85NP	0.571**	0.000	0.100	0.361	0.362
1900–2004	276A	0.757**	0.000	−0.123*	0.041	0.599
1900–2004	17P	0.636**	0.006	−0.181	0.488	0.533†
1900–2004	259NP	0.701**	0.000	0.088	0.159	0.646†
1820–2004	321A	0.734**	0.000	−0.081	0.147	0.513†
1820–2004	23P	0.468*	0.024	−0.362	0.089	0.412†
1820–2004	298NP	0.662**	0.000	0.047	0.418	0.604†

The table shows results for Spearman rank correlations (Spearman's rho, r_S), two-sided probability values, and whether the correlation is significant (*), highly significant (**), or not, and either positive or negative. The final column shows the mean validity ratio for the cohort, and cases (†) where 95% binomial confidence limits indicate distinct ranges. **Abbreviations**: **A**, all authors; **NP**, non-prolific authors only; **P**, prolific authors only.

Dinosaurs may be an extreme example, and studies of other extinct and living groups may confirm that the group has been subject to some 'extreme' behavior. Since Victorian times there has been considerable cachet associated with naming a new dinosaur species, and that is still true today. For example, a new species of dinosaur is more likely to grace the pages of *Nature* or *Science*, and to attract wide press interest, than a new species of almost any group other than fossil hominids. Further, specimens are often incomplete, and variables such as sexual dimorphism, ontogeny, geographic variation, and individual variation are poorly understood. This, combined with the time and money involved in securing a dinosaur specimen, may constitute unusual pressures on systematists to err on the side of naming a new species, even if they might personally prefer to be cautious.

Of course, this study assumes that the current status of revision is correct, but that is of course debatable. It is likely that enthusiastic revisers might themselves be lumpers, and they would then be glad to sink many earlier names in their revisions. This then acts as a bias against the fine subdivision of species. If the reviser, on the other hand, were a splitter, then s/he might 'discover' many previously unrecognized species, and determine that conservative systematists were in the wrong and that the enthusiastic namers of species were correct. In such a case, the fact that prolific authors named many more dinosaurs currently regarded as erroneous could turn their fortunes around, and they could then be regarded as having behaved correctly, whereas the conservative and unprolific authors of names might be criticized for their caution. Only time will tell whether the current stance in which prolific naming is frowned upon is correct or not.

The three most prolific namers of dinosaurs (Marsh, Cope, Huene) did some fieldwork themselves, but relied more on their extensive networks of collectors and contacts who supplied them with abundant material (Marsh, Cope), or made materials available to them when they traveled the world (Huene). Living active namers, such as Bonaparte, Sereno, Dong, and Xu, may suffer from the enthusiasm of earlier prolific authors, although all the evidence suggests that this is not the case. However, it is hard to make a meaningful comparison between the work of people operating today, and those operating before, because the more recently named (say, post-1990) species have not necessarily been subject to such thorough revision as earlier-named taxa: a kind of 'pull of the recent' effect (Raup, 1972) masks the current trends.

If this study highlights any message for the future, it is that systematists ought to spread their interests widely. A motivation simply to name new species is dangerous, and can lead to overhopeful establishment of new taxa on the basis of inadequate material. A broader range of interests, including evolutionary and paleobiological studies, as well as engagement in substantial systematic works in which older taxa are compared and revised, perhaps means that the practitioner would be more cautious in naming new taxa. The evident improvement in the quality of type specimens (Benton, 2008a), in communication between workers in different countries, in the use of photographs of material, in journal refereeing standards, and in regulation by the ICZN and other bodies should mean that taxonomic standards overall continue to improve, and that many of the errors of the past ought not to be repeated. Further, the striking trend towards multiple authorship of new dinosaurian names (Fig. 2C) surely means that naming decisions are discussed and tested among the group, and that it is now rather rare for an author to work alone.

CONCLUSIONS

This study has shown that prolific authors of dinosaur names, whether 'prolific' means a lifetime total of more than 40, more than ten, or even two or more names, have been more likely to name invalid new species than authors of a single name. This is somewhat counter-intuitive because it might be expected that authors of single names would tend to be dilettantes, dabbling in the field and liable to make mistakes, whereas prolific authors would have a stock of experience that would ensure they made fewer mistakes.

The wide divergence in current validity of dinosaur species names between prolific and less prolific authors is not a result of the choice of different specimen qualities, different dinosaurian sizes, different ages, or different geographic source regions. Nor is the difference an artifact of different eras of work, say between scientists active in Victorian times, and scientists active

today. The difference must reflect broadly different attitudes to systematics, most likely the ever-present tension between 'splitters' and 'lumpers': prolific namers of new dinosaur species must have been on the whole splitters, whereas the revisers of names in recent compilations (e.g., Weishampel et al., 1990, 2004) must be more often than not lumpers.

It will be instructive to find whether such imbalances exist in the nomenclatural history of other taxa living and extinct, and whether current methods and standards act better to regulate some of the over-enthusiastic naming of species that happened in the past.

Supplementary Material—Additional information can be found in the online supplementary data file that accompanies this paper.

ACKNOWLEDGMENTS

LITERATURE CITED

Alroy, J. 2002. How many named species are valid? Proceedings of the National Academy of Sciences of the Unites States of America 99:3706–3711.

Alkroy, J. 2003. Taxonomic inflation and body size distributions in North American fossil mammals. Journal of Mammalogy 84:431–443.

Benton, M. J. 2008a. Fossil quality and naming dinosaurs. Biology Letters 4:729–732.

Benton, M. J. 2008b. How to find a dinosaur, and the role of synonymy in biodiversity studies. Paleobiology 34:516–533.

Bock, W. J., and J. Farrand Jr. 1980. The number of species and genera of Recent birds: a contribution to comparative systematics. American Museum Novitates 2703:1–29.

De Grave, S. 2003. Historical patterns in the description of northeast Atlantic Decapoda. Journal of Crustacean Biology 23:669–677.

Gaston, K. J., and L. A. Mound. 1993. Taxonomy, hypothesis-testing and the biodiversity crisis. Proceedings of the Royal Society of London, Series B 251;139–142.

Grant, R. E. 1980. The human face of the brachiopod. Journal of Paleontology 54:499–507.

Isaac, N. J. B., and A. Purvis. 2004. The 'species problem' and testing macroevolutionary hypotheses. Diversity and Distribution 10:275–281.

Li, Q., K.-q. Gao, J. Vinther, M. D. Shawkey, J. A. Clarke, L. D'Alba, Q. Meng, D. E. G. Briggs, and R. O. Prum. 2010. Plumage color patterns of an extinct dinosaur. Science 327:1369–1372.

May, R. M. 1990. How many species? Philosophical Transactions of the Royal Society of London, Series B 330:292–304.

May, R. M., and S. Nee. 1995. Taxonomy—the species alias problem. Nature 378:447–448.

Nussbaum, R. A., and M. Wilkinson. 1989. On the classification and phylogeny of caecilians (Amphibia: Gymnophiona), a critical review. Herpetology Monographs 3:1–42.

Purvis, A., and A. Hector. 2000. Getting the measure of biodiversity. Nature 405:212–219.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177:1065–1071.

Raup, D. M. 1991. The future of analytical paleobiology: pp. 207–216 in N. L. Gilinsky and P. M. Signor (eds.), Analytical Paleobiology, Short Courses in Paleontology, No. 4. Paleontological Society, Knoxville, Tennessee.

Solow, A. R., L. A. Q. Mound, and K. J. Gaston. 1995. Estimating the rate of synonymy. Systematic Biology 44:93–96.

Tarver, J. E., S. J. Braddy, and M. J. Benton. 2007. The effects of sampling bias on Paleozoic faunas and implications for macroevolutionary studies. Palaeontology 50:177–184.

Wang, S. C., and P. Dodson. 2006. Estimating the diversity of dinosaurs. Proceedings of the National Academy of Sciences of the United States of America 103:13601–13605.

Weishampel, D. B., P. Dodson, and H. Osmólska (eds.). 1990. Dinosauria. University of California Press, Berkeley, California, 733 pp.

Weishampel, D. B., P. Dodson, and H. Osmólska (eds.). 2004. Dinosauria, second edition. University of California Press, Berkeley, California, 861 pp.

Wilson, E. O. 1992. The Diversity of Life. Penguin, London, 424 pp.

Wilson, E. O. 2003. The encyclopedia of life. Trends in Ecology and Evolution 18:77–80.

Wortley, A. H., and P. W. Scotland. 2004. Synonymy, sampling and seed plant numbers. Taxon 53:478–480.

Zhang, F., S. L. Kearns, P. J. Orr, M. J. Benton, Z. Zhou, D. Johnson, X. Xu, and X. Wang. 2010. Fossilized melanosomes and the colour of Cretaceous dinosaurs and birds. Nature 463:1075–1078.