Increasing numbers of bird species result from taxonomic progress, not taxonomic inflation

George Sangster*,#†
National Museum of Natural History, PO Box 9517, 2300 RA Leiden, The Netherlands

The impact and significance of modern taxonomy on other fields in biology have been subjects of much debate. It has been proposed that increasing numbers of vertebrate species are largely owing to ‘taxonomic inflation’. According to this hypothesis, newly recognized species result from reinterpretations of species limits based on phylogenetic species concepts (PSCs) rather than from new discoveries. Here, I examine 747 proposals to change the taxonomic rank of birds in the period 1950–2007. The trend to recognize more species of birds started at least two decades before the introduction of PSCs. Most (84.6%) newly recognized species were supported by new taxonomic data. Proposals to recognize more species resulted from application of all six major taxonomic criteria. Many newly recognized species (63.4%) were not based exclusively on PSC-based criteria (diagnosability, monophyly and exclusive coalescence of gene trees). Therefore, this study finds no empirical support for the idea that the increase in species is primarily epistemological rather than data-driven. This study shows that previous claims about the causes and effects of taxonomic inflation lack empirical support. I argue that a more appropriate term for the increase in species is ‘taxonomic progress’.

Keywords: taxonomy; species limits; species criteria; species concepts

1. INTRODUCTION

The number of described vertebrate species is increasing rapidly (Haffer 1992; Glaw & Köhler 1998). The underlying cause of this increase and its impact on other fields in biology have been subjects of recent debate (e.g. Dubois 1998; Hanken 1999; Isaac et al. 2004; Padial & de la Riva 2006). A major point of controversy is whether recent increases in the number of vertebrate species are problematic (Chaitra et al. 2004; Isaac et al. 2004; Tattersall 2007) or represent progress (Groves 2001; Dubois 2003). Isaac et al. (2004) recently applied the term ‘taxonomic inflation’ to cases in which many existing subspecies are raised to the species level. These authors argued that the increase is caused by a change in the species concept, rather than ‘new discoveries’. Isaac et al. (2004) suggested that a recent trend away from the biological species concept (BSC) towards the phylogenetic species concept (PSC) represents the main cause of the increase. They regarded the growth of species numbers owing to the elevation of subspecies to species as ‘unnatural’ (Isaac et al. 2005, p. 280).

One characteristic of taxonomic inflation identified by Isaac et al. (2004) is that taxonomic changes are biased towards certain groups, which these authors attribute to the charisma, rarity or ease of study of these groups. They warned that such biases affect macroecological studies and conservation biology. Isaac et al. (2004) indicated that taxonomic inflation could compromise comparative studies because newly recognized taxa are effectively ‘pseudoreplicates’.

Isaac et al. (2004) based their claims on an informal analysis of taxonomic changes in mammals and birds. Their analysis re-affirmed that species numbers are increasing, but did not provide any evidence in support of their specific claims, namely that the increase (i) is epistemological rather than data-driven and (ii) is biased towards charismatic, rare or easily studied taxonomic groups.

The taxonomic inflation thesis has generated much discussion (e.g. Agapow & Sluys 2005; Harris & Froufe 2005; Knapp et al. 2005; Köhler et al. 2005; Padial & de la Riva 2006). Some authors have asserted that the species-level taxonomy of vertebrates is data-driven but offered no quantitative data on the role of new taxonomic data (Padial & de la Riva 2006; Dubois 2008). So far, only two quantitative studies have addressed some of the claims made by Isaac et al. (2004). Köhler et al. (2005) provided evidence that the increase in amphibian diversity in Madagascar is largely owing to intensified exploration and application of molecular and bioacoustic techniques, rather than to the elevation of subspecies to species rank. Padial & de la Riva (2006) showed that the increase in the number of recognized amphibian species began two decades before the formal introduction of the PSC in the early 1980s. They also noted that only a small fraction of the subspecies that have been raised to species rank since 1980 were due to explicit adoption of the PSC and/or use of phylogenetic techniques. However, it is not clear whether these observations are representative of other taxonomic groups such as birds and mammals on which Isaac et al. (2004) based their claims.

*george.sangster@nrm.se
†Present address: Department of Vertebrate Zoology, Swedish Museum of Natural History, PO Box 50007, 104 05 Stockholm, Sweden.

Received 7 April 2009
Accepted 22 May 2009

This journal is © 2009 The Royal Society
Here, I test whether taxonomic studies of birds support the idea that increasing numbers of bird species result from epistemological changes rather than an increased availability of new data. I also examine whether the increase in bird species is biased towards charismatic species. Specifically, I test the following predictions. First, if the increase in species numbers results from ‘a recent trend away from the broad-brush biological species concept (BSC) towards more fine-grained phylogenetic species concepts’ (Isaac *et al.* 2004, p. 464), then one would expect that the increase is recent and did not start before the early 1980s when various PSCs were formalized (Cracraft 1983; Donoghue 1985). Second, if the increase in species numbers is caused primarily by reinterpretations of previous data under the PSC, as claimed by Isaac *et al.* (2004), then one would expect that, after the introduction of the PSC, a large portion of newly proposed species are based on previous data. Third, if the PSC is causing the increase in species, then many splits should be based exclusively on the taxonomic criteria of the PSC (i.e. diagnosability, monophyly and exclusive coalescence). Fourth, if the increase in bird species is biased towards the charismatic groups, one would expect a proportionally greater increase in the number of charismatic species than of non-charismatic species. Finally, I present evidence in support of an alternative hypothesis that attributes differences in the rate at which species numbers increase in various taxonomic groups to historical biases in the application of the polytypic species concept.

2. MATERIAL AND METHODS

(a) Taxonomic criteria and new data

The effects of taxonomic criteria and new data on the increase in bird species between 1950 and 2007 were tested using a dataset with proposals to change the rank of at least one taxon. This time span includes the period in which ‘evolutionary systematics’ is believed to have dominated taxonomy (Mayr 1982; Vernon 1993) and the introduction of various PSCs (Cracraft 1983; Donoghue 1985). This period is therefore well suited to test whether increasing numbers of species can be attributed to the introduction of PSCs.

Taxonomic proposals were located in seven major ornithological journals: *The Auk, Bulletin of the British Ornithologists’ Club, The Condor, Emu, Ibis, Ostrich and The Wilson Bulletin* (renamed *The Wilson Journal of Ornithology* in 2006). These journals were selected because each regularly publishes taxonomic papers, has a wide geographical coverage, serves a broad community of ornithologists and has a complete run in the study period, forming a total of 406 journal-years. Both regular issues and supplements were included. All publications in these journals that were likely to include some kind of proposal were examined. Proposals for the placement of a nominal taxon into the synonymy of another taxon (*or vice versa*) were excluded because such proposals do not involve a ranking issue. Proposals were also excluded if the scientific names of the relevant taxa are unknown or unspecified. Proposals in book reviews, symposium abstracts and essays in which a particular taxonomic problem is used as an example were also excluded. Taxonomic recommendations in the reports of the taxonomic committees of the British Ornithologists’ Union (published in *Ibis*) and the American Ornithologists’ Union Committee on Classification and Nomenclature (published in *The Auk*) were excluded because these represent reviews and are not primary taxonomic research literature.

It was recorded whether the taxonomic proposal referred to (i) a change from subspecies to species (a ‘split’) or (ii) a change from species to subspecies (a ‘lump’). To control for the effect of a general increase in taxonomic activity, the proportion of splits (i.e. the number of splits divided by the total number of taxonomic proposals) was used rather than the absolute number of splits.

To determine whether taxonomic changes are a reinterpretation of previous evidence or an effect of new evidence, it was noted whether new taxonomic information relevant to the proposal was included. New information may refer to new distributional data, new evidence of reproductive isolation or new examinations of morphological, acoustic, ecological, behavioural or molecular data. If no new taxonomic information was included, the proposal was regarded as a reinterpretation of previous data.

If a rationale for the taxonomic rank of the focal taxon was presented by the authors, the rationale was categorized as one or more of six categories of ranking criteria: diagnosability, degree of difference, monophyly, exclusive coalescence, adaptive zone and reproductive isolation. These six ranking criteria were selected because these criteria feature prominently in discussions over species concepts and each represents the primary criterion of one or more species concepts (e.g. Mayden 1997; de Queiroz 2007).

The ranking criterion was determined on the basis of the criteria that were actually used by the authors even if they have stated that their case is based on a different criterion or species concept. For instance, if the authors stated that they used the BSC but the case is based on diagnostic differences only (without a case for reproductive isolation), the taxonomic criterion was scored as ‘diagnosability’.

(i) Diagnosability

Hypotheses were considered to be based on this criterion if the focal taxon was ranked based on the possession (or lack) of unique, fixed character states or a unique combination of character states. Although many taxonomic revisions include a diagnosis, this does not necessarily mean that the taxa were ranked (i.e. as species or subspecies) on the basis of the criterion of diagnosability. Therefore, the rationale was categorized as diagnosability only if ranking was explicitly based on this criterion.

(ii) Degree of difference

Hypotheses were considered to be based on this criterion if the author considered the relevant taxa to differ ‘too much’ to be treated as subspecies, or too little to be considered as species. Differences may refer to morphological, biometric, behavioural or genetic characters or distances.

(iii) Monophyly

Hypotheses were considered to be based on this criterion if the taxonomic ranks of two or more taxa were based on their phylogenetic relationships (i.e. their position relative to each other in the ‘species tree’).
(iv) Exclusive coalescence
Hypotheses were considered to be based on this criterion if ranking was based on exclusive coalescence of gene trees (i.e. reciprocal monophyly or the existence of ‘phylogroups’).

(v) Adaptive zone
Hypotheses were considered to be based on this criterion if taxa were ranked on the basis of ecological differences or the occupation of different niches.

(vi) Reproductive isolation
Ranking on the basis of reproductive isolation can take many forms including evidence for (or for lack of) interbreeding, hybridization, gene flow, fusion of populations, differences in mate choice, reduced hybrid fitness or combinations of these.

To determine whether recent taxonomic changes may be viewed as a correction of past mistakes in the application of the polytypic species concept, for each proposed split it was recorded whether the focal taxon was originally described as a species or a subspecies.

(b) Charisma and degree of polytypy
The effects of charisma and degree of polytypy on the increase in bird species were tested using a second dataset. This dataset comprised 1132 taxa that became recognized as species between 1951 and 2007 owing to the elevation of subspecies to species rank. The numbers of species recognized by Mayr & Amadon (1951; \( n = 8590 \)) and Clements (2007; \( n = 9935 \)) were used as the starting point, from which 20 extinct species in Mayr & Amadon (1951) and 233 species newly described between 1951 and 2007 listed in Clements (2007) were excluded. To allow comparisons between the two taxonomies, family taxa recognized in the Peters checklist were used (Peters 1931–1951; Mayr & Greenway 1960, 1962; Mayr & Paynter 1964; Paynter 1967–1970; Mayr & Cottrell 1979, 1986; Traylor 1979).

To determine whether the increase in species is biased towards ‘charismatic’ groups, as suggested by Isaac et al. (2004), avian family taxa were scored for three characteristics: (i) body size, (ii) morphological distinctiveness and (iii) familiarity among non-biologists, using a three-point scale (0, 1, 2). Low values for these characteristics denote species that are small, indistinct and unfamiliar, respectively. The total number of points (0–6) was used as a measure of the ‘charisma’ of the species in these families. Only families with more than 60 species were used for statistical analysis.

To determine whether there is a relationship between the degree of polytypy of family taxa and the increase in species numbers, (i) the proportion of polytypic species in each family and (ii) the mean number of subspecies per polytypic species were determined for each family. Families with more than 60 species were used for statistical analysis. Peters’s checklist was used as a reference (Peters 1931–1951; Mayr & Greenway 1960, 1962; Mayr & Paynter 1964; Paynter 1967–1970; Mayr & Cottrell 1979, 1986; Traylor 1979).

(c) Statistical analysis
To test statistically whether there is any relationship between two categorical variables (with two levels), Fisher’s exact test (two-tailed) was used, with type I error rate set at 5 per cent. The relationship between two continuous variables was examined using linear regression and significance was tested with ANOVA.

3. RESULTS
The dataset with taxonomic proposals between 1950 and 2007 included 747 proposals to change the taxonomic rank of species and subspecies, of which 448 were splits (60.0%) and 299 were lumps (40.0%). From the 1950s, taxonomic proposals show a steady increase in the proportion of splits (figure 1). A slight decrease in the 2000s resulted from a single controversial study that proposed 17 lumps (Penhallurick & Wink 2004; see also Rheindt & Austin 2005).

Most proposals were supported by new data (76.4%; 571/747). However, proposals for splits were significantly more often supported by new data than those for lumps: 84.6 per cent (379/448) versus 64.2 per cent (192/299; \( p < 0.001 \)). Only 15.4 per cent of the proposed new species were based on re-examination of previous information. The proportion of splits supported by new data has increased steadily during the study period (figure 2). The decrease in the 2000s was caused by a single study that split 11 taxa on the basis of previous information (Cotterill 2006). There was a marked difference between the period preceding the introduction of the PSC and the subsequent period in the proportion of splits that were supported by new data. Up to and including 1983, 71.7 per cent (134/187) of all splits were supported by new data, whereas after 1983 93.9 per cent (245/261) of them were supported by new data. This difference was highly significant (\( p < 0.001 \)).

Multiple criteria were used to delimit species (table 1). Diagnosability was the most frequently applied criterion (60.5%; 452/747). Other ‘phylogenetic’ criteria were used less often: monophyly (6.2%; 46/747) and exclusive coalescence (2.5%; 19/747). Of all splits, 83.7 per cent (375/448) were wholly or partially based on phylogenetic criteria, but only 36.6 per cent (164/448) were exclusively based on phylogenetic criteria. Consequently, 63.4 per cent of all splits did not result from the application of phylogenetic criteria only.

The proportion of splits and lumps differed between taxonomic criteria (table 1). Exclusive coalescence (94.7%), monophyly (89.1%) and adaptive zone (84.3%) most often led to the separation of species.
Degree of difference was the criterion that least often led to the separation of species (52.8%). Exclusively phylogenetic criteria more often led to splits than exclusively non-phylogenetic criteria (76.6 versus 31.5%; \( p < 0.001 \)). Overall, only 10.3 per cent (46/448) of the splits were based on reinterpretations of previous information using phylogenetic criteria.

In families with more than 60 species (\( n = 43 \)), the relative increase in species numbers was significantly related to both the proportion of polytypic species (\( r^2 = 0.009, F = 4.479, p < 0.05 \); figure 3a) and the mean number of subspecies per polytypic species (\( r^2 = 0.148, F = 7.138, p < 0.05 \); figure 3b). A large proportion of the species that were elevated from the subspecies to species level during the study period were originally described as species but subsequently downgraded to subspecies rank and included in a polytypic species (77.9%; 349/448).

### 4. DISCUSSION

#### (a) Increasing numbers of species: epistemological or data-driven?

The taxonomic inflation thesis is based on the claim that most newly recognized species result from reinterpretations of species limits based on PSCs, rather than new discoveries, and implies that the increase in species numbers is more epistemological than data-driven (Isaac et al. 2004). The results of the present study are inconsistent with these views and provide strong evidence that the increase in the number of bird species is primarily data-driven. A very high proportion of splits was based, at least in part, on new taxonomic information. The recognition of additional species resulted from the application of several criteria, not just the diagnosability criterion of the PSC. Furthermore, the increase in the number of recognized species predates the introduction of the PSC (Cracraft 1983) by at least two decades. These findings demonstrate that increases in species numbers do not primarily result from reinterpretations of previous data owing to reclassifications under the PSC.

The results of this study are in agreement with those of Padial & de la Riva (2006), who noted that increases in species numbers of amphibians preceded the introduction of PSCs and that the influence of the PSC is much less than presumed by Isaac et al. (2004). These results provide support for the view that species-level taxonomy, in general, is data-driven (Padial & de la Riva 2006; Dubois 2008). However, this study is the first to quantify the importance of new taxonomic information and the relative contribution of different taxonomic criteria. The finding that the increase in species is accompanied by an increase in taxonomic knowledge indicates that the pejorative term ‘taxonomic inflation’ is inappropriate.

#### (b) The impact of past taxonomic biases on recent increases in species

Proponents of the taxonomic inflation thesis have argued that charismatic groups and groups that are easy to study would show greater increases in species numbers than other groups. According to this view, unequal taxonomic activity (among taxonomic groups and geographical regions) would introduce biases in comparative studies (Isaac et al. 2004). The present study, however, did not find a bias towards charismatic groups. Instead, this study indicates that recent increases in the number of taxonomically recognized bird species may be a result of past biases in the application of the polytypic species concept.

A point often overlooked in discussions about taxonomic instability is that, in the 1900s to 1940s, many thousands of bird species have been downgraded to subspecific rank and combined into large variable polytypic species (Haffer 1992). Thus, whereas 18 939 species of birds were recognized in 1909 (Sharpe 1909), only 8590 species were recognized by 1951 (Mayr & Amadon 1951), a reduction of 55 per cent in just over 40 years. Most of these rearrangements were made...

---

Table 1. Differences between taxonomic criteria in the proportion of splits and lumps in taxonomic studies of birds, 1950–2007.

<table>
<thead>
<tr>
<th>criterion</th>
<th>sample</th>
<th>lump</th>
<th>split</th>
</tr>
</thead>
<tbody>
<tr>
<td>all hypotheses</td>
<td>747</td>
<td>299 (40.0%)</td>
<td>448 (60.0%)</td>
</tr>
<tr>
<td>reproductive isolation</td>
<td>232</td>
<td>77 (33.2%)</td>
<td>155 (66.8%)</td>
</tr>
<tr>
<td>diagnosability</td>
<td>452</td>
<td>100 (22.1%)</td>
<td>352 (77.9%)</td>
</tr>
<tr>
<td>degree of difference</td>
<td>216</td>
<td>102 (47.2%)</td>
<td>114 (52.8%)</td>
</tr>
<tr>
<td>adaptive zone</td>
<td>70</td>
<td>11 (15.7%)</td>
<td>59 (84.3%)</td>
</tr>
<tr>
<td>monophyly</td>
<td>46</td>
<td>5 (10.9%)</td>
<td>41 (89.1%)</td>
</tr>
<tr>
<td>exclusive coalescence</td>
<td>19</td>
<td>1 (5.3%)</td>
<td>18 (94.7%)</td>
</tr>
<tr>
<td>PSC-based criteria only(^a)</td>
<td>214</td>
<td>50 (23.4%)</td>
<td>164 (76.6%)</td>
</tr>
<tr>
<td>non-PSC-based criteria only(^b)</td>
<td>165</td>
<td>113 (68.5%)</td>
<td>52 (31.5%)</td>
</tr>
<tr>
<td>both PSC and non-PSC criteria</td>
<td>264</td>
<td>53 (20.1%)</td>
<td>211 (79.9%)</td>
</tr>
<tr>
<td>taxonomic rationale not stated</td>
<td>104</td>
<td>83 (79.8%)</td>
<td>21 (20.2%)</td>
</tr>
</tbody>
</table>

\(^a\)Diagnosability, monophyly, exclusive coalescence.

\(^b\)Degree of difference, adaptive zone, reproductive isolation.
species (Minelli 1993) or have been applied only rarely (e.g. typic species and subspecies have been much less popular been applied rigorously. In other taxonomic groups, poly-

...birds and mammals, the polytypic species concept has show a greater increase in species numbers than groups with a lower degree of polytypy (figure 3). I suggest that the observed trend in species numbers may have been caused by past biases in the application of the polytypic species concept. Groups with a higher degree of polytypy required a relatively greater degree of taxonomic activity since the mid-twentieth century to overturn a higher number of lumps in these groups during the first half of the twentieth century. This view has two important corollaries. First, apparent biases in taxonomic activity may have little to do with the charisma or ease of study of these groups. Second, the elevation of subspecies to species rank in certain vertebrate groups actually helps to reduce taxonomic biases.

A similar line of reasoning may be applied to biodiver-
sity hotspots. Isaac et al. (2004) suggested that higher taxonomic activity in hot spots would make distinctive populations in these areas more likely to be designated as species and make hotspots appear even hotter. However, most hotspots are rich in islands or isolated mountain ranges (Myers et al. 2000), and such areas are known to contain higher numbers of allopatric populations, which are generally designated as different subspecies (e.g. Mayr & Diamond 2001; Phillimore & Owens 2006). If the effect of the polytypic species concept has been greater in hotspots than in other areas, it is to be expected that a relatively higher number of valid species has been suppressed in hotspots. Thus, higher levels of taxonomic activity in hotspots would be justified. A greater increase in species numbers in hotspots than in other areas would reduce a previous taxonomic bias against hotspots.

(c) Newly recognized species are not pseudoreplicates
Isaac et al. (2004) viewed newly recognized species as pseudoreplicates of other species rather than as valuable entities for ecological research. The idea that newly recog-
nized species are pseudoreplicates, and hence inappropriate units in comparative studies, is based on the assumption that newly recognized species are ecologically identical to their close relatives. However, taxonomic studies examined for the present study often revealed previously unrecognized patterns of differentiation. Many newly recognized species of birds differed from their close relatives in traits such as habitat preference (Freitag & Robinson 1993), diet (Perrin 2005), moults (Rohwer & Manning 1990), migratory behav-
iour (Outlaw et al. 2003), breeding system (Zimmer & Whittaker 2000), sexual dimorphism (Buckley & Buckley 2004), body size (Clark & Banks 1992) and flight ability (Kennedy & Spencer 2000). Thus, rather than undifferen-
tiated lookalikes, newly recognized species are often morphologically and ecologically distinct, and merit separate treatment in comparative studies.

(d) Taxonomic progress
Isaac et al. (2004) suggested that subspecies that have been raised to species rank do not represent ‘real value’. In fact,
their use of the term taxonomic inflation implies that new, finer-grained species taxonomies are ‘less value for money’ than traditional, more broadly circumscribed species limits. Based on the results of the present study, I argue that the growth of species numbers results from progress in taxonomy and benefits other disciplines. First, the increase is primarily a result of new empirical findings and thus reflects an increase in knowledge. Second, in many cases, the recognition of additional species represents the overturn of poorly documented lumps. The taxonomic limits of many species now finally have an empirical basis. Third, the dissection of polytypic species reduces biases towards groups in which species have been less rigorously combined into polytypic species. Fourth, newly recognized species often reveal previously overlooked diversity in life-history traits. As a result, newly recognized species add precision to comparative studies. Refined species taxonomies may inform biogeographic studies of microendemic biota (Wilmé et al. 2006), enable more precise estimation of cladogenesis (e.g. Tobias et al. 2008), inform the study of adaptation (Zink & McKittrick 1995) and help to identify conservation priorities (Daugherty et al. 1990; Cracraft et al. 1998). Even if taxonomic change has undesirable practical consequences (e.g. taxonomic and nomenclatural instability), end-users of taxonomy, such as ecologists and conservationists, have good scientific reasons to support clarification of species limits.

I am grateful to J. Martin Collinson, José M. Padial and David M. Watson for providing valuable comments on the manuscript.

REFERENCES

tree.2005.04.001
Cotterill, F. P. D. 2006 Taxonomic status and conservation importance of the avifauna of Katanga (south-east Congo Basin) and its environs. Ostrich 77, 1–21.
Groves, C. 2001 Why taxonomic stability is a bad idea, or why are there so few species of primates (or are there?). Evol. Anthropol. 10, 192–198. (doi:10.1002/evan.10005)


