A revision of bird skin preparation aimed at improving the scientific value of ornithological collections

Short title: Improving the preparation of bird skins for scientific use

José Carrillo-Ortiz[1], Santi Guallar[1], Jessica Martínez-Vargas[1,2], Javier Quesada[1]

[1] *Department of Vertebrates, Museu de Ciències Naturals de Barcelona – Natural Science Museum of Barcelona, Passeig Picasso s/n, 08003 Barcelona, Spain.*

[2] *Myrmex. Serveis Tècnics a les Ciències Naturals S.L. Carrer de Bailèn, 148 3r 1a, 08037 Barcelona, Spain.*

Corresponding author:

Javier Quesada, Department of Vertebrates, Museu de Ciències Naturals de Barcelona – Natural Science Museum of Barcelona, Passeig Picasso s/n, 08003 Barcelona, Spain.

Email: jquesadal@bcn.cat

Wordcount (without references): 3744

Abstract

The methods used to preserve bird skins in museums have a potentially crucial impact on the feasibility and use of these specimens as a source of biological knowledge, although this subject is rarely broached. Study skins of birds are usually prepared with folded wings and straight legs to facilitate storage in the collection; yet, this method can hamper the measurement and examination of certain important features such as wing-feather moult. To make consultation easier for ornithologists, alternative preparation methods such as the splitting of wings and tarsi from the rest of the animal have been proposed by curators. Our aim was to study whether or not preparing bird specimens with spread limbs makes consultation simpler. First, we used two different methods to prepare two specimens each of two common European passerine species: 1) 'traditional' (folded wings and straight tarsi) and 2) 'spread' (limbs spread on one side of the body). Then, we asked 22 experienced ornithologists to identify moult limits and take three biometric measurements (wing chord, length of the third primary feather and tarsus length) from all four specimens. Subsequently, we asked which preparation method they preferred for obtaining data. The 'spread' preparation was preferred for moult, third primary feather length and tarsus length, whilst the 'traditional' preparation was preferred for wing chord. Data obtained from the folded and spread preparations were very highly repeatable within each method but only moderately to highly repeatable between methods. One of the handicaps with the 'spread' preparation is the increase in storage space required, a factor that should be taken into account before it is employed. Nevertheless, this specimen preparation technique can greatly facilitate

consultation and therefore improve the scientific value of ornithological collections.

Keywords

biometry, bird study skin, moult, natural history museums, taxidermy, zoological collections.

Introduction

Natural history collections are repositories of an important biological heritage and have traditionally played a pivotal role in improving knowledge of the natural sciences. For instance, zoological collections are crucial in the acquisition of and increase in knowledge of bird ecology, evolution and systematics.[1,2] The study of bird skins in zoological collections enables us to split subspecies of bird species and identify fundamental biological aspects such as differences in sex and age, moult limits and biometry,[3,4] phylogeny[5,6] and evolutionary processes.[7,8] However, the methods used for preparing bird skins may affect the feasibility and use of these specimens as a source of biological information.

Bird specimens are traditionally prepared with folded wings and straight legs (Figure 1A) due, historically, to the need for quick and efficient preparation in the field, for specimens to be robust enough to withstand shipping, and for space-efficient packing and storage.[9–12] This last is the final goal of the specimen in being amenable to researchers for consultation. This method of preparation, however, does not necessarily take into account the use of specimens for scientific purposes, and may in fact hamper studies of bird biometrics, aging, moult and the identification of key plumage characteristics. For example, the examination of secondary feathers in specimens prepared using the ‘traditional’ method is problematic,[13] and measuring the wing chord (frequently used for determining age, species or subspecies) is complicated if specimens have folded wings and increases the risk of specimen damage.[14,15] Examining feather moult in specimens with folded wings can be challenging or even almost impossible, as many feathers remain hidden below the tightly

84 packed feathers that overlay them.[16] In addition, the evaluation of feather
85 colouration and wear is difficult in specimens with folded wings.[17] Therefore, the
86 correct determination of certain commonly used wing parameters (e.g. moult,
87 biometric measurements and colouration) is impossible for certain parts of the
88 wing unless folded-wing specimens are mutilated.[18,19,20]

89 To facilitate research and, in particular, biometric and moult studies using
90 bird specimens, a variety of alternative preparation methods have been
91 proposed by museum curators and employed, including the idea of severing
92 one extended wing.[16] In addition to facilitating the collection of moult and
93 biometric data, these proposals aim to facilitate taxonomic studies and the
94 evaluation of wing wear, shape and colour.[18,19] However, to date few
95 quantitative studies have ever attempted an objective evaluation of zoological
96 preparation methods,[21] despite the widespread use of collections and the calls
97 made to use quantitative methods to improve the protocols used in zoological
98 preparation.[21,22] Here, we publish for the first time a quantitative approach that
99 objectively assesses and compares two preparation methods for bird
100 specimens. This novel study involved the cooperation of field ornithologists and
101 museum curators, and was designed to improve the ease with which relevant
102 biological information can be obtained from ornithological collections. Our aim
103 was to compare the ease of consultation of bird specimens prepared using two
104 different methods: folded wings and straight tarsi ('traditional' *preparation*) with
105 limbs spread out on one side of the body ('spread' preparation). We requested
106 expert passerine ringers to take biometric and moult data from two specimens
107 of a small-sized passerine, the Eurasian blackcap (*Sylvia atricapilla* Linnaeus,
108 1758), and two of a medium-sized passerine, the song thrush (*Turdus*

philomelos Brehm, 1831), of which one specimen of each was prepared using the 'traditional' method and one using the 'spread' method. We also evaluated the consistency of the biometric measurements obtained from both methods by assessing the repeatability within each method (Intramethod Ri, for 'traditional' and 'spread' preparations) and between methods (Intermethod Ri, 'traditional' *versus* 'spread' preparations).[23] Finally, and given that the 'spread' preparation requires greater storage space, we compare the storage space required for specimens prepared using these two methods.

Material and methods

We prepared four passerine specimens, two of the Eurasian blackcap (voucher numbers: MZB 2017-0505 and MZB 2017-0508) and two of the song thrush (voucher numbers: MZB 2017-0504 and MZB 2017-0507). These species are very frequent in the migration period of the **northeastern** part of the Iberian Peninsula, they are frequently ringed by ornithologists based there,[24] and they also are regularly donated to the Natural Science Museum of Barcelona (MCNB)**,** added to the fact that they represent two common species of small and medium body size birds, respectively, explain why they were **chosen for** the present study**.** One specimen of each species was prepared using the 'traditional' method (folded wings and straight tarsi, see Figure 1A) (MZB 2017-0507 and MZB 2017-0508) and one using the 'spread' method (limbs spread on the left-hand side of the body, see Figure 1B) (MZB 2017-0504 and MZB 2017-0505).

Museum specimens dry out when frozen and stored for long periods (i.e years or even decades), especially when they are not properly insulated, which

134 can affect both the ease of preparation and their final quality (i.e. probability of
135 breakage and their flexibility).[22] To ensure similar freezing conditions, we only
136 used specimens that had been collected in autumn 2015 in the NE Iberian
137 Peninsula as accidental deaths and subsequently donated to the MCNB by
138 wildlife recovery centres or ornithologists. Since we were interested in
139 identifying wing-moult limits, we selected individuals hatched during the
140 previous breeding season (EURING 5 or SY) that had undergone a partial moult
141 during their first year.[18] In addition, we only selected males to avoid possible
142 differences caused by the confounding effects of plumage dichromatism.[18]
143
144 *Measurement of observer preferences*
145 We carried out a preference survey amongst 22 ornithologists, all
146 experienced passerine ringers but not all with previous experience of working
147 with zoological collections. However, given that experience with bird skins may
148 play an important role in assessing the feasibility of alternative preparation
149 methods, we took into account ornithologists' experience (or lack of) with bird
150 skins in order to obtain a better understanding of the advantages/disadvantages
151 of each preparation method when extracting the required data. All of these
152 expert ringers had previous experience with these two species or with similarly
153 sized species as all fulfilled the EURING-based standards for ringers (e.g.
154 Catalan ringers must ring at least 500 individuals of 50 species to become an
155 'expert ringer' and have more than two years of experience). Nevertheless, our
156 aim was not to evaluate their expertise with these species but, rather, determine
157 which preparation technique made consultation easier, regardless of the
158 species considered and bearing in mind that they were a homogeneous group

of experienced observers. For each specimen, these expert ringers were asked to fill out a moult card[25] and take three biometric measurements:[15,26] maximum wing chord, length of the third primary feather (hereafter 'P3 length') and tarsus length, all information that is frequently sought by museum-based studies,[15,19,27] and it is known by ringers of Spanish Ringing schemes (SYLVIA, PASSER).[26,28] Subsequently, they were requested to judge for each measurement whether one preparation was easier to handle than the other by scoring each as follows: 1: slightly easier; 2: easier; and 3: markedly easier. The small sample size used in this study can be explained by the fact that we were interested in assessing the ease of consultation of each preparation method, and by the limitations of the collection. One observer only measured the song thrush but not the Eurasian blackcap so sample sizes differ between species (i.e. $N_{\textit{Turdus philomelos}}$ = 22; $N_{\textit{Sylvia atricapilla}}$ = 21).

We calculated the mean score and 95% bootstrapped intervals (size of the posterior sample = 4000).[29] Finally, to compare the scores, we used the Wilcoxon rank-sum (also known as the Mann-Whitney U), which is employed when a categorical scoring of preference is similar to Likert scores[30]. All analyses were carried out in R (version 3.6.3).[31]

Repeatability of measures within and between methods

An important assumption for validating the scientific use of the 'spread' preparation method is that the wing measurements obtained from this alternative model are as repeatable as those obtained from a folded wing. For this reason, we evaluated the repeatability of the biometric measurements used in this study for both species. Repeatability (Ri) is a statistical measurement of

the reliability of repeated measurements of a single characteristic of the same specimen, and is used to quantify statistically the consistency of equivalent measurements of a particular object.[32] Additionally, it can be used as an indicator of the consistency of methods.[23,32,33] Repeatability values lie **in** the range 0–1, where the measurement error calculation equals 1-Ri. Thus, a high repeatability value (>0.70) indicates a consistent measurement and a low measurement error, although values ≥0.90 (very high repeatability) are generally sought in the case of biometric measurements.[34,35]

To assess the consistency of biometric measurements using the same method (Intramethod Ri), we measured the wing chord, P3 length and tarsus length from the same body side (left-hand) in a set of specimens from the MCNB collection obtained between 2009 and 2018 (traditional method: $N_{Turdus\ philomelos}$= 10; $N_{Sylvia\ atricapilla}$= 8; spread method: $N_{Turdus\ philomelos}$= 10; $N_{Sylvia\ atricapilla}$= 8). To assess the repeatability between methods (Inter-method Ri) we compared both wing and tarsus measurements from the same individual, prepared following the traditional method (right-hand body side) and the spread method (left-hand body side) ($N_{Turdus\ philomelos}$= 10; $N_{Sylvia\ atricapilla}$= 8). All **of** the measurements were taken by the same observer (SG).

To calculate the repeatability, we used the rptR package in R which is based on the estimation of repeatabilities within the framework of the mixed effects model. This approach uses random-effect predictors to estimate variances at different hierarchical levels.[36]

Finally, in order to analyse the trade-off using the ‘spread’ preparation method (i.e. bearing in mind its drawbacks regarding storage space), we assessed the percentage of storage loss with the alternative preparation

method compared to the traditional one by counting the number of specimens prepared according to each method that could be accommodated in a standard skin storage box (400 mm long by 300 mm wide by 120 mm deep; Euro container EU4312L Box®, Rajapack™) in the MCNB.

Results

The preliminary analyses showed that results were consistent between and within species. Thus, we pooled the scores for the two species and report here the results for the whole sample. We found that the 'spread' preparation was preferred for scoring moult limits and for measuring P3 and tarsus lengths. The preference for the 'spread' format when measuring P3 length was less than for the two other categories, which provided clearly higher scores. Conversely, the 'traditional' preparation method was preferred for measuring chord length (Table 1, Figure 2).

When we considered whether or not observers had previous experience working with collections, similar results in method preference were obtained. There were no differences in preference between experienced observers and non-experienced observers (all Ps>0.11), except for tarsus length ($W=133.74$; $P<0.05$). However, all observers chose the spread method for this trait ($X_{Experienced}=1.07$; $X_{Non\text{-}experienced}=2.31$).

Repeatability analyses showed that all measurements had high (>0.70) or very high (>0.90) repeatabilities; the 'spread' preparation had similar or higher repeatabilities than the folded preparation for both species (Table 2). However, we obtained moderate to high repeatability when we compared the two

methods. The Intermethod Ri was higher in the larger species (i.e. song thrush) than in the smaller one (i.e. Eurasian blackcap).

The results showed that for both species there was an increase in the storage space required when specimens were prepared using the 'spread' method. We could store 17 song thrushes and 19 Eurasian blackcaps prepared using the 'folded' method in the same space (Euro Box®), compared to 12 song thrushes and 14 Eurasian blackcaps prepared according to the 'spread' method. This implies a increase by 42% and 36% in storage space requirements by due to 'spread' method for the two species, respectively.

Discussion

The method used to prepare bird specimens destined for zoological collections has an important effect on how our biological heritage is used as a source of scientific knowledge. Regardless of the experience of the observer with zoological collections, the 'spread' preparation of bird skins was notably preferred when examining moult limits and measuring tarsus length, and even appeared to provide a slight advantage in the measurement of P3 length. This preparation method allows wing moult to be scored without damaging specimens since researchers do not have to unfold feathers to expose those lying beneath. This alternative technique is also likely to facilitate other biometric measurements since, for example, measurement of the wing chord is not hindered as it only requires the unfolding of one wing. Unlike the preparation of extended wings severed from the body, a method used in museums and universities such as the Burke Museum in Seattle,[37] the Cleveland Museum of Natural History,[38] the Australian Museum (L. R. Tsang, personal

258 communication), the University of Puget Sound (P. H. Wimberger & G. Shugart,
259 personal communication) and the National Museum of Ireland (P. Viscardi,
260 personal communication), the ‘spread’ preparation allows researchers to obtain
261 moult and biometric information from the whole specimen. Furthermore, the
262 ‘spread’ preparation also displays the whole wing, which simplifies the
263 inspection of feather variation, form and colour, and is ideal for artists,[20]
264 illustrators of field guides[39] and for studies of plumage coloration.[7]
265 Both preparation methods provide consistent biometric measurements, as
266 the very high repeatability values for each method demonstrate. Nevertheless, a
267 relevant question here is how comparable or close to one another (i.e. precise)
268 are the data obtained from the alternative ‘spread’ method of preparation and
269 those obtained using the ‘traditional’ method, in addition to how accurate those
270 measurements are with regard to the actual morphometric in question. In the
271 context of this study, precision is important given that the ‘spread’ method could
272 negatively affect long-term studies of traits in collections that have been
273 prepared using the ‘traditional’ method. Our results suggest that measurements
274 resulting from the ‘spread’ method are partially comparable to those obtained
275 from the ‘traditional’ one. However, it is worth highlighting two aspects of this
276 question. First, since the idea is to unfold only one wing (left-hand one), the
277 right-hand wing can still be measured in the same way as in the ‘traditional’
278 preparation. Secondly, to evaluate the Intramethod Ri, we would need to assess
279 at least two independent measurements of the *same* wing after preparation by
280 both the ‘spread’ and ‘folded’ techniques, which is a drawback. The only
281 possibilities for doing so are 1) to compare the same bird before and after
282 preparing the wing/tarsus (however, previous studies have shown that

283 significant biometric differences between fresh and prepared specimens occur
284 in some traits[15]); and 2) to compare both wings/tarsi from the same sides. We
285 used this second approach, even though wing/tarsus asymmetry could mask
286 the repeatability of results. Thus, the measurement error obtained in our
287 analysis results from the sum of the asymmetry between wings/tarsi and the
288 actual measurement made by the observer. Nevertheless, to overcome the fact
289 that this new 'spread' design could compromise the utility of the measurements
290 in longitudinal studies, correction factors could be developed,[15,40] which would
291 allow for each species' morphometric correction from a 'spread' specimen to a
292 'folded' specimen.
293 Although this alternative preparation **provides** new opportunities for
294 obtaining biometric data more effectively (particularly important for moult and
295 coloration analysis), this method may interfere with the availability of other
296 scientific data (e.g. studies of bilateral asymmetry). This issue, however, is a
297 recurrent trade-off in collection management and is also of concern when
298 preparing a single wing (**as** in other museums) or a particular element from a
299 specimen (e.g. skeleton).[41]
300 Spread-wing preparation entails important advantages in research, and
301 also has advantages for collection management: it reduces the risk of damage
302 to prepared specimens when consulted by experts, and having a single voucher
303 for each specimen avoids double labelling and minimizes errors, which is
304 particularly important in large collections. However, it should be acknowledged
305 that spread-wing specimens also represent challenges for collection
306 management. One drawback of the 'spread' method is that it increases the
307 chances that the specimen will be damaged since unfolded wings are more

fragile. Another shortcoming is that it requires more storage space (36–42% in our study) than the ‘traditional’ preparation, which can be a severe limiting factor in the case of museums under increasing storage pressure.[9] There are, however, two ways of minimising this latter impact: 1) specimens can be prepared with their unfolded limbs always on the same side, and 2) specimens can be placed in drawers with an imbricated lay-out, that is, with the body of one specimen lying on the unfolded wing of the adjacent specimen. This second strategy could significantly increase the risk that specimens will be damaged and so precautions should be taken to minimize its impact. It is also important to acknowledge that the ‘spread’ method could be particularly problematical when preparing larger taxa, even if these species may also benefit from this preparation technique. Another disadvantage associated with the ‘spread’ preparation is that it seems to require slightly more material during the preparation work (e.g. needles and wire) and more time (20–25%, JC-O&JM-V, pers. obs.). Despite these considerations, we still lack a thorough cost-benefit analysis of the storage issues associated with the implementation of the ‘spread’ preparation. Moreover, the ‘spread’ preparation might end up being more time-consuming as more time will be needed to repair specimens in the future (which are more likely to be damaged), to retrieve or replace specimens that overlap each other, and to pack or unpack specimens for shipping. Nevertheless, as we have seen, this method provides several functional advantages. In general terms, it is important to acknowledge the existence of a compromise between the ease of use of specimens (which is vital for their scientific usefulness), their long-term preservation so they will be available to

future researchers, and storage practicalities in museums. A balanced perspective is therefore encouraged.

Natural science museums and the repositories of research collections (e.g. research centres and universities) worldwide strive to increase the value of their collections by increasing the scientific use of the biological heritage they harbour.[42] This study highlights how research in zoological preparation can help to improve preparation techniques, and so optimise the scientific use of our accumulated ornithological heritage. We have shown that the 'spread' preparation facilitates the consultation of bird skin collections and also reduces the risk of damage to prepared specimens when handled for data acquisition. Although this format could result in more damage to specimens when they are retrieved and returned to their storage drawers, the fact that it actually reduces the risk of damage when measurements are carried out constitutes the key advantage of this style of preparation. We recommend the implementation of this preparation method in bird skin collections if storage space is not an important limiting factor.

Acknowledgements

We would like to thank the Catalan Ornithological Institute (ICO) for permission to carry out the survey amongst expert ringers during the Delta Birding Festival 2016. We would also like to thank Laura Roqué, a preparer of zoological material, for her support. Mike Lockwood, Fernando Mateos and Hannah Findlay made useful comments on an earlier version of this manuscript. We finally would like to thank Jim Reynolds and two anonymous referees for their comments on this manuscript.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclose the receipt of the following financial support for the research, authorship, and/or publication of this article: this work was financially supported by the Natural Science Museum of Barcelona (JQ) and the Ministry of Economy and Competitiveness (grant number PTA 2014-09302-I; JC-O).

ORCID iD

Santiago Guallar https://orcid.org/0000-0002-1121-1968

Jessica Martínez-Vargas https://orcid.org/0000-0003-4689-5228

Javier Quesada https://orcid.org/0000-0002-6010-8473

Author contributions

Study conception and design: JC-O, JQ; development of methodology: JC-O, SG, JQ; data collection: JC-O, JQ, SG, JM-V; formal analysis of the study data: SG; drafting of the manuscript: JM-V, JC-O, SG, JQ; critical revision of the manuscript: JC-O, JM-V, SG, JQ; research supervision and project administration: JQ.

References

1. Suarez AV and Tsutsui ND. The value of museum collections for research and society. *BioScience* 2004; 54: 66–74.
2. Holmes MW, Hammond TT, Wogan GO, et al. Natural history collections as windows on evolutionary processes. *Mol Ecol* 2016; 25: 864–881.
3. Perrins CM and Snow DW. *The complete birds of the western Palearctic*. Oxford: Oxford University Press, 1998.
4. Engelmoer M and Roselaar CS. *Geographical variation in waders*. Dordrecht: Springer Science & Business Media, 2012.
5. Burns KJ. A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (Thraupidae): the role of female versus male plumage. *Evolution* 1998; 52: 1219–1224.
6. Stoddard MC and Prum RO. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. *Am Nat* 2008; 171: 755–776.
7. Hill GE and McGraw KJ. *Bird coloration volume 2: function and evolution*. Cambridge: Harvard University Press, 2006.
8. Dunn PO, Armenta JK and Whittingham LA. Natural and sexual selection act on different axes of variation in avian plumage color. *Sci Adv* 2015; 1: e1400155.
9. Amadon D. The use of scientific study skins of birds. *Curator* 1958; 1: 77–80.
10. Fuller T. Storage methods for taxidermy specimens. *WAAC Newsletter* 1992; 14: 18–23.
11. Ramotnik CA. Handling and care of dry bird and mammal specimens. *Conserve O Gram* 2006; 11/9: 1–5.

12. Eck S, Fiebig J, Fiedler W, et al. *Measuring Birds Vögel Vermessen*. Wilhelmshaven: Deutsche Ornithologen-Gesellschaft, 2011.
13. Rohwer S and Rohwer VG. Primary molt in Gruiforms and simpler molt summary tables. *PeerJ* 2018; 6: e5499.
14. Kelm H. Beitragzurmethodik des flügelmessens. *J Ornith* 1970; 111: 482–494.
15. Jenni L and Winkler R. The feather-length of small passerines: a measurement for wing-length in live birds and museum skins. *Bird Study* 1989; 36: 1–15.
16. Winker K. Obtaining, preserving, and preparing bird specimens. *J Field Ornithol* 2000; 71: 250–297.
17. Jenni L and Winkler R. *Moult and ageing of European passerines*. London: Academic Press, 1994.
18. Svensson L. *Identification guide to European passerines*. 4th ed. Stockholm: British Trust for Ornithology, 1992.
19. Pyle P. *Identification guide to North American birds: a compendium of information on identifying, ageing, and sexing "near-passerines" and passerines in the hand*. Slate Creek Press, 1997.
20. Dwight JJ. The moult of the North American Tetraonidæ (quails, partridges and grouse). *Auk* 1900; 17: 34–51.
21. Simmons JE. Collection care and management: history, theory, and practice. In: Macdonald S, Leahy HR, Witcomb A, et al. (eds) *The international handbooks of museum studies*. New Jersey: Wiley-Blackwell, 2013, pp. 221–247.

22. Roqué L, Del Canto I, Carrillo-Ortiz J, et al. Effects of prolonged freezing on the preparation quality of fresh material destined to bird collections. In: *10th International Meeting of European Bird Curators*. Paris, France, 17–19 October 2017.
23. Quesada J, Aurell-Garrido J, Gago S, et al. Measurement errors in 3D models used in osteometric data research with freeware: a test using skulls of the Algerian hedgehog (*Atelerix algirus*). *Vertebr Zool* 2016; 66: 411–418.
24. ICO. "SIOC: servidor d'informació ornitològica de Catalunya". Barcelona: Institut Català d'Ornitologia (ICO), www.sioc.cat (2020, accessed 23 October 2020).
25. Ginn HB and Melville DS. *Moult in birds - BTO guide 19*. Tring: British Trust for Ornithology, 1983.
26. Institut Català d'Ornitologia (ICO). *Oficina Catalana d'Anellament – Estàndards d'anellament*. Barcelona: Catalan Ornithological Institute, 2003.
27. Totterman SL. Random measurement error and specimen shrinkage in short-tailed shearwaters *Puffinus tenuirostris*. *Mar Ornithol* 2016; 44: 11–20.
28. Pinilla J. *Manual para el anillamiento científico de aves*. Madrid: SEO/BirdLife y DGCN-MIMAM, 2000.
29. Bååth R. *Bayesboot: an implementation of Rubin's (1981) Bayesian bootstrap, R package version 0.2.2*, https://CRAN.R-project.org/package=bayesboot (2018, accessed 31 March 2020).
30. Leon AC. 3.12 - Descriptive and Inferential Statistics. In Bellack, AS Hersen,M (eds) Comprehensive Clinical Psychology. Pergamon: Elsevier, 1998, pp 243-285

455 31. R Core Team. *R: a language and environment for statistical computing,*
456 https://www.R-project.org/ (2018, accessed 31 March 2020).
457 32. Quesada J and Senar JC. Comparing plumage colour measurements
458 obtained directly from live birds and from collected feathers: the case of the
459 great tit *Parus major*. *J Avian Biol* 2006; 37: 609–616.
460 33. Figuerola J and Senar JC. Measurement of plumage badges: an evaluation
461 of methods used in the Great Tit *Parus major*. *Ibis* 2000; 142: 482–484.
462 34. Harper DGC. Some comments on the repeatability of measurements.
463 *Ringing Migr* 1994; 15: 84–90.
464 35. Senar JC. La medición de la repetibilidad y el error de medida. *Etologuía*
465 1999; 17: 53–64.
466 36. Nakagawa S and Schielzeth H. A general and simple method for
467 obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol*
468 *Evol* 2013; 4: 133–142.
469 37. Burke Museum in Seattle. Ornithology at the Burke Museum
470 www.burkemuseum.org/collections-and-research/biology/ornithology (2020,
471 accessed 9 October 2020).
472 38. Cleveland Museum of Natural History. Collection & Database,
473 www.cmnh.org/ornithology/collection-database (2020, accessed 14 October
474 2020).
475 39. Rasmussen PC. On producing *Birds of South Asia*. *Indian Birds* 2005; 1:
476 50–56.
477 40. Winker K. Specimen shrinkage in Tennessee warblers and "Traill's"
478 flycatchers. *J Field Ornithol* 1993; 64: 331–336.

41. Cato PS and Jones C. *Natural history museums: directions for growth*. Lubbock: Texas Tech University Press, 1991.

42. Collar NJ, Fisher CT and Feare CJ. Why museums matter: Avian archives in an age of extinction. *Bulletin of the British Ornithologists' Club* 2003; 123A: 1-360.

Table 1. Wilcoxon rank-sum value analysis used to evaluate preferences by experienced ornithologists for 'folded' or 'spread' bird skin preparations. Negative values between means imply a preference for the traditional ('folded') preparation.

	N	W	Difference of mean	P-value
Moult limits	**43**	**1682.00**	**2.26**	*******
Wing chord	**43**	**1370.00**	**-2.19**	*******
P3 length	**43**	**1170.00**	**0.63**	*****
Tarsus length	**42**	**1527.00**	**1.80**	*******

**** P<0.05 ** P<0.01 *** P<0.001***

Table 2. Repeatability (Ri) values of three biometric measures within each method ('folded' and 'spread' preparation) and a comparison of both methods ('folded' *vs* 'spread' preparation). Ri>0.70 is considered as a high repeatability.

	WITHIN-METHOD Ri						BETWEEN-METHOD Ri		
	Folded preparation			Spread preparation			Folded *vs* Spread prep.		
	Ri	SE	P-value	Ri	SE	P-value	Ri	SE	P-value
Eurasian blackcap									
Wing chord	0.97	0.03	***	0.93	0.07	***	0.54	0.25	0.07
P3 length	0.83	0.16	**	0.81	0.17	**	0.66	0.23	*
Tarsus length	0.72	0.19	*	0.96	0.05	***	0.43	0.26	0.14
Song thrush									
Wing chord	0.54	0.23	*	0.98	0.02	***	0.53	0.24	*
P3 length	0.93	0.06	***	0.94	0.06	***	0.70	0.19	**
Tarsus length	0.93	0.07	***	0.90	0.09	***	0.72	0.19	**

** P<0.05 ** P<0.01 *** P<0.001*

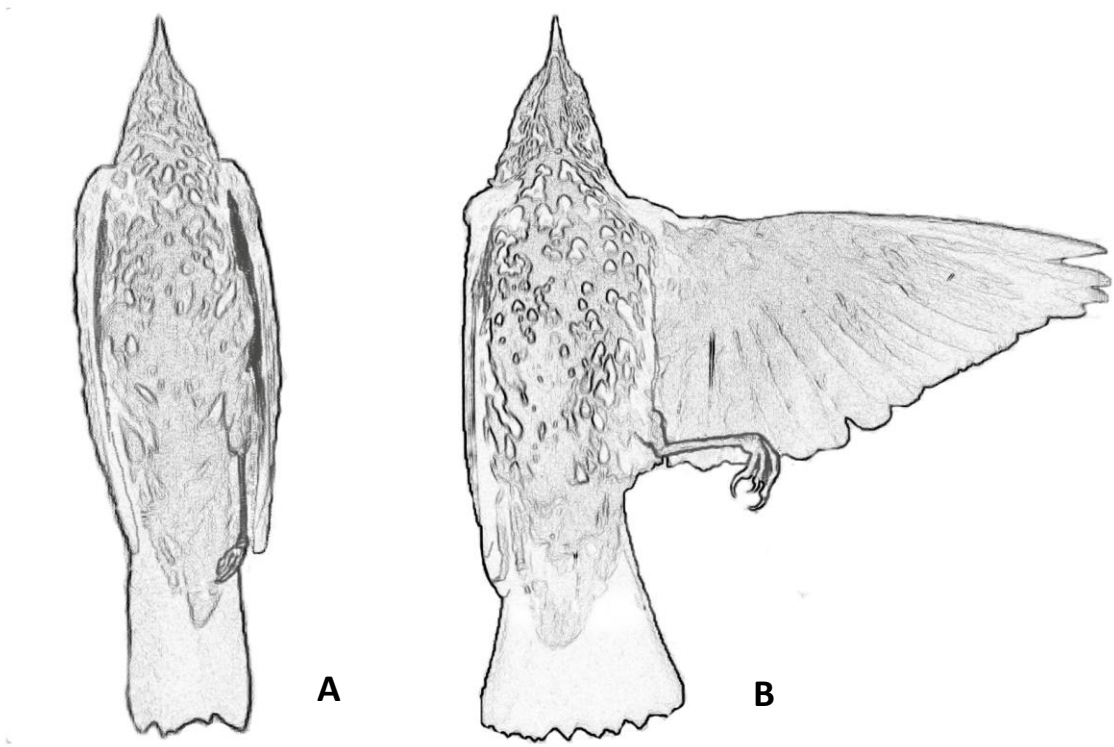

Figure 1. Example of two specimens of song thrush in ventral view prepared with the A) 'traditional' folded method (folded wings and straight tarsi) and the B) 'spread' method (spread left-hand wing and bent left-hand tarsus).

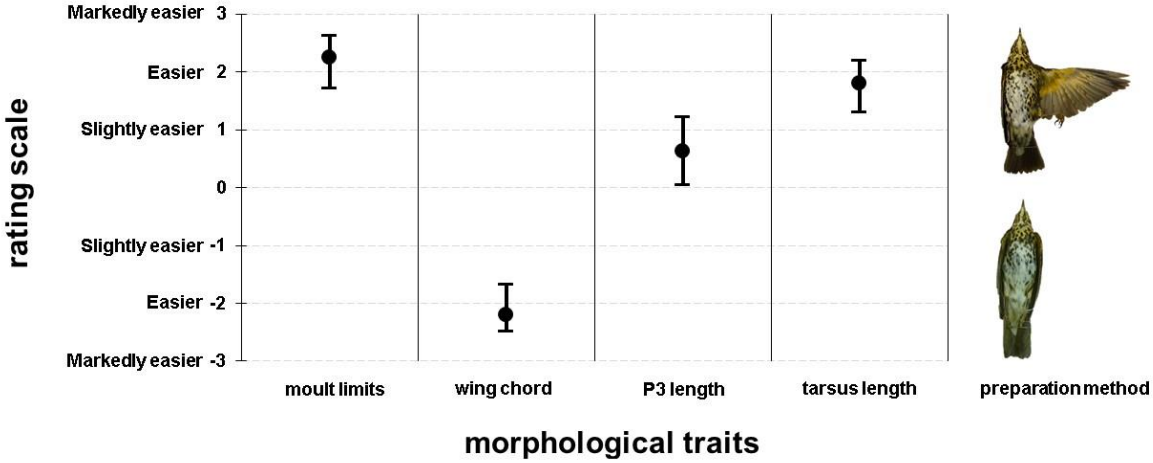

Figure 2. Ornithologists' preferences between the two specimen preparation methods: 'spread' (one unfolded wing and tarsus on one side of the body) and 'folded' (folded limbs). Mean and 95% Bayesian bootstrapped intervals (N=4000).