

## Birds that are more commonly encountered in the wild attract higher public interest online

Mittermeier, John C.; Roll, Uri; Matthews, Tom; Correia, Ricardo; Grenyer, Richard

DOI:  
[10.1111/csp2.340](https://doi.org/10.1111/csp2.340)

License:  
Creative Commons: Attribution (CC BY)

*Document Version*  
Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*  
Mittermeier, JC, Roll, U, Matthews, T, Correia, R & Grenyer, R 2021, 'Birds that are more commonly encountered in the wild attract higher public interest online', *Conservation Science and Practice*, vol. 3, no. 5, e340. <https://doi.org/10.1111/csp2.340>

[Link to publication on Research at Birmingham portal](#)

### **Publisher Rights Statement:**

Mittermeier, JC, Roll, U, Matthews, TJ, Correia, R, Grenyer, R. Birds that are more commonly encountered in the wild attract higher public interest online. *Conservation Science and Practice*. 2021:e340. <https://doi.org/10.1111/csp2.340>

### **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# Birds that are more commonly encountered in the wild attract higher public interest online

John C. Mittermeier<sup>1</sup>  | Uri Roll<sup>2</sup>  | Thomas J. Matthews<sup>3,4</sup> |  
Ricardo Correia<sup>5,6</sup>  | Rich Grenyer<sup>1</sup>

<sup>1</sup>School of Geography and The Environment, University of Oxford, Oxford, UK

<sup>2</sup>Mitrani Department of Desert Ecology, The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Beersheba, Israel

<sup>3</sup>School of Geography, Earth and Environmental Sciences, and Birmingham Institute of Forest Research, University of Birmingham, Edgbaston, Birmingham, UK

<sup>4</sup>CE3C – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group and Univ. dos Açores – Depto de Ciências e Engenharia do Ambiente, Angra do Heroísmo, Portugal

<sup>5</sup>Helsinki Lab of Interdisciplinary Conservation Science (HELICS), Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland

<sup>6</sup>Helsinki Institute of Sustainability Science (HELSUS), University of Helsinki, Helsinki, Finland

## Correspondence

Rich Grenyer, School of Geography and The Environment, University of Oxford, Oxford, OX1 3QY, UK.  
Email: richard.grenyer@ouce.ox.ac.uk

## Abstract

Large body size, the defining characteristic of “charismatic megafauna,” is often viewed as the most significant correlate of higher public interest in species. However, common, local species (many of which are not large) can also generate public interest. We explored the relative importance of body size versus local occurrence in patterns of online interest in birds using a large sample of digital human-wildlife interactions (367 million Wikipedia pageviews) that included more than 10,000 bird species and a range of cultural and geographic contexts (represented by 25 Wikipedia language editions). We compared interest in Wikipedia, as measured by pageviews, with a bird's body size and its regional observation frequency (using data from eBird.org). We found that local species (i.e., those that occur in the wild in the country responsible for the majority of a Wikipedia language edition's pageviews) attract more pageviews than global species. Both body size and observation frequency had a positive correlation with Wikipedia pageviews across languages, but eBird observation frequency explained more of the variance in pageviews on average. In a model that included both observation frequency and body size, observation frequency was a significantly better predictor of pageviews than body size in 24 of 25 languages. Our results demonstrate that the opportunity to encounter birds in the wild is a significant correlate of increased online interest in birds across multiple linguistic and geographic contexts. This relationship provides insight into why some species attract greater interest than others and emphasizes the overlooked potential of common species in conservation marketing.

## KEYWORDS

bird conservation, citizen science, conservation culturomics, eBird, geography, Wikipedia

## 1 | INTRODUCTION

Understanding why some species command greater public interest than others have long been of interest to

conservationists (Kellert, 1982; Lorimer, 2007; Macdonald et al., 2015). Identifying the biological and ecological traits of species that correlate with increased public interest can benefit conservation marketing and be used to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Conservation Science and Practice published by Wiley Periodicals LLC. on behalf of Society for Conservation Biology

identify potential flagship species (Smith, Verissimo, Isaac, & Jones, 2012; Verissimo et al., 2013). Previous studies have evaluated the role of biological traits such as coloration (Lišková & Frynta, 2013), venomousness (Roll et al., 2016), and the perceived attractiveness of species (Gunnthorsdottir, 2001) in determining public interest. While a variety of traits have been recognized as correlating with increased interest, large body size is frequently identified as the most important (Berti, Monsarrat, Munk, Jarvie, & Svenning, 2020; Clucas, McHugh, & Caro, 2008; Macdonald et al., 2015). Indeed, the widespread use of the term “charismatic megafauna” to describe species that attract high public interest suggests that large body size is often viewed as synonymous with increased public interest in the conservation community.

The importance of large body size suggests that direct interactions with animals in the wild may be relatively unimportant to determining interest in species. While large species are strongly represented in media (Clucas et al., 2008) and captive environments (Martin, Lurbiecki, Joy, & Mooers, 2014), they are often rare and infrequently encountered in the wild. In extreme examples, species such as tigers (*Panthera tigris*) and giant pandas (*Ailuropoda melanoleuca*) attract substantial public interest despite only a tiny minority of people ever observing them in the wild. Conversely, direct encounters with species in the wild do seem to drive public interest in some situations. The commonness of butterfly and bird species in the UK and Poland correlates with the attention these species receive online (Zmihorski, Dziarska-Palac, Sparks, & Tryjanowski, 2013). In Brazil, species that overlap with areas of higher human population density receive greater interest online, as measured by Google search frequency (Correia, Jepson, Malhado, & Ladle, 2016). This relationship between people's ability to directly encounter a species in the wild and their interest in it has relevance for conservation. If direct encounters are relatively unimportant, then emphasizing large and charismatic species as well as virtual encounters through various forms of media should be the top priority for engaging greater public interest. In this scenario, public interest in a species is unlikely to decline if the species becomes rarer or even goes extinct in the wild. Conversely, if direct encounters with wild animals are important for generating public interest, then approaches that emphasize direct encounters with species in the wild may be more effective for increasing public engagement in conservation. Furthermore, if public interest is closely linked to wild encounter rates then the declining numbers of wild animals (Dirzo et al., 2014) and trends of increasing urbanization and detachment from nature (Miller, 2005; Soga & Gaston, 2016) could lead to decreased interest in some species.

Conservation culturomics is a new research area that uses online digital data to investigate questions around human-nature interactions that are relevant to conservation (Ladle et al., 2016). Previous studies have used conservation culturomic methods to assess people's interest in reptiles (Roll et al., 2016) and Brazilian birds (Correia et al., 2016), and to compare seasonal patterns in biodiversity awareness (Mittermeier, Roll, Matthews, & Grenyer, 2019). Here we used these methods to assess the relationship between online interest, body size and the frequency with which people have direct encounters species in the wild. Using birds as case study, we assessed patterns of public interest (as measured by Wikipedia pageviews) across a large number of digital interactions and a range of cultural contexts (25 different Wikipedia languages). We compared Wikipedia pageview data to information on the frequency with which birds are observed in the wild derived from eBird, an online database of avian distributions. With more than 700 million records, eBird is the world's largest biodiversity-related community science project and provides occurrence data for birds around the world (Levatich & Ligocki, 2020; Sullivan et al., 2009; Wood, Sullivan, Iliff, Fink, & Kelling, 2011). Multiple studies have used eBird data to monitor patterns in avian distribution and abundance at a variety of scales (eBird, 2020).

We tested three hypotheses around the importance of direct encounters in generating online interest in birds. First, we assessed whether the geographic distribution of birds impacted their Wikipedia pageviews. If wild encounters are unimportant, we predicted that the local occurrence of bird species would have a minimal impact on their pageviews; if direct encounters are important, we predicted pageviews would be skewed toward local species (i.e., those that occurred in the wild in a given region). Second, we tested for correlations between pageviews and body size/eBird frequency. We predicted a positive correlation between pageviews and body size if indirect encounters are important and a positive correlation between pageviews and eBird frequency if direct encounters with species are important. Third, in instances where we had both body size and eBird frequency data, we tested the relative importance of both variables in predicting Wikipedia pageviews in a combined model. Here, we predicted that body size would explain more variation in pageviews if direct encounters are less important and eBird frequency would explain more variation if direct encounters are more important.

Previously, Mittermeier et al. (2019) identified temporal correspondence between eBird sightings and Wikipedia pageviews for a sample of migratory birds in Italy, Germany, Sweden, and the United States. Their study demonstrated that, for some migratory birds, the physical presence

of the species in a region correlated with increased public interest. This study builds on these findings by evaluating the relationship between the observation frequency and online interest across a much larger range of languages (25 language editions), geographic regions (25 regions including both temperate and tropical countries) and species (more than 10,000 bird species, nearly the entire global diversity). Thus, we provide a generalized understanding of the relationship between wild encounter frequency and public interest in birds irrespective of the phenomenon of migration and test the importance of the relationship across a range of linguistic and geographic contexts.

## 2 | METHODS

### 2.1 | Data selection and extraction

We used the number of pageviews that a page receives in Wikipedia as a measure of online interest (e.g., Mittermeier et al., 2019; Roll et al., 2016). Wikipedia does not include geographic information with its pageviews, but Wikipedia editions are constructed in different languages and summary data are available that list the proportion of pageviews each language edition receives by country (Zachte, 2020). Following previous studies, we used these country-level summaries as a coarse proxy for geography by pairing Wikipedia language editions with the geographic region that accounts for the majority of a language edition's pageviews (Generous, Fairchild, Deshpande, Del Valle, & Priedhorsky, 2014; Mittermeier et al., 2019).

We selected 25 Wikipedia editions for non-artificial languages that, as of June 22, 2019, had over 100,000 articles, a Wikipedia editing depth higher than 10 (a measure of the language edition's quality; Wikimedia 2019), and more than 50% of the language's pageviews originating from a single country (Appendix S1; Wikipedia, 2020; Zachte, 2020). Selecting language editions with more than 50% of their pageviews originating from a single country provided higher confidence in the geographic origin of the pageviews in our data set. It also resulted in several widely spoken languages with large Wikipedia editions not being included in our data set (e.g., Spanish, English).

We identified pages for bird species in Wikipedia using the Wikidata Query Service (<https://query.wikidata.org/>) to extract a list of entities tagged with an eBird taxon ID (Wikidata property: P3444) on June 23, 2019. Wikidata is a secondary database that collects structured data for Wikimedia projects, including all Wikipedia language editions (<https://www.wikidata.org>). We cross-referenced our list of Wikidata entities with the eBird/Clements global avian taxonomy (Clements et al., 2018) to ensure that non-bird pages were not

included. We obtained page links for Wikipedia pages in languages that met our criteria and downloaded pageviews for each page for the period between July 1, 2015 and June 22, 2019 (1,453 days) using "pageviews" in R (Keyes & Lewis, 2016).

To obtain eBird data, we downloaded the eBird Basic Data set (version April 2019) with records for all species and all years (eBird, 2019) for the regions associated with each of our 25 Wikipedia language editions. In most cases, we defined a region as a single country, however, in rare instances where the distribution of a language corresponded strongly to a specific subnational region, we downloaded data for that region rather than for the entire country (e.g., Tamil Wikipedia was paired with eBird records from Tamil Nadu rather than all of India). We limited our analyses to regions with a minimum of 10,000 unique sampling events in eBird, and to species that appeared in more than 10 sampling events in the region's eBird data set. These thresholds helped minimize biases present in smaller eBird data sets and reduce instances where incorrectly identified species had not yet been removed via eBird's review process (Wood et al., 2011). Bird pages in each Wikipedia language edition were assigned as either "local" if they occurred on the eBird list for the associated region or "global" if they did not (eBird, 2019). For local species, we calculated the observation frequency of a species as the total reports of that species in the region divided by the total unique sampling events in the region (Sullivan et al., 2009). This approach provided a rough overall measure of observation frequency for all local species in region. It did not account for seasonal variations (such as for migratory species) or for how species were distributed within the region. Since global species by definition had no sampling events in a region, they did not have an eBird frequency for that region in our data set. We obtained bird body mass data from Dunning (2008).

### 2.2 | Data analysis

We explored the relationship between Wikipedia pageviews, eBird sightings, and avian body mass across each of the 25 language-region pairs in our data set. Since language editions varied substantially in their views, we assessed patterns in each language separately. Furthermore, since our data were sparse (we only had eBird frequency data for local species, and we did not have body mass data for all species) we analyzed data for each language in four separate analyses. First, for all of the bird Wikipedia pages in a language, we looked at the overall pattern between pageviews for global as opposed to local species and compared the median pageviews for local

and global species using a Wilcoxon rank-sum test. Second, for all the birds with Wikipedia pages in a language for which we had eBird sighting frequency (i.e., the local species that occurred in the region associated with the language), we tested the relationship between eBird sighting frequency as a predictor and Wikipedia pageviews as a response variable using a simple linear regression model. Third, for all the birds with Wikipedia pages in a language for which we had body mass data (this included both local and global species), we modeled the relationship between body mass (predictor) and Wikipedia pageviews (response) using a simple linear regression model. Since we had body mass data for both global and local species in each language, we tested whether a species being classified as local or global influenced the relationship between body mass and pageviews using an analysis of covariance (ANCOVA). Fourth, for the birds with Wikipedia pages in a language for which we had both eBird frequency data and body mass data (i.e., local species with body mass data) we used a combined linear regression model with eBird frequency and body mass as predictors of Wikipedia pageviews. In the combined model, we compared the relative influence of eBird frequency and body mass as predictors of Wikipedia pageviews using variation partitioning (Oksanen et al., 2017). For all analyses, we normalized variables using a logit transformation for proportions (eBird frequency) and a log transformation for counts and continuous variables (Wikipedia pageviews, body mass). Final models were tested for assumptions of normality and homoscedasticity of residuals.

### 3 | RESULTS

#### 3.1 | Data selection and extraction

Our data set contained 78,415 Wikipedia pages for birds across 25 languages (Appendix S2). In total, 10,174 bird species had at least one Wikipedia page in our data set (96.1% of the total bird species; Clements et al., 2018). Numbers of bird pages per language varied from 103 (Hindi) to 10,103 in Dutch (mean pages per language 3,137, *SD* 3,034). Wikipedia pages in our data set received 367 million views over the sampling period (views per language 417,000–70.9 million; mean 14.7 million, *SD* 18.5 million). With the exception of Brazil (which accounted for 79.7% of the pageviews for Portuguese Wikipedia), all of our geographic regions were in Europe (15 regions) and Asia (nine regions).

Our eBird data set contained 2.3 million unique sampling events with records of 4,340 bird species. Unique sampling events per region ranged from 10,600 (South

Korea) to 817,000 (India; mean 91,900, *SD* 169,000). eBird species richness per region ranged from 313 species (Czech Republic) to 1,716 (Brazil; mean overall 615, *SD* 374). Not all species that occurred in a region had a Wikipedia page in the associated language. The proportion of eBird species in a region with Wikipedia pages varied from 7.62% of species (Hindi—India) to >98% of species in eleven language–region pairs (mean overall 82.5%, *SD* 25.6%). We obtained body mass data for 9,406 species.

#### 3.2 | Wikipedia pageviews for local vs. global species

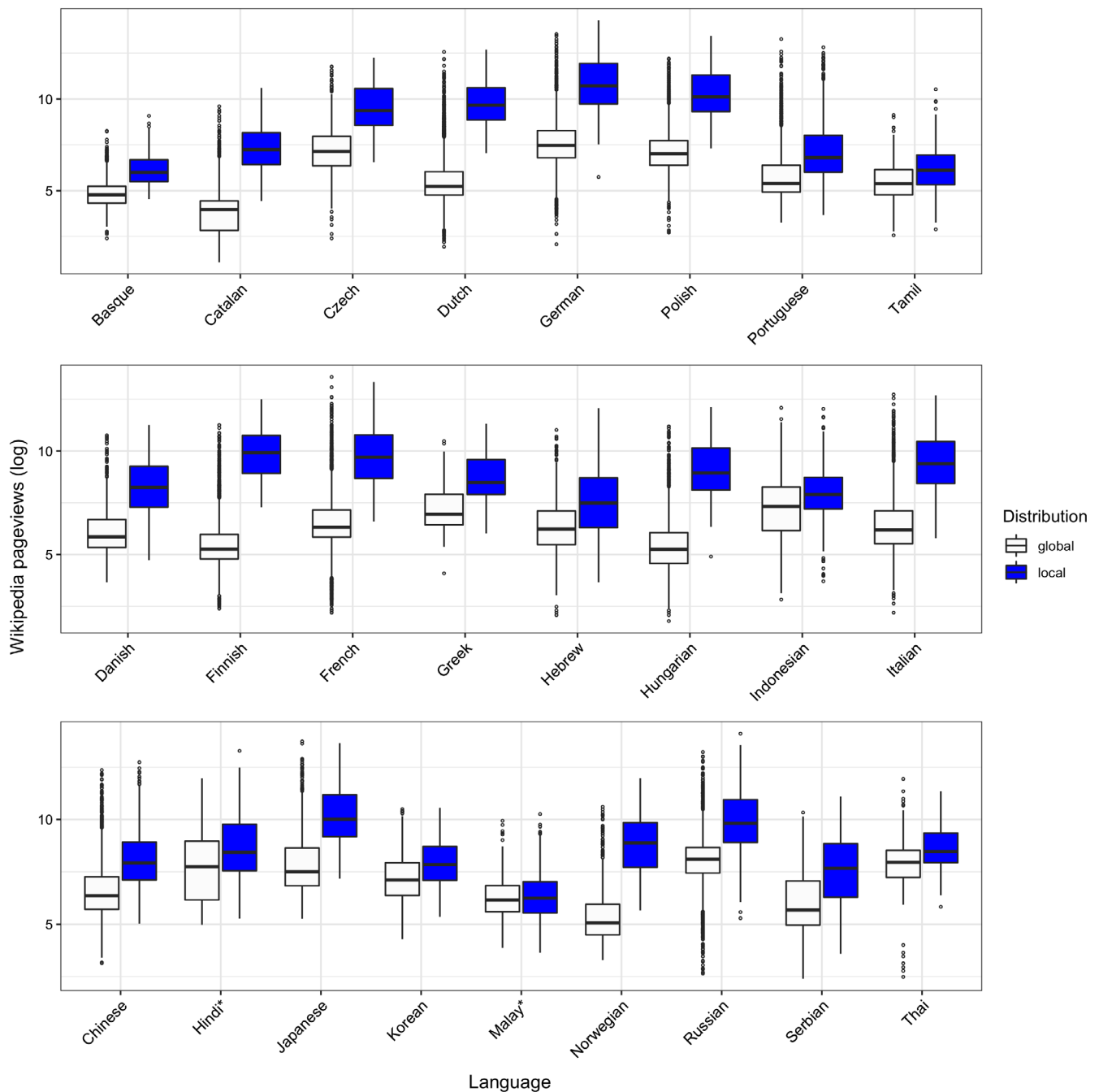
Our data set contained many more Wikipedia pages that we classified as global (70,250 global vs. 8,165 local). Despite this, pages for local species received more views than those for global species (218 million vs. 149 million pageviews; mean pageviews for local species 26,700, *SD* 74,900; mean pageviews for global species 2,120, *SD* 14,300). The mean pageviews for local species was higher than the mean pageviews for global species in all 25 languages, with the difference between the means being statistically significant in all but two languages (Wilcoxon rank-sum  $p < .01$ ; mean pageviews local 723–131,000; mean pageviews global 106–16,600; Figure 1, Appendix S3). The effect size of the difference in mean pageviews for local as opposed to global species was either large or moderate in 20 languages (large defined as >0.5, moderate as 0.3–0.5).

#### 3.3 | eBird frequency as a predictor of Wikipedia pageviews

For local birds in our data set (i.e., those for which we had eBird frequency data) we tested the relationship between eBird frequency as a predictor and Wikipedia pageviews as a response variable using a linear regression model. The number of species on a region's eBird list with Wikipedia pages ranged from 82 to 1,138 (mean 399, *SD* 216; total local pages across all languages 8,165). eBird frequency showed a significant positive correlation with Wikipedia pageviews in all 25 languages ( $p < .01$ ). Across languages, a linear regression model with eBird frequency as a predictor explained between 6.9 and 49.0% of the variance in Wikipedia pageviews for bird pages (adj.  $R^2 = .07$ –.49, mean 0.25, *SD* .13; Figure 2, Table 1).

The most frequently reported birds in eBird in each language received a high proportion of the Wikipedia pageviews for birds. Among local species, the 20 most frequently recorded species in eBird accounted for 20.2 to

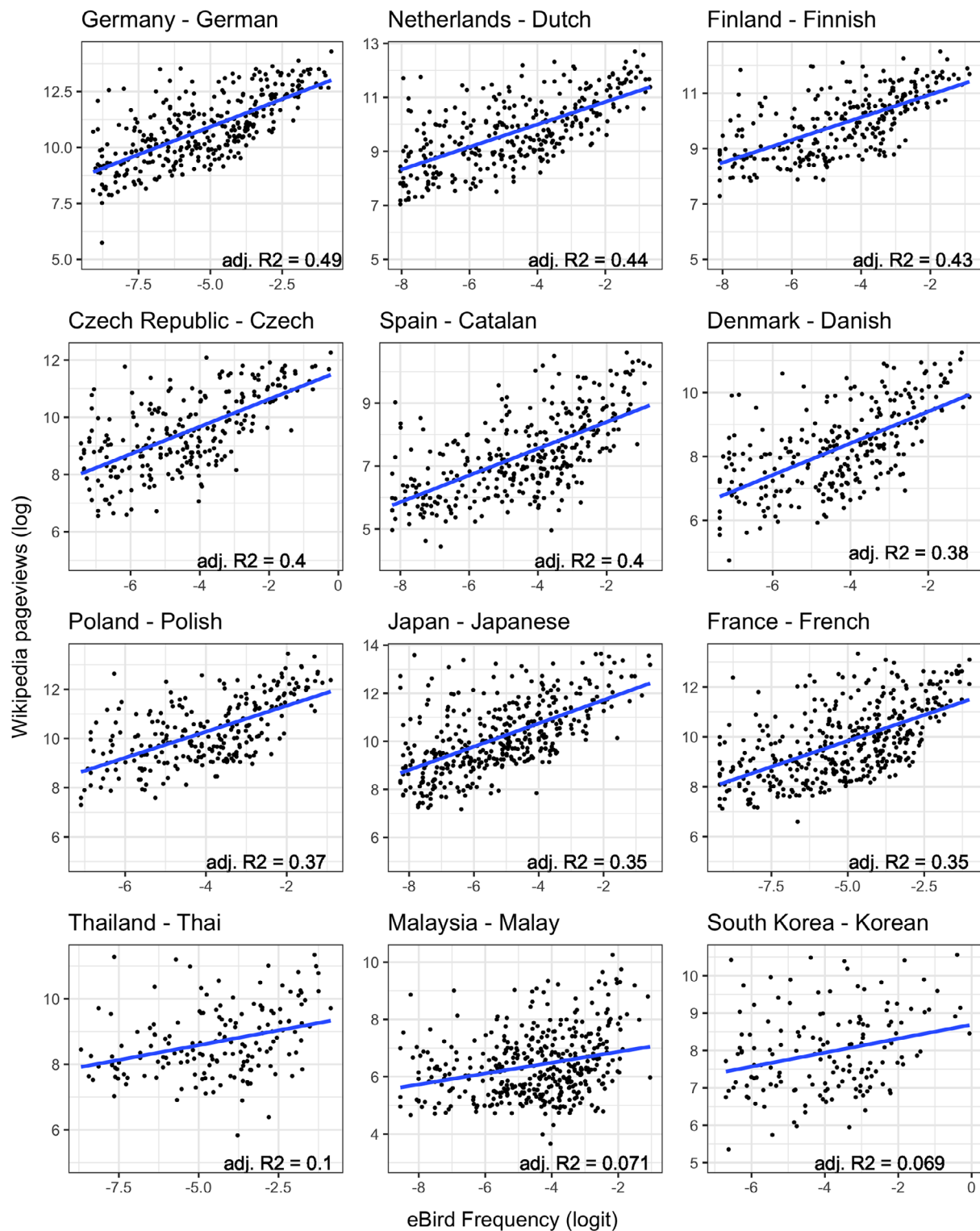




**FIGURE 1** Bird species that occur in the region responsible for the majority of a Wikipedia language edition's pageviews ("local" species) attract more pageviews than bird species that do not occur in the region ("global" species). Points indicate outliers occurring > 1.5x the interquartile range beyond the median. Statistical significance of the difference between means calculated using a Wilcoxon rank-sum test with  $p < 0.01$ ; two language editions where difference between means was not statistically significant are marked with an asterisk (Malay and Hindi)

49.8% of the pageviews for local bird species, while making up 3.14 to 20.1% of the local bird species pages (mean proportion of views 40.0%,  $SD$  8.24%; mean proportion of species 13.1%,  $SD$  5.09%). This pattern was also evident when considering the total bird Wikipedia pages in a language (i.e., both local and global). Across languages, the

20 species with the highest eBird frequency accounted for 3.16 to 19.6% of the total Wikipedia pageviews for bird species, while consisting of only 0.20 to 4.20% of the total bird species pages (mean proportion of views 13.36%,  $SD$  4.66%; mean proportion of species 1.29%,  $SD$  1.14%; Appendix S4).



**FIGURE 2** eBird observation frequency correlates positively with Wikipedia pageviews for local bird species across multiple Wikipedia language editions. In a linear model with eBird observation frequency (logit transformed) as a predictor of pageviews (log transformed), observation frequency predicted between 49% and 6.9% of the variance in pageviews (top three rows show 9 language editions with the highest adj. R2 in the linear model; bottom row shows 3 languages with the lowest adj. R2 in the linear model; see Table 1 for full results)

**TABLE 1** Model results for a linear model with eBird observation frequency (logit transformed) as a predictor of pageviews (log transformed) to pages for local bird species in 25 Wikipedia language editions

Language	Estimate ( $\beta$ )	SE	p-value	Model adj. $R^2$
German	0.50	0.03	7.16E-51	.49
French	0.42	0.03	9.58E-40	.35
Dutch	0.42	0.03	1.61E-40	.44
Russian	0.50	0.05	3.57E-24	.20
Italian	0.40	0.04	5.06E-21	.25
Polish	0.53	0.04	6.82E-28	.37
Japanese	0.49	0.03	2.02E-39	.35
Chinese	0.23	0.02	1.83E-19	.16
Portuguese	0.38	0.02	1.71E-53	.19
Serbian	0.29	0.06	1.62E-06	.14
Catalan	0.42	0.03	9.62E-38	.40
Norwegian	0.43	0.04	5.53E-28	.34
Indonesian	0.35	0.04	5.34E-18	.14
Finnish	0.41	0.03	7.01E-35	.43
Korean	0.19	0.05	7.67E-04	.07
Hungarian	0.45	0.04	1.89E-20	.28
Czech	0.48	0.04	1.51E-28	.40
Basque	0.24	0.02	8.19E-20	.25
Malay	0.19	0.03	2.68E-08	.07
Danish	0.50	0.04	2.55E-29	.38
Hebrew	0.35	0.05	1.26E-12	.16
Greek	0.26	0.06	8.94E-06	.11
Thai	0.18	0.04	2.57E-05	.10
Hindi	0.23	0.06	5.80E-04	.13
Tamil	0.24	0.03	8.68E-12	.16

### 3.4 | Body mass as a predictor of Wikipedia pageviews

For species in our data set for which we had body mass data, we tested the relationship between body mass (as a predictor) and Wikipedia pageviews (as a response) using a linear regression model. The number of species per language for which we had body mass data varied from 99 to 9,403 (mean 2,993, *SD* 2,875; 74,830 pages total). Body mass showed a significant positive correlation with Wikipedia pageviews in 21 of 25 languages ( $p < .01$ ; Table 2). In languages where the relationship was significant, body mass described between 2.1 and 24.5% of the variance in pageviews ( $p < .01$ , adj.  $R^2 = .02$ –.25, mean 0.11, *SD* 0.08). The interaction between local–global and body mass was significant in 12 of the 25 languages ( $p < .01$ ). In instances where the relationship

**TABLE 2** Model results for a linear model with body mass (log transformed) as a predictor of pageviews (log transformed) to pages for local and global bird species in 25 Wikipedia language editions

Language	Estimate ( $\beta$ )	SE	p-value	Model adj. $R^2$
German	0.36	0.01	5.1E-167	.17
French	0.39	0.01	.00	.20
Dutch	0.42	0.01	.00	.25
Russian	0.28	0.01	5.2E-89	.12
Italian	0.32	0.01	2.8E-104	.12
Polish	0.26	0.02	1.4E-63	.10
Japanese	0.18	0.02	4.2E-16	.04
Chinese	0.36	0.01	9.2E-130	.21
Portuguese	0.43	0.01	1.9E-289	.23
Serbian	1.0E-03	0.05	0.99	.00
Catalan	0.13	0.01	3.80E-32	.03
Norwegian	0.34	0.02	4.00E-62	.13
Indonesian	−0.05	0.03	0.06	.00
Finnish	0.39	0.01	1.2E-291	.20
Korean	0.13	0.03	2.2E-05	.04
Hungarian	0.41	0.01	1.5E-296	.21
Czech	0.17	0.03	9.6E-12	.04
Basque	0.07	0.01	4.5E-53	.03
Malay	0.09	0.03	2.3E-04	.02
Danish	0.13	0.03	8.9E-08	.03
Hebrew	0.12	0.02	1.1E-07	.02
Greek	0.00	0.04	0.94	.00
Thai	0.11	0.04	0.01	.02
Hindi	0.03	0.10	0.77	−.01
Tamil	0.18	0.03	4.5E-09	.07

was significant, body mass had a stronger positive relationship with pageviews for global species than for local ones in all but one language (local mean coefficient 0.18, *SD* 0.11, global mean coefficient 0.34, *SD* 0.05; Appendix S5).

### 3.5 | eBird frequency versus body mass as predictors of Wikipedia pageviews for local species

For all species for which we had both body mass and eBird frequency data (i.e., local species with body mass data; 2,563 species with 8,014 pages), we fitted a linear regression model that included both eBird frequency and body mass as predictors of Wikipedia pageviews. This joint model was significant in all 25 languages ( $p < .01$ ) and explained between 10.5 and 54.9% of the variance in



Language	Model adj. $R^2$	Frequency   mass	Mass   frequency	Difference
German	.55	0.52	0.06	0.46
Dutch	.46	0.44	0.02	0.42
Czech	.45	0.45	0.05	0.39
Catalan	.43	0.42	0.04	0.39
Polish	.41	0.4	0.05	0.35
Finnish	.51	0.42	0.08	0.34
Japanese	.38	0.36	0.03	0.33
Danish	.47	0.39	0.1	0.29
French	.47	0.39	0.12	0.28
Norwegian	.39	0.33	0.05	0.28
Hungarian	.32	0.29	0.05	0.24
Basque	.29	0.27	0.05	0.22
Italian	.36	0.3	0.11	0.19
Hebrew	.22	0.2	0.06	0.13
Hindi	.12	0.12	−0.01	0.13
Indonesian	.13	0.13	0.0	0.13
Russian	.32	0.23	0.12	0.11
Greek	.18	0.15	0.08	0.07
Serbian	.24	0.17	0.11	0.06
Thai	.25	0.2	0.14	0.06
Chinese	.33	0.21	0.17	0.04
Tamil	.33	0.22	0.17	0.04
Malay	.13	0.09	0.06	0.04
Korean	.11	0.07	0.04	0.03
Portuguese	.44	0.18	0.26	−0.08

Note: Adjusted  $R^2$  of Frequency | Mass and Mass | Frequency calculated using variation partitioning (Oksanen et al., 2017).

**TABLE 3** Model adjusted  $R^2$  results for a linear model with both body mass and eBird frequency as predictors of pageviews to pages for local bird species in 25 Wikipedia language editions

Wikipedia pageviews for birds across languages (adj.  $R^2$  = .10–.55, mean 0.33,  $SD$  0.13). We used partition of variance using partial linear regression to assess the relative contributions of body mass and eBird frequency as explanatory variables. For all languages except one (Portuguese), eBird frequency explained more of the variance in pageviews than body mass (adj.  $R^2$  frequency | mass 0.07–0.52, mean 0.28,  $SD$  0.13; adj.  $R^2$  mass | frequency −0.01–0.26, mean 0.08,  $SD$  0.06; Table 3). Often the difference between the two was substantial; in 14 languages eBird frequency explained over three times more variance than body mass.

## 4 | DISCUSSION

We outlined three hypotheses for assessing the relative importance of a physical trait (body size) versus the potential to encounter a species in the wild in

determining public interest in birds. For all three hypotheses, our results strongly supported the importance of encounter frequency over body size in determining public interest online. First, we observed a clear geographical pattern in the distribution of Wikipedia pageviews for birds, with local species attracting more interest than pages for global species across languages (Figure 1). Second, while both body size and eBird frequency (Figure 2; Tables 1 and 2) correlated positively with increased pageviews; eBird frequency explained more of the variance in pageviews across languages (mean adj.  $R^2$  eBird frequency 0.25,  $SD$  0.13; mean adj.  $R^2$  body size 0.11,  $SD$  0.08). Third, when eBird frequency and body size were compared directly in a combined model, the former explained significantly more variance in pageviews in 24 of 25 languages in our data set (mean adj.  $R^2$  frequency | mass 0.28,  $SD$  0.13; mean adj.  $R^2$  mass | frequency mean 0.08,  $SD$  0.06; Table 3). In several language editions, the difference between eBird frequency and

body size as predictors of Wikipedia pageviews in this combined model was substantial: in German Wikipedia eBird frequency explained 52% of the variance in Wikipedia pageviews for local species when used as the sole predictor in the combined linear model, while mass explained only 6% of the variance.

Fitting the established wisdom in conservation (e.g., Berti et al., 2020) our results found that body size correlated positively with online interest. Notably, some of the most-viewed pages in several languages were rarely encountered, global species. In a striking example, the Dodo (*Raphus cucullatus*), a relatively large bodied, extinct species, received the most pageviews of any bird page in both French and Portuguese Wikipedias. Thus, indirect representations such as appearances in media and popular culture clearly drive interest in some species and can even determine the most viewed species pages overall. However, when assessed across larger numbers of species, our results clearly showed that body size was secondary to the potential for direct wild encounters in predicting Wikipedia pageviews for birds. While high public interest in locally common species has been shown in previous studies using online data (e.g., Correia et al., 2016; Zmihorski et al., 2013), we demonstrated that this pattern is present across a wide range of cultural and geographic contexts.

The positive correlation between wild encounter rates and online interest could prove relevant for conservation policy in two ways. First, from a methodological standpoint it demonstrates the importance of taking encounter frequency into account when measuring public interest in species. Second, it highlights the importance of local and frequently encountered species in attracting people's attention to the natural world. Few if any of the 20 most frequently observed birds in the regions in our data set would qualify as "charismatic megafauna" under most criteria, and yet they accounted for between 10 and 20% of the total bird pageviews in most languages. As a result of the interest they attract, these commonly observed species could act as entry points for people's interactions with biodiversity. Conservation initiatives that improve opportunities for people to interact with common, local species (such as the development of public parks and green spaces) could help increase public support for conservation.

There are several caveats to consider when interpreting our results. Our method of matching Wikipedia languages to regions provided only rough geographic resolution. For each language, a percentage of pageviews came from outside of the associated region. Furthermore, species that are frequently encountered in one part of a region may be absent from another, particularly in large regions such as Brazil. Using culturomic resources with

more fine-scale geographic resolution could be an approach to investigating these patterns at finer scales (e.g., Correia et al., 2016). eBird data also have important biases (Sullivan et al., 2009). eBird users may be more likely to visit certain locations, report some species over others, and the coverage and types of users may differ between regions. It is also important to consider the ways in which people use Wikipedia as opposed to other online resources (Correia et al., 2021). Wikipedia pageviews reflect people's attention and their searches for additional knowledge, but not necessarily their preferences. High interest in a species could equally result from people liking a species as it could from people considering it is a pest and wanting to remove it. Thus, additional context is required understand the drivers behind high pageviews for species. Finally, it is possible that interactions with wild animals are more important for birds than for other taxonomic groups. Birds are more easily observed and identified in the wild than many other organisms and generate unique forms of human-nature interaction through activities such as birdwatching and bird-feeding (Cocker, Tipling, Elphick, & Fanshawe, 2013). It is possible that observation frequency may have a stronger positive correlation with online interest for birds than it does for other groups of organisms. Future studies may be able to address these caveats and expand upon our results to investigate the mechanisms underlying the high online interest in local bird species.

Our results highlighted the novel insights that are possible using the large analytical scales enabled by conservation culturomics methods. The relationship between sighting frequency and increased public interest may be obscured when public interest in species is assessed using small or unrepresentative samples. In smaller samples, traits of extreme outliers (such as the Dodo) may outweigh broadscale patterns in the data. Thus, our findings highlighted the value of large-scale analyses across many species and over broad cultural and geographic contexts. Access to large digital sources such as Wikipedia, together with powerful new analytical tools hold much promise for further novel research insights using conservation culturomics (Correia et al., 2021).

Future studies can build on our results by using more focused studies to investigate the causal relationships between public interest and people's wild encounters with species. If there is a causal relationship between wild encounters with species and increased interest, as the correlations we identified suggest, this relationship may hint at an upcoming challenge for conservation: populations of many species are declining at alarming rates resulting in fewer opportunities for people to encounter species in the wild (Dirzo et al., 2014; Inger

et al., 2015; Rosenberg et al., 2019). These trends are exacerbated by urbanization and the “extinction of experience” that results from increasingly limited interactions with wild nature (Miller, 2005; Soga & Gaston, 2016). These trends could result in an “extinction of interest” in some species as opportunities for direct interactions with them decline. Reinforcing the value of common species in conservation could help to combat this (Gaston, 2010). Creating chances for people to engage with local and common species and through activities such as bird feeding are often feasible and inexpensive (Cox & Gaston, 2016; Miller, 2005). By promoting these wild interactions, the relationship between wild encounters and public interest could offer opportunities to increase people's awareness of nature and engage broader support for conservation initiatives.

## ACKNOWLEDGMENTS

We are grateful to Wikipedia and eBird for allowing open access to their databases and to the many volunteers whose contributions to Wikipedia and eBird made this study possible.

## CONFLICT OF INTEREST

There are no competing interests.

## AUTHOR CONTRIBUTIONS

John C. Mittermeier, Uri Roll, and Richard Grenyer designed the study and research questions. John C. Mittermeier and Thomas J. Matthews designed the methods and statistical analyses. John C. Mittermeier compiled and analyzed the data and prepared the figures. John C. Mittermeier, Uri Roll, Thomas J. Matthews, and Ricardo Correia wrote the manuscript.

## DATA AVAILABILITY STATEMENT

All data used in this study is open access.

## ETHICS STATEMENT

No ethics approval was required for this research.

## ORCID

John C. Mittermeier  <https://orcid.org/0000-0001-5061-7699>

Uri Roll  <https://orcid.org/0000-0002-5418-1164>

Ricardo Correia  <https://orcid.org/0000-0001-7359-9091>

## REFERENCES

- Berti, E., Monsarrat, S., Munk, M., Jarvie, S., & Svenning, J. C. (2020). Body size is a good proxy for vertebrate charisma. *Biological Conservation*, 251, 108790.
- Clements JF, Schulenberg TS, Iliff MJ, Roberson D, Fredericks TA, Sullivan BL, Wood CL. 2018. The eBird/Clements checklist of birds of the world: v2018. Available from <http://www.birds.cornell.edu/clementschecklist/download/>.
- Clucas, B., McHugh, K., & Caro, T. (2008). Flagship species on covers of US conservation and nature magazines. *Biodiversity and Conservation*, 17, 1517–1528.
- Cocker, M., Tipling, D., Elphick, J., & Fanshawe, J. (2013). *Birds and people*. London, UK: Jonathan Cape.
- Correia, R. A., Jepson, P. R., Malhado, A. C. M., & Ladle, R. J. (2016). Familiarity breeds content: Assessing bird species popularity with culturomics. *PeerJ*, 4, 1–15.
- Correia, R. A., Ladle, R., Jaric, I., Malhado, A. C. M., Mittermeier, J., Roll, U., ... Di Minin, E. (2021). Digital data sources and methods for conservation culturomics. *Conservation Biology*. <https://doi.org/10.1111/cobi.13706>
- Cox, D. T. C., & Gaston, K. J. (2016). Urban bird feeding: Connecting people with nature. *PLoS One*, 11, 1–13.
- Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, 345, 401–406.
- Dunning, J. B. (2008). *CRC handbook of avian body masses* (2nd ed.). Boca Raton, FL: CRC Press.
- eBird. 2019. eBird Basic Dataset. Version: EBD\_relApril-2019. Cornell Lab of Ornithology, Ithaca, NY.
- eBird. 2020. Publications. Available from <https://ebird.org/science/publications>.
- Gaston, K. J. (2010). Valuing common species. *Science*, 327, 154–155.
- Generous, N., Fairchild, G., Deshpande, A., Del Valle, S. Y., & Friedhorsky, R. (2014). Global disease monitoring and forecasting with Wikipedia. *PLoS Computational Biology*, 10, e1003892.
- Gunnthorsdottir, A. (2001). Physical attractiveness of an animal species as a decision factor for its preservation. *Anthrozoös*, 14, 204–214.
- Inger, R., Gregory, R., Duffy, J. P., Stott, I., Voříšek, P., & Gaston, K. J. (2015). Common European birds are declining rapidly while less abundant species' numbers are rising. *Ecology Letters*, 18, 28–36.
- Kellert, S. R. (1982). Factors in endangered social and perceptual species management. *The Journal of Wildlife Management*, 49, 528–536.
- Keyes O, Lewis J. 2016. pageviews: an API client for wikimedia traffic data. Available from <https://cran.r-project.org/package=pageviews>.
- Ladle, R. J., Correia, R. A., Do, Y., Joo, G. J., Malhado, A. C. M., Proulx, R., ... Jepson, P. (2016). Conservation culturomics. *Frontiers in Ecology and the Environment*, 14, 269–275.
- Levatic T, Ligocki S. 2020. EOD - eBird Observation Dataset. Cornell Lab of Ornithology. Occurrence dataset. Available from [www.gbif.org](http://www.gbif.org).
- Lišková, S., & Frynta, D. (2013). What determines bird beauty in human eyes? *Anthrozoös*, 26, 27–41.
- Lorimer, J. (2007). Nonhuman charisma. *Environment and Planning D: Society and Space*, 25, 911–932.
- Macdonald, E. A., Burnham, D., Hinks, A. E., Dickman, A. J., Malhi, Y., & Macdonald, D. W. (2015). Conservation inequality and the charismatic cat: *Felis felis*. *Global Ecology and Conservation*, 3, 851–866.
- Martin, T. E., Lurbiecki, H., Joy, J. B., & Mooers, A. O. (2014). Mammal and bird species held in zoos are less endemic and less threatened than their close relatives not held in zoos. *Animal Conservation*, 17, 89–96.

- Miller, J. R. (2005). Biodiversity conservation and the extinction of experience. *Trends in Ecology & Evolution*, 20, 430–434.
- Mittermeier, J. C., Roll, U., Matthews, T. J., & Grenyer, R. (2019). A season for all things: Phenological imprints in Wikipedia usage and their relevance to conservation. *PLoS Biology*, 17, e3000146.
- Oksanen J. F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs, Helene Wagner 2017. *vegan: Community Ecology Package*. Available from <https://cran.r-project.org/package=vegan>.
- Roll, U., Mittermeier, J. C., Diaz, G. I., Novosolov, M., Feldman, A., Itescu, Y., ... Grenyer, R. (2016). Using Wikipedia page views to explore the cultural importance of global reptiles. *Biological Conservation*, 204, 42–50.
- Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., ... Marra, P. P. (2019). Decline of the north American avifauna. *Science*, 1313, 1–10.
- Smith, R. J., Veríssimo, D., Isaac, N. J. B., & Jones, K. E. (2012). Identifying Cinderella species: Uncovering mammals with conservation flagship appeal. *Conservation Letters*, 5, 205–212.
- Soga, M., & Gaston, K. J. (2016). Extinction of experience: The loss of human-nature interactions. *Frontiers in Ecology and the Environment*, 14, 94–101.
- Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142, 2282–2292.
- Verissimo, D., Pongilluppi, T., Cintia, M., Santos, M., Develey, P. F., Fraser, I., ... MacMilan, D. C. (2013). Using a systematic approach to select flagship species for bird conservation. *Conservation Biology*, 28, 269–277.
- Wikipedia. 2020. List of Wikipedias. Available from [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias).
- Wood, C., Sullivan, B., Iliff, M., Fink, D., & Kelling, S. (2011). eBird: Engaging birders in science and conservation. *PLoS Biology*, 9, e1001220.
- Zachte E. 2020. Wikimedia Traffic Analysis Report - Page Views per Wikipedia Language - Breakdown. Available from <https://stats.wikimedia.org/wikimedia/squids/SquidReportPageViewsPerLanguageBreakdown.htm>.
- Zmihorski, M., Dziarska-Palac, J., Sparks, T. H., & Tryjanowski, P. (2013). Ecological correlates of the popularity of birds and butterflies in internet information resources. *Oikos*, 122, 183–190.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Mittermeier JC, Roll U, Matthews TJ, Correia R, Grenyer R. Birds that are more commonly encountered in the wild attract higher public interest online. *Conservation Science and Practice*. 2021;e340. <https://doi.org/10.1111/csp2.340>