Check for updates

Diversity and Distributions

WILEY

Diversity and Distributions

RESEARCH ARTICLE OPEN ACCESS

Detections of Rare Species Lead Citizen Scientists to Initiate Data Recording

Louis J. Backstrom[1,2] | Rachel L. Drake[1,2] | Hannah Worthington[1,2] | Alison Johnston[1,2]

[1]Centre for Research Into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK | [2]School of Mathematics and Statistics, University of St Andrews, St Andrews, UK

Correspondence: Louis J. Backstrom (ljb38@st-andrews.ac.uk)

Received: 12 March 2025 | **Revised:** 24 September 2025 | **Accepted:** 30 September 2025

Editor: Susan Jarvis

Funding: Louis Backstrom and Rachel Drake received PhD scholarship funding from CREEM, the Centre for Research into Ecological and Environmental Modelling.

Keywords: BirdTrack | citizen science | community science | eBird | GAMLSS | observation process | observer preferences | occupancy model | sampling bias | taxonomic bias

ABSTRACT

Aim: Citizen science data are increasingly used to monitor biodiversity but come with several challenges that can impair accurate ecological conclusions. We explore how observers' preferences for certain species over others bias the initiation of survey efforts and assess the extent of this sampling bias in semi-structured citizen science datasets. We investigate the effects of this bias on occupancy model-based species distribution models and offer suggestions for mitigating this when analysing citizen science data.

Location: Great Britain, with methods applicable to citizen science datasets worldwide.

Methods: We assess observer species preferences in list initiation via two methods: (1) indirectly through exploration of the relationship between species rarity and survey duration and (2) directly through analysis of the first-recorded species on surveys. We use these results to assess the impact of list-initiation sampling bias on occupancy models of 132 common breeding birds across Britain.

Results: We find evidence for list-initiation sampling bias in British eBird and BirdTrack data. This bias is driven by observer preferences for certain species over others, with species preferences correlated with species rarity. This bias was stronger in short-duration surveys, and removing short surveys to remove these biased lists had limited impacts on occupancy models.

Conclusions: Citizen science schemes projects that allow observers freedom in how they observe and record biodiversity consequently have more heterogeneous datasets. Observer species preferences lead to biased initiation of surveys and consequent overreporting of some species. Although we find limited effects of this bias on occupancy models in our study, we nevertheless suggest analysts consider its possible effects when exploring and analysing citizen science data.

1 | Introduction

Species observations recorded by volunteer naturalists are an important data source in biodiversity monitoring, with billions of records contributed by citizen scientists (also called community scientists) across broad spatial, temporal and taxonomic extents. Despite this wealth of information, citizen science data have well-recognised caveats that reduce their utility in biological research, management and policy (Burgess et al. 2017). It is therefore crucial to improve our understanding

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.

© 2025 The Author(s). *Diversity and Distributions* published by John Wiley & Sons Ltd.

of how these challenges impact ecological inference and the effectiveness of analytical approaches to mitigate their impact (Dickinson et al. 2010; Johnston et al. 2023; Stuber et al. 2022). One major challenge is the effect of observer behaviours on the data collection process, as variation in how citizen scientists observe and record wildlife can confound the underlying biological processes that are of interest to analysts, leading to poor inference and potentially worsened outcomes for biodiversity.

As individuals, citizen scientists have preferences that shape their behaviour and influence the data they contribute. Many preferences are taxonomic in nature, with observers favouring certain species over others based on attributes such as rarity (Arazy et al. 2024; Bowler et al. 2022; Johnston et al. 2023; Randler et al. 2023) and novelty (Bowler et al. 2022; Goldstein and Stoudt 2025), with morphological traits such as body size and colour also influencing the observation process (Callaghan et al. 2021; Randler et al. 2023; Stoudt et al. 2022; van Tongeren et al. 2023). We expect to see preferred species appearing more in these datasets, leading to bias in citizen science data.

A key feature used in citizen science projects to understand taxonomic bias is information on whether or not a list of species from a survey represents a 'complete checklist.' A complete checklist includes all the species that the observer could detect and identify during their survey (Kelling et al. 2019). Data from complete checklists are less taxonomically biased as the influence of observer species preferences is reduced and common species are less likely to be omitted (Johnston et al. 2021; Snäll et al. 2011). However, taxonomic bias may still influence complete checklists, as some complete lists may be started by the detection of an interesting species. Species that prompt an observer to initiate a list will therefore lead to biased sampling of biodiversity even within this complete checklist framework, due to an overrepresentation of detections of interesting species. Additionally, the interest of an observation to trigger a checklist will vary not only by species, but also spatially and temporally for each species. For example, the first observation of a species each year may have particular interest.

In this article, we aim to (1) demonstrate conceptually how observer preferences can lead to biased list initiation, (2) assess the extent of this list-initiation bias in two semi-structured citizen science datasets across Britain, (3) measure the effects of this bias on a common application of biodiversity citizen science data, occupancy model-based species distribution models and (4) present recommendations for analysts and citizen science project coordinators wishing to mitigate the effects of observer preferences and their related biases.

2 | Materials and Methods

2.1 | Datasets and Study Area

We explored the effects of list-initiation bias in two citizen science projects across Great Britain in 2021. The two datasets we investigated were eBird (Cornell Lab of Ornithology 2024) and BirdTrack (British Trust for Ornithology 2024), which are both semi-structured citizen science projects that collect observations of birds via 'checklists' (lists of species observed during one visit to a site) submitted by birders. While the two datasets are similar in structure, we analysed them separately for two reasons: (1) eBird records checklist durations to 1-min precision, whereas BirdTrack rounds durations to 5-min intervals and (2) eBird does not record the order in which species are added to checklists (list entry order), whereas a subset of the BirdTrack dataset does. Given that most citizen science projects do not record list entry order, but many do record precise checklist durations, we expect eBird to be more representative of semi-structured citizen science datasets than BirdTrack, and thus we focus on eBird in this paper and use BirdTrack primarily to validate our results. We selected eBird and BirdTrack data from 2021 as this corresponded with the period when the BirdTrack mobile app tracked list entry order (version 3, first released early 2019, superseded by version 4 in late 2022). Full details of the eBird and BirdTrack datasets, and the filters we applied to them, as well as details of the environmental datasets we used in species distribution modelling, are provided in Appendix S1. The filtered eBird dataset contained 3,600,798 observations of 388 species across 171,474 checklists, and the BirdTrack dataset contained 1,316,710 observations of 394 species across 63,186 checklists.

2.2 | Modelling Species Rarity and Checklist Duration With eBird

We explored the relationship between species rarity and checklist duration in the eBird dataset, as an indication of rarity-driven list-initiation bias. We expected rare species to be overrepresented on very short-duration checklists, due to observers' tendencies to initiate a short, complete survey upon encountering a species of interest. To explore this relationship, we first calculated a rarity score for each eBird checklist (hereafter 'checklist rarity score') and used this as the response variable in our models. We defined the checklist rarity score as the overall reporting rate (prevalence) of the rarest species on each checklist. However, checklists with different numbers of species reported have different ranges of possible rarity scores. All checklists have the same lower bound, as any list can record the rarest species in the dataset. However, checklists of different lengths have different upper bounds. If we consider the upper bound to be all species on a list being the most common species possible, then a checklist with just two species could at most have the 'rarest' species be the second most common species in the dataset, whereas a checklist with 10 species could at most have the tenth most common species as the rarest species. Therefore, for each checklist, we calculated the checklist rarity score as the prevalence of the rarest species on the checklist divided by the maximum possible checklist rarity score for a checklist of that number of species and then subtracted this score from 1. All checklist rarity scores were therefore in the interval [0, 1]. Checklist rarity scores close to 0 indicate a checklist with only common species, and scores close to 1 indicate a checklist with at least one very rare species. This process is illustrated in Figure S1 in Appendix S2.

We wanted to explore how checklist rarity varied with duration, and we hypothesised that very short-duration checklists would have greater rarity (scores closer to 1) than expected by chance, due to the detection of rare species driving the initiation of short-duration checklists. We modelled this relationship using a beta distribution and Generalised Additive Models (GAMs) for

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Location, Scale, and Shape (GAMLSS), which we chose due to their flexibility, particularly their ability to fit non-linear relationships between the response variable and predictors, while also modelling the variance as a function of the predictors (Rigby and Stasinopoulos 2005).

A key consideration worth noting is that we expect longer duration checklists to record more species, and therefore more rare species by chance, so it was important to consider this relationship in a null case with no species preferences driving list initiation. To understand how our fitted relationship between checklist rarity score and duration deviated from the relationship we would expect to see if there were no effect of observer species preferences on list initiation (a null model), we generated randomised checklists with no preference. We created these via permutation in such a manner to (1) preserve the prevalence of all species across the observed and permuted datasets, (2) preserve the number of species on each checklist across both datasets, but (3) redistribute species identities randomly across permuted checklists. A full description of the permutation method used for our null models is provided in Appendix S2. We fit beta-distributed GAMLSS models on two versions of the eBird dataset: (1) the observed dataset and (2) the permuted dataset with species observations randomly shuffled between checklists. In both cases, we used checklist rarity score as the response variable, duration of checklist was an explanatory variable for mu (location) and sigma (scale), and we added log (duration) as an additional explanatory variable for mu, to allow for greater variability in mu at low duration values. Treating duration as the sole explanatory variable allowed us to directly estimate its relationship with rarity, so that we could later use checklist duration as a data filter to mitigate any bias that might be present. Because permuting the full dataset (171,474 checklists) was computationally prohibitive, we instead subsampled without replacement a random unweighted selection of 10,000 checklists and fit GAMLSS models on the observed and permuted versions of these subsampled datasets (see figure in Appendix S2). To account for sampling variation during subsampling, we repeated this subsampling process 1000 times (for a total of 2000 models) and calculated 95% confidence intervals for the two modelled relationships by taking the 2.5th and 97.5th percentiles of the model fits for the two dataset versions.

2.3 | Estimating Species-Specific Observer Preferences in eBird

We estimated observer species preferences in the eBird dataset by comparing the rate at which species occurred on short-duration checklists in the subsampled observed and permuted datasets. Because we hypothesised that the observed checklists would tend to have rarer species than the permuted checklists at short durations, and vice versa at longer durations, we expected to see a crossing over in the observed and permuted models of checklist rarity score and duration. As such, we defined short-duration checklists as those with durations less than the median crossover point at which the fitted rarity-duration curves of the observed model crossed over the model from the permuted data.

After calculating the median crossover duration of the rarity-duration models, we explored how often species occurred on short-duration checklists. We aggregated these rates across the 1000 subsamples, allowing us to compare the average short-duration checklist occurrence rate between the observed and permuted datasets for each species. For example, Goldcrests (*Regulus regulus*) occurred on 7065 out of 263,316 short-duration checklists across all subsamples in the observed dataset (an average short-duration checklist occurrence rate of 2.7%), and on 7941 out of 263,316 in the permuted dataset (a rate of 3.0%). We could therefore define an index of observer preference, which we call the eBird preference ratio, by the ratio between the number of times a species occurred on short-duration checklists in the observed datasets (which are affected by observer preferences) and in the permuted datasets (which are not). In the Goldcrest example, this preference ratio was 2.7%/3.0% = 0.89 or 1:1.12. To estimate precision in the eBird preference ratios for each species, we bootstrap resampled the 1000 subsamples and recalculated the preference ratio for each species in each bootstrap iteration. We repeated this process 1000 times and calculated a 95% bootstrap confidence interval by taking the 2.5th and 97.5th percentiles of the preference ratios for each species. As we hypothesised that rarity would influence observers' preferences, we explored the relationship between species' reporting rates in the eBird dataset and their eBird preference ratios. We did this by fitting a Gaussian GAM with log (eBird preference ratio) as the response variable and a Duchon smooth spline on log (reporting rate) as the explanatory variable, with smoothing parameters $k = 4$ (a maximum of 3 degrees of freedom) and gamma = 1.4. A diagram outlining each step of the eBird modelling process is provided in Figure S3 in Appendix S2.

2.4 | Estimating Species-Specific Observer Preferences in BirdTrack

To validate our analysis of the eBird dataset, we conducted a similar study of observer preferences in BirdTrack data. We leveraged list entry order information in the BirdTrack dataset to explore which species were listed first on each checklist, estimating species preferences by analysing the frequency with which species were the first-recorded observation on lists.

For each species, we calculated the proportion of sightings that were the first-recorded observation on the checklist. For example, Goldcrests were reported on 8092 checklists and were the first-recorded species on 230 of those—a first-on-list frequency of 2.8%. We compared these observed frequencies with those expected for each species. For example, on checklists with 10 species recorded, assuming no observer preferences, each species would be expected to be the first-recorded species 1/10th of the time. A species' expected first-on-list frequency is thus the average of the reciprocal number of species recorded on each checklist that recorded the species; that is:

$$\text{expected first on list frequency} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{S_i},$$

where N is the number of checklists that reported the species (for Goldcrest, 8092), and S_i is the number of species on the ith checklist that reported the species (for Goldcrest: 31, 14, 10, ..., 31). For most species, the expected first-on-list frequency was

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

between 3% and 6% (for Goldcrest: 4.4%), depending on the richness of the checklists that reported the species. We defined our index of observer preference, which we call the BirdTrack preference ratio, by the ratio between the observed and expected first-on-list frequencies for each species (for Goldcrest: 2.8%/4.4% = 0.64 or 1:1.55). Once again, we wanted to estimate variation in the BirdTrack preference ratios for each species, so we bootstrap resampled the 63,186 checklists in the BirdTrack dataset and recalculated the preference ratio for each species in each bootstrap dataset. We repeated this process 1000 times and calculated a 95% bootstrap confidence interval by taking the 2.5th and 97.5th percentiles of the preference ratios for each species. As with eBird, since we hypothesised that rarity would contribute to observers' species preferences, we explored the relationship between species' reporting rates in the BirdTrack dataset and their BirdTrack preference ratios, using the same Gaussian GAM as described above.

2.5 | Comparing Species Preferences Across eBird and BirdTrack

We compared the results of our analyses of species preferences in the eBird and BirdTrack datasets to validate the eBird preference ratio, which is a less direct but more widely applicable measure of observer preferences than the BirdTrack preference ratio. To do so, we tested the correlation between log (eBird preference ratio) and log (BirdTrack preference ratio) using the R function *cor.test()*. We tested the correlation on (1) all species for which we were able to estimate preference ratios in both datasets ($N = 242$; 56% of the 429 total species recorded across both datasets) and (2) a set of common breeding species ($N = 132$). We focused on common species as these had higher precision in the estimates of preference ratios. We defined common species as all those with overall (January–December) and breeding season (May–August) reporting rates > 1% in both eBird and BirdTrack datasets. To estimate uncertainty around both of these correlation coefficients, we took each of the 1000 bootstrap eBird and BirdTrack preference ratios for each species and calculated correlation coefficients for each bootstrap iteration, with the bounds of the 95% confidence intervals being the 2.5th and 97.5th percentiles of the distribution of these coefficients.

2.6 | Additional Analyses of Observer Preferences

To estimate the effect of the complete checklist protocol on observer species preferences in the eBird dataset, we modelled the relationship between species' rarity and the ratio of their reporting frequencies on incomplete versus complete checklists, following the method in Johnston et al. (2023). We then compared species' incomplete: complete checklist ratios with their eBird preference ratios. Since observers have near-complete freedom in the observation process for incomplete checklists, but are more constrained when conducting complete checklists, the relationship between the two is informative as to how much the complete checklist protocol mitigates observer preferences. Full methods for this analysis are detailed in Appendix S3.

We also explored the effect of novelty on observer species preferences in the BirdTrack dataset, by classifying observations according to whether or not they were an observer's first recorded observation of that species for the year and comparing the BirdTrack preference ratios of first-of-year and not first-of-year observations for each species. A full description of the methods and results for this analysis is provided in Appendix S4.

2.7 | Effects of Observer Preferences on Species Distribution Models

We explored the effects of list-initiation bias on species distribution models, a popular application of citizen science data. In particular, we fit single-season, single-species spatial occupancy models (hereafter, occupancy models) with eBird data during the breeding season (May–August) for the 132 common breeding species described above. We defined a 3 × 3 km grid over the study area (Great Britain) and treated each grid cell as a single site, even if observations came from multiple locations within the cell. We fit models on three different datasets for each species: one with all eBird data included (Full), one with only short-duration checklists included (those most affected by list-initiation bias; Short-only), and one with short-duration checklists excluded (Long-only). By comparing these three models, we aimed to assess the impact of list-initiation bias on the outputs of occupancy modelling for each species. We used 10 min as the cutoff duration for the Short/Long-only models, rather than the eBird-derived 3-min crossover duration, to be conservative in removing most list initiation-biased lists from the Long-only data. After creating each dataset, to reduce the effect of spatial sampling bias and make the occupancy models more efficient, we randomly subsampled each grid cell to have at most 10 checklists to use as 'repeat visits' to the cell in the occupancy framework, following the approach of Johnston et al. (2021). We fit the 396 models (132 species × 3 datasets) using the R function *occu()* from the package *unmarked* (Fiske and Chandler 2011). We fit an intercept-only submodel for detection, and a submodel for occupancy with latitude, longitude, median elevation, standard deviation of elevation, percentage land coverage of the seven landcover classes and coastline (true/false) as covariates. Full details of the environmental datasets we used in occupancy modelling are provided in Appendix S1.

We evaluated the effects of list-initiation biased data by comparing the average estimated occupancy probability across Britain for each dataset and species. To test if there was an overall effect of eBird preference ratio on the difference between models, we fit a simple linear model (Gaussian GLM) with the mean absolute difference in estimated occupancy (across all grid cells in Britain) between Full and Long-only models for each species as the response variable, and log (eBird preference ratio) as the explanatory variable.

3 | Results

3.1 | Modelling Species Rarity and Checklist Duration With eBird

We identified significant differences in the relationships between duration and rarity for observed and permuted datasets (Figure 1). As hypothesized, short checklists reported rarer

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

species than expected, with a slight underrepresentation of rarer species on longer lists, relative to the permuted data. Across all 1000 subsample iterations, the median crossover duration between observed and permuted models of checklist rarity was 3.3 min with a 95% confidence interval of 2.2–4.5 min. Thus, in this dataset, the effect of rarity-biased list initiation is evident in checklists shorter than 5 min.

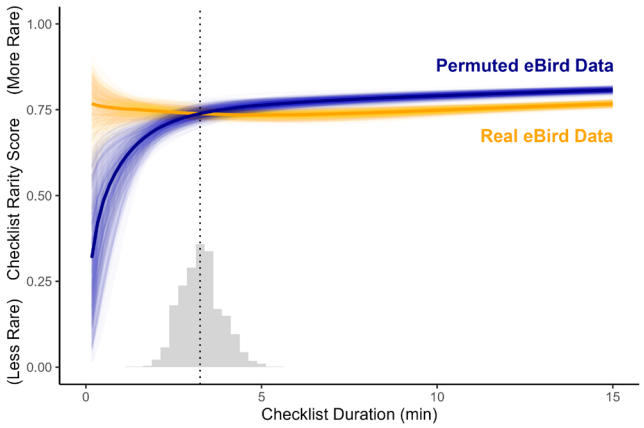


FIGURE 1 | Comparison of models of checklist rarity in British eBird data. We fit GAMLSS models on 1000 subsamples of the eBird dataset, with checklist duration used as the predictor of checklist rarity score (a normalised measure of the maximal rarity of species observed on a checklist). Orange lines show the model fits for the subsampled real (observed) eBird datasets. Blue lines show the model fits for the permuted eBird datasets, in which species observations are randomly shuffled between checklists to eliminate the effects of observer preferences. Thin lines indicate each of the 1000 model fits for the real and permuted datasets, whereas thick lines indicate the median of all 1000 model fits for both. The median crossover point (dotted line) between these two models is at 3.3 min with 95% confidence interval 2.2–4.5 min (the distribution of these crossover durations is indicated by the grey histogram); eBird checklists shorter than this duration tend to have rarer-than-expected species, which we believe to be a result of biased list initiation driven by observer preferences for rare species.

3.2 | Estimating Species-Specific Observer Preferences in eBird

Comparing occurrence rates between the observed and permuted datasets on checklists 1–3 min in duration allowed us to estimate eBird preference ratios for 259 (of 388 total) species (Figure 2). The species for which preference ratios could not be estimated were all rare (the most common being Hawfinch *Coccothraustes coccothraustes*, with a reporting rate of 0.2%). As expected, there was a negative relationship between the log eBird preference ratio and log reporting rate, with rarer species being recorded on short-duration checklists significantly more often than expected (Figure 2), although this trend was only evident across particularly rare species with reporting rates < 1%. Among 151 more common species (reporting rate > 1%), preference ratios varied from 1:3.7 to 3.9:1 (i.e., from approximately four times less frequently recorded on short-duration checklists than expected to approximately four times more frequently than expected). The species with the five highest and lowest ratios are presented in Table 1; a full list is provided in Table S1 in Appendix S5.

3.3 | Estimating Species-Specific Observer Preferences in BirdTrack

Comparing the observed and expected frequency of first-on-list observations allowed us to estimate BirdTrack preference ratios for 295 (of 394 total) species (Figure 3), which was a more direct

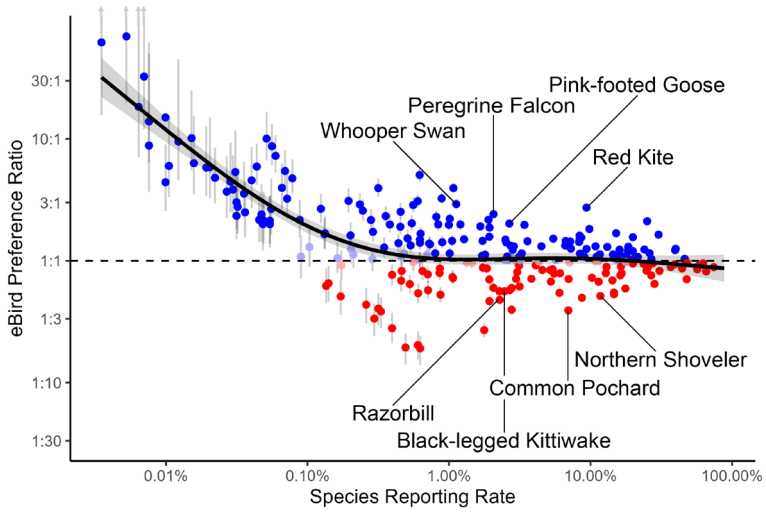


FIGURE 2 | Relationship between species reporting rate and eBird preference ratio for species in the British eBird dataset. Preference ratios with bootstrap 95% confidence intervals were able to be estimated for 259 (of 388 total) species (arrowheads indicate that the upper bound of the interval is outside the axis limits of the plot). Species points are coloured according to their estimated preference ratio: Those estimated to be overpreferred in short-duration observed checklists (compared with expected) are blue, whereas those that are underpreferred in short-duration observed checklists (compared with expected) are red; species whose preference ratios were not significantly different from 1:1 (indicated by the dashed line) are paler. There is a negative relationship between the two; rarer species (those with reporting rates < 1%) have higher preference ratios. We have labelled eight common species with reporting rates > 1%, four that occurred in short duration lists more than expected and four that occurred in short-duration lists less than expected.

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

measure of list initiation compared with the inference from short-duration eBird lists. The species for which preference ratios could not be estimated were all rare (the most common being Swan Goose *Anser cygnoides*, with a reporting rate of 0.04%). As with the eBird analysis, there was a negative relationship between the log BirdTrack preference ratio and log reporting rate (Figure 3); this trend was most evident across particularly rare species with reporting rates < 1%. Among 149 more common species (reporting rate > 1%), preference ratios varied from 1:2.8 to 2.2:1. The species with the five highest and lowest ratios are presented in Table 2; a complete table is presented in Table S1 in Appendix S5.

3.4 | Comparing Species Preferences Across eBird and BirdTrack

To validate the results of our analysis of preferences in the eBird dataset, we compared preference ratios from the eBird analysis with those from BirdTrack. We did this across (1) all 242 species for which we were able to estimate preference ratios in both datasets and (2) a set of 132 common breeding species. In both cases, there was a significant positive correlation between eBird and BirdTrack preference ratios (Figure 4; Pearson's r for all species 0.49 ($p \ll 0.001$; 95% confidence interval [0.451; 0.567]), for common breeding species 0.38 ($p \ll 0.001$; 95% confidence

TABLE 1 | Common birds in the British eBird dataset ranked by their eBird preference ratio (the five highest- and lowest-ranked species are shown here).

Rank	Common name	Scientific name	eBird preference ratio	Reporting rate (%)
1	Western Barn Owl	*Tyto alba*	3.9:1	1.1
2	Whooper Swan	*Cygnus cygnus*	2.9:1	1.1
3	Red Kite	*Milvus milvus*	2.7:1	9.3
4	Peregrine Falcon	*Falco peregrinus*	2.4:1	2.1
5	White-throated Dipper	*Cinclus cinclus*	2.2:1	1.0
147	Razorbill	*Alca torda*	1:2.1	2.3
148	Garden Warbler	*Sylvia borin*	1:2.1	1.9
149	Western Yellow Wagtail	*Motacilla flava*	1:2.5	2.8
150	Common Pochard	*Aythya ferina*	1:2.6	7.0
151	Common Scoter	*Melanitta nigra*	1:3.7	1.8

Note: The preference ratio is the ratio of how often a species occurs on short-duration checklists in the observed (real) versus permuted (simulated) datasets. A high ratio (e.g., 3.9:1) indicates the species was reported on short lists 3.9 times more often than expected based on the null permuted data. A low ratio (e.g., 1:3.7) indicates that the species was reported 3.7 times less often than expected. The reporting rate is the average occurrence of the species across all lists in the dataset.

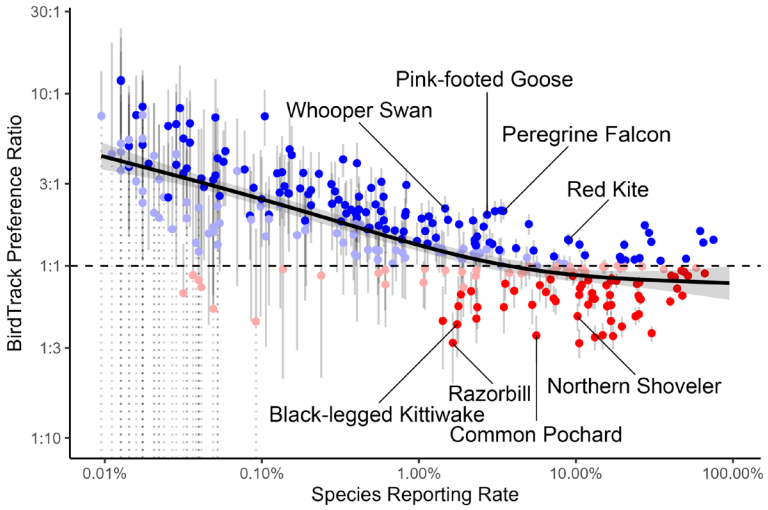


FIGURE 3 | Relationship between species reporting rate and BirdTrack preference ratio for species in the British BirdTrack dataset. Preference ratios with bootstrap confidence intervals were able to be estimated for 295 (of 394 total) species (dotted confidence intervals indicate that the lower bound of the interval is 0). Species points are coloured according to their estimated preference ratio: Those estimated to be overpreferred (they are the first-reported species on checklists more often than expected) are blue, whereas those that are underpreferred (they are the first-reported species on checklists less often than expected) are red; species whose preference ratios were not significantly different from 1:1 (indicated by the dashed line) are paler. There is a negative relationship between the two; rarer species have higher preference ratios. We have labelled eight common species with reporting rates > 1%, four that were the first-reported species on checklists more than expected and four that were the first-reported species on checklists less than expected.

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Common birds in the British BirdTrack dataset ranked by their BirdTrack preference ratio (the five highest- and lowest-ranked species are shown here).

Rank	Common name	Scientific name	Preference ratio	Reporting rate (%)
1	Whooper Swan	*Cygnus cygnus*	2.2:1	1.5
2	Pied Avocet	*Recurvirostra avosetta*	2.1:1	3.3
3	Peregrine Falcon	*Falco peregrinus*	2.1:1	3.5
4	Common Cuckoo	*Cuculus canorus*	2.1:1	3.1
5	Pink-footed Goose	*Anser brachyrhynchus*	2.0:1	2.7
145	Common Pochard	*Aythya ferina*	1:2.5	5.6
146	Reed Bunting	*Emberiza schoeniclus*	1:2.6	17
147	Common Gull	*Larus canus*	1:2.6	13
148	Razorbill	*Alca torda*	1:2.8	1.6
149	Great Black-backed Gull	*Larus marinus*	1:2.8	11

Note: The BirdTrack preference ratio is the ratio of how often a species was the first reported on a list, relative to how often it would be expected to be first on a list by chance. A high ratio (e.g., 2.2:1) indicates the species was first on a list 2.1 times more often than expected by chance. A low ratio (e.g., 1:2.8) indicates that the species was reported first on a list 2.8 times less often than expected by chance. The reporting rate is the average occurrence of the species across all lists in the dataset.

interval [0.331; 0.417])); species that were preferred by eBird users tended to be preferred by BirdTrack users, even though we quantified observer preference via different methods across the two datasets. Thus, even in the absence of information on list entry order, we have demonstrated that list initiation behaviours can be identified using the approach we present here for eBird data.

3.5 | Additional Analyses of Observer Preferences

We estimated the effect of the complete checklist protocol on observer species preferences in the eBird dataset by modelling the relationship between species' rarity and the ratio of their reporting frequencies on incomplete versus complete checklists. This relationship was very similar to the one found by Johnston et al. (2023); there is a strong effect of rarity, with less frequently encountered species being disproportionately likely to occur on incomplete checklists (Figure S4 in Appendix S3). Furthermore, when assessing the relationship between species' incomplete: complete checklist ratios and their eBird preference ratios, we found a significant positive correlation between the two across all species, but no correlation when looking solely at common species (Figure S6). This suggests that the complete checklist protocol effectively mitigates observer preferences against common species but does not mitigate observer preferences towards rare species. Full results for this analysis are detailed in Appendix S3.

We also explored the effect of novelty on observer species preferences in the BirdTrack dataset by comparing the BirdTrack preference ratios of first-of-year and not first-of-year observations for each species. As expected, preference ratios were strongly positively correlated between first-of-year and subsequent observations for each species, but most species had higher preference ratios for first-of-year observations than subsequent observations. Full results for this analysis are provided in Appendix S4.

3.6 | Effects of Observer Preferences on Species Distribution Models

We estimated breeding season distributions with occupancy models for 132 common breeding species. For each species, we produced models using three subsets of the eBird dataset: the full dataset, and subsets including either only short checklists (less than 10 min in duration) or only long checklists (at least 10 min in duration). Following data pre-processing, the Full models were fit on 5609 sites, whereas the Long-only models were fit on 5481 (a 2.3% contraction), and the Short-only models on just 651 (an 88% contraction). Models fit to Full data and Long-only data converged for all 132 species, whereas models fit to Short-only data failed to converge for many species. This was primarily due to the much smaller dataset size for Short-only models: of the 53,978 breeding season checklists used, only 8% were classified as short duration. Occupancy models showed very little difference in estimated occupancy between the Full and Long-only datasets for most species. The mean absolute difference in estimated occupancy probability across Britain between models from Full and Long-only datasets ranged between 0.001 and 0.032, with a mean difference of 0.01. There was a weak positive relationship between species' eBird preference ratios and their mean absolute difference between Full and Long-only models, with observer-preferred species tending to have slightly larger differences between the estimated occupancy from the two models (Figure 5). A selection of plots showing estimated occupancy probability and differences between the three models is provided in Figures S8–S13 in Appendix S6.

4 | Discussion

In this article, we explored the effects of observer species preferences on list initiation in semi-structured citizen science data. We demonstrated how observer preferences can lead to list-initiation bias, developed methods to quantify list-initiation bias in ecological datasets, and applied these methods to the

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

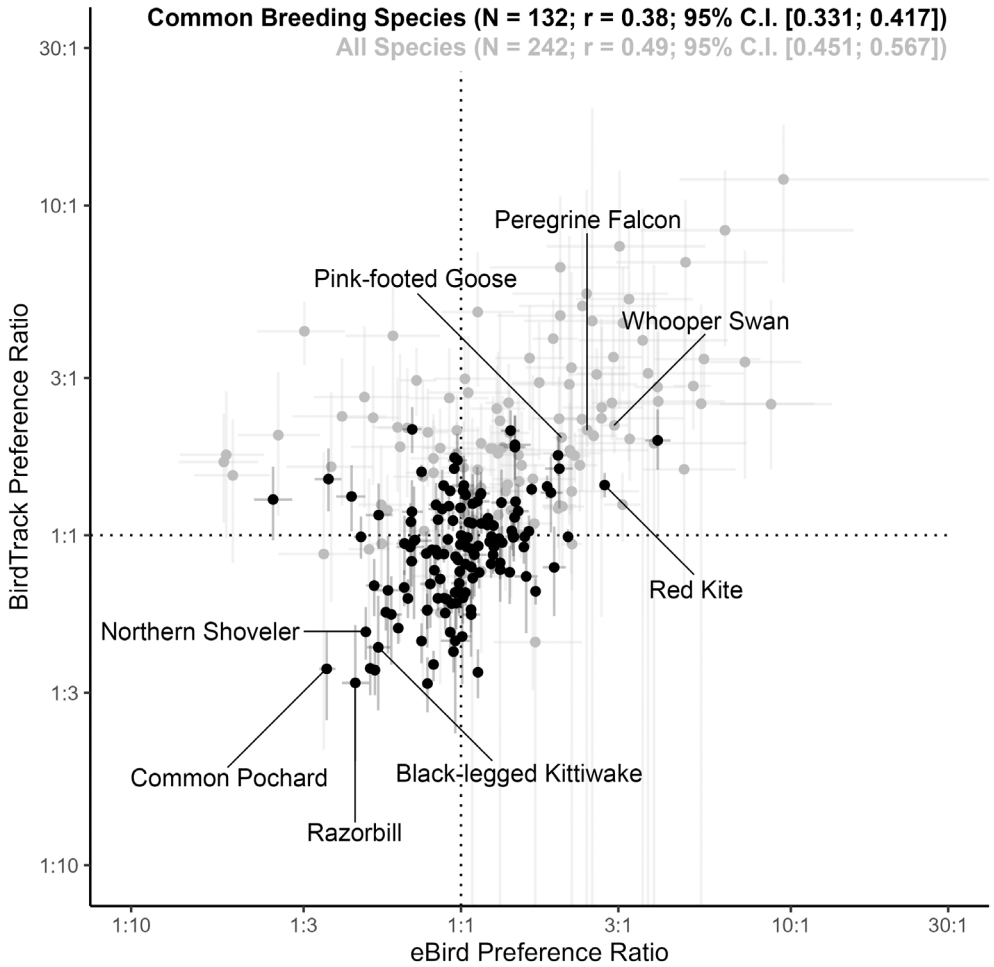


FIGURE 4 | Correlation between preference ratios derived from eBird and BirdTrack datasets. The two methods used to produce preference ratios produce broadly comparable scores for most species, suggesting that observer behaviours and preferences are similar across the two datasets.

British eBird and BirdTrack datasets, finding evidence for biased list initiation in both. The bias evident in short-duration checklists was driven by observers' species preferences, which were correlated with rarity. We also measured the effects of biased data on occupancy models fit to eBird data but found little evidence of an effect when comparing models fit with the full dataset and those fit with the filtered version of the dataset with very short-duration checklists removed, mostly due to the small fraction of data affected by list-initiation bias. We now explore this bias further and conclude by presenting recommendations for analysts and citizen science project coordinators wishing to mitigate the effects of observer preferences and their related biases.

4.1 | Why and How List-Initiation Bias Occurs

List-initiation bias occurs when citizen scientists conduct a complete survey they otherwise would not have, due to the opportunistic detection of a species of interest. We found evidence of this bias in short-duration checklists which were less than 5 min in length. We expect this is because among complete surveys, opportunistically initiated checklists tend to be shorter in duration (e.g., a few minutes long), whereas planned checklists are often much longer. For example, the opportunistic observation of an interesting species such as Peregrine Falcon *Falco peregrinus* would be unlikely to initiate an hour-long checklist that would not have happened otherwise. We have identified three factors that determine whether a species observation is notable; all three have been noted in the literature, but mostly in the context of unstructured citizen science data rather than the complete checklist semi-structured data we analysed here.

One important predictor of list initiation we found was rarity. Very short-duration eBird checklists reported rarer species than expected by chance (Figure 1), and among all observations on these short-duration checklists there was a strong relationship between species rarity and occurrence relative to our null expectation (Figure 2). We validated these results by analysing data from BirdTrack and found that rarer species were more likely to be the first-recorded species on checklists in that dataset (Figure 3). We also calculated the index of observer preference presented in Johnston et al. (2023), which compares reporting rates on incomplete versus complete eBird checklists (i.e., presence-only versus detection-nondetection data). In this analysis (Figure S4 in Appendix S3), rarer species were reported relatively more often on incomplete checklists than complete checklists, suggesting that eBird users place greater value on sightings of rare species than those of common species, and that these preferences are mitigated but not removed by the complete checklist

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

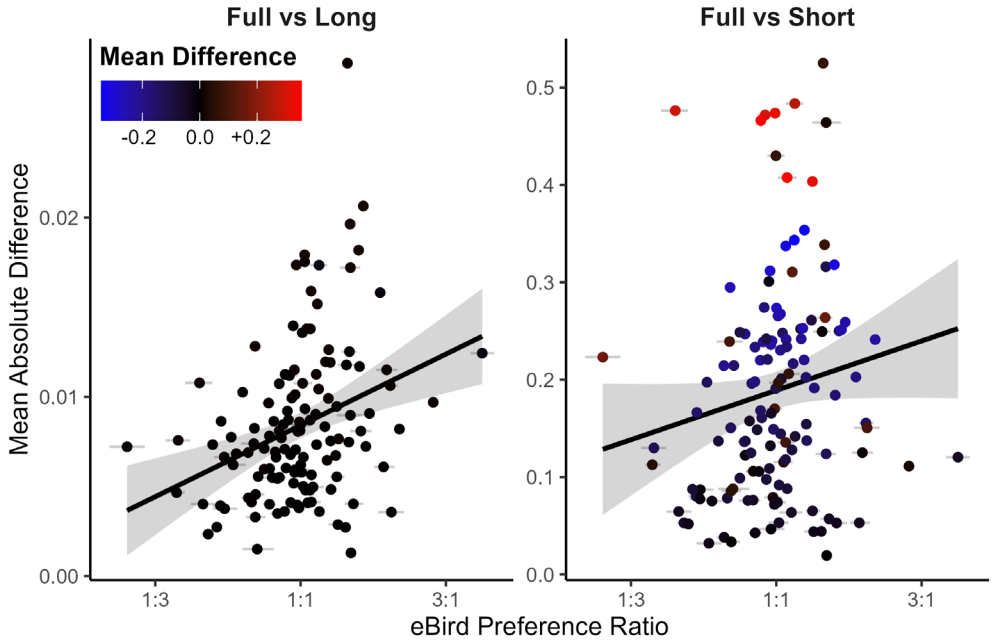


FIGURE 5 | Relationship between eBird preference ratio and the mean absolute difference in estimated occupancy between Full versus Long-only (left panel) and Full versus Short-only (right panel) models. Distributions were estimated using occupancy models for 132 common breeding species across Britain using three subsets of the eBird dataset: Full (all checklists included), Long-only (only checklists 10 min or longer included) and Short-only (only checklists shorter than 10 min included). Each point represents a species; points are coloured by the mean difference (Full minus Long or Full minus Short) in estimated occupancy between the two models. Note that both mean differences and mean absolute differences for Full versus Long-only tend to be much smaller than those for Full versus Short-only (the *y*-axis scales differ between the two panels). The black line shows the fit of a simple linear model with the mean absolute difference in estimated occupancy (across all grid cells in Britain) between Full and Long-only models for each species as the response variable, and log (eBird preference ratio) as the explanatory variable; the shaded ribbon denotes the 95% confidence interval for the estimated fit. The mean absolute difference between models for each species is the mean of the absolute differences between cell estimates of occupancy across Britain and the mean difference between models for each species is the mean of the differences between cell estimates of occupancy across Britain.

protocol (Figures S5 and S6). Observers' preferences for rare species are well documented, especially among birdwatchers (Booth et al. 2011; Randler et al. 2023), but while the effects of rarity preferences have been previously explored in unstructured citizen science schemes (Arazy et al. 2024; August et al. 2020; Bowler et al. 2022; Stoudt et al. 2022), this is the first demonstration of this bias within complete checklists in semi-structured schemes.

Novelty has also been identified as a factor influencing observer preferences in a handful of previous studies (Bowler et al. 2022; Goldstein and Stoudt 2025). We did not explore the effect of novelty on list initiation in detail in this study, but did identify a relationship between observers' first observations of species for the year and their BirdTrack preference ratio (Figure S7 in Appendix S4). This suggests that observers' motivations to add species to their personal year or life tallies (McFarlane 1994) influence their contributions to citizen science datasets, and that first-of-year observations of species are more likely to lead to list initiation than subsequent observations.

A third set of factors that potentially influences observer preferences is species traits. However, we did not conduct any formal trait-based analyses, and so we consider this to be a hypothesis in need of further exploration. As with rarity, species characteristics have been noted to affect observers' preferences (Schuetz and Johnston 2019), especially in unstructured citizen science data (Callaghan et al. 2021; Stoudt et al. 2022; van Tongeren et al. 2023), but this has received less attention in analyses of complete checklists in semi-structured datasets.

4.2 | Impacts of List-Initiation Bias on Species Distribution Models

We found limited evidence for an effect of list-initiation bias on occupancy models fit using eBird data. When we removed short-duration checklists (those most impacted by list-initiation bias) from our datasets, for most species the resultant models were very similar to those fit with the full dataset (with short-duration checklists included). The difference in estimated Britain-wide occupancy between Full-dataset and Long-only models tended to be slightly larger for more observer-preferred species (Figure 5), but the effect was minor and varied considerably in magnitude across species. This lack of support for our hypothesis is explained by the relatively small proportion of checklists removed when filtering out short-duration lists: in our case, less than 10% of checklists (and an even smaller fraction of sites) were removed as part of this process, meaning that the Long-only models were fit to very similar datasets as the Full-dataset models for most species. In some cases, and depending on the inferential goals of the analysis (van Strien et al. 2013), the removal of shorter lists may result in much greater reductions in dataset size and possibly larger effects on model outputs, and this remains an area for future research.

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.3 | Limitations and Future Study

We only explored observer preferences in one region (Britain) and year (2021) in this paper. Exploratory analyses of other regions (not presented here) suggested that the short-duration rarity effect (Figure 1) was widespread, but the exact crossover point varied. Future study could explore this variation in observer species preferences across time and space in further detail, in addition to exploring potential interactions between the observer preference-driven sampling bias we identify here and other spatiotemporal sampling biases. Furthermore, we were only able to analyse a subset of the total British BirdTrack dataset for 2021 (checklists submitted via the mobile app), which may not be representative of the broader BirdTrack dataset for that year.

With the exception of rarity, we did not explore the factors driving observers' species preferences in detail in this paper; choosing instead to focus on the overall effect of these preferences on list initiation. Future studies could explore species-specific preferences in greater depth, for example via trait-based or phylogenetic explorations of the relationship between species and observer preferences. Trait- and novelty-based observer preferences are often noted in opportunistic citizen science datasets, and we identified several traits that may have influenced observers' list initiation tendencies in the two semi-structured datasets we analysed, but these require further investigation to confirm causal relationships between species traits, observer preferences and citizen science list initiation.

One drawback of our methods to infer list-initiation bias is that they require metadata that may not always be available. Our eBird analysis required precise duration metadata; some datasets do not contain this information and thus this method would not work. For example, the BirdTrack mobile app rounds checklist durations to 5-min intervals, which we found was too imprecise to effectively estimate a crossover duration. Our BirdTrack analysis required list entry order metadata; very few datasets contain this information and thus this method cannot be used unless list entry order (or at minimum, documentation of the first-entered species on each list) is included in data collection.

Finally, although we found limited effects of list-initiation bias on occupancy models in our study, we expect that this bias may cause problems in certain circumstances, and it would thus be beneficial to explore alternative methods of mitigating them during analysis, for example by exploring the relationship between list length and presence of rare species on checklists, or testing the effect of including list length as a covariate during modelling (van Strien et al. 2013).

4.4 | Recommendations

For analysts wishing to minimise the impact of observer preferences when analysing their own datasets, we recommend the removal of incomplete and/or short-duration checklists. Alternatively, an analysis could include an indicator variable for a 'short-duration' checklist in the detectability submodel of an occupancy model. Where possible, we suggest that analysts calculate the specific crossover duration for their data, as this allows for more precise removal of those checklists most impacted by list-initiation bias. Otherwise, the 10-min cut-off used in our occupancy models is likely a conservative but sufficient approach. Short-duration checklists usually comprise a small proportion of the total dataset size and are sometimes removed anyway since they typically have lower information value than longer duration checklists because they provide a less comprehensive picture of the species present at a site.

For citizen science project coordinators wanting to help analysts mitigate the effects of any observer species preferences in their platforms, we recommend the collection of metadata that allows analysts to more directly infer list-initiation bias, for example by directly recording list entry order (as was the case for the BirdTrack mobile app, which we take advantage of here), or by asking users to indicate the first-observed species on each checklist they submit. Alternatively, it could be communicated to users that it is preferable to denote a checklist as incomplete or incidental when it has been initiated due to an opportunistic detection of a species of interest.

5 | Conclusions

Citizen science data provide a wealth of potentially valuable information that can be used to support biodiversity monitoring, research and management (Pocock et al. 2018). Semi-structured citizen science schemes allow observers freedom in how they observe and record biodiversity, often resulting in increased popularity and data volumes at the cost of additional analytical challenges associated with variation in observer behaviours (Kelling et al. 2019). In this study, we explored how observer preferences lead to biased initiation of citizen science surveys and found that list-initiation bias results in over- and under-reporting of some species in complete checklists. This bias was evident in very short-duration checklists less than 5 min long and had limited impacts on occupancy models fit to the eBird data we tested. Nevertheless, we urge analysts and project coordinators to consider this bias and its possible effects when dealing with citizen science data and encourage future work to better understand and mitigate the effects of observer preferences in citizen science.

Author Contributions

Louis J. Backstrom: conceptualization, methodology, formal analysis, writing – original draft, visualisation, project administration. **Rachel L. Drake:** methodology, writing – review and editing. **Hannah Worthington:** conceptualization, methodology, writing – review and editing, supervision, project administration. **Alison Johnston:** conceptualization, methodology, writing – review and editing, supervision, project administration.

Acknowledgements

We thank the many thousands of contributors to the eBird and BirdTrack projects, and are grateful to Philipp Boersch-Supan, Elizabeth Tatham and Simon Gillings at the British Trust for Ornithology for facilitating the use of the BirdTrack dataset. We thank Wesley Hochachka, Ben Sheldon, Richard Fuller, Philipp Boersch-Supan and many members of CREEM for helpful discussions that shaped the direction and analysis of this paper. To meet institutional and research funder open access requirements, any accepted manuscript arising shall be open access under a Creative Commons Attribution (CC BY) reuse license with zero embargo.

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Ethics Statement

Secondary data ethics approval was obtained for this project to cover the use of citizen science data from eBird and BirdTrack following the processes of the Animal Welfare and Ethics Committee at the University of St Andrews.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Code, processed data and outputs for all analyses are provided on GitHub: https://github.com/Louis-Backstrom/CS-ObserverPreferences. Raw eBird data are available at https://ebird.org/data/download and raw BirdTrack data at https://app.bto.org/data-request/new-data-request.jsp. During peer review, an anonymised version of the GitHub repository is available at https://anonymous.4open.science/r/ObserverPreferences.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ddi.70103.

Biosketch

The authors work together at the Centre for Research into Ecological and Environmental Modelling (CREEM), developing methods to derive robust ecological insights from citizen science data.

References

Arazy, O., K. Kaplan-Mintz, D. Malkinson, and Y. Nagar. 2024. "A Local Community on a Global Collective Intelligence Platform: A Case Study of Individual Preferences and Collective Bias in Ecological Citizen Science." *PLoS One* 19, no. 8: e0308552. https://doi.org/10.1371/journal.pone.0308552.

August, T., R. Fox, D. B. Roy, and M. J. O. Pocock. 2020. "Data-Derived Metrics Describing the Behaviour of Field-Based Citizen Scientists Provide Insights for Project Design and Modelling Bias." *Scientific Reports* 10, no. 1: 11009. https://doi.org/10.1038/s41598-020-67658-3.

Booth, J. E., K. J. Gaston, K. L. Evans, and P. R. Armsworth. 2011. "The Value of Species Rarity in Biodiversity Recreation: A Birdwatching Example." *Biological Conservation* 144, no. 11: 2728–2732. https://doi.org/10.1016/j.biocon.2011.02.018.

Bowler, D. E., N. Bhandari, L. Repke, et al. 2022. "Decision-Making of Citizen Scientists When Recording Species Observations." *Scientific Reports* 12, no. 1: 11069. https://doi.org/10.1038/s41598-022-15218-2.

British Trust for Ornithology. 2024. *BirdTrack*. British Trust for Ornithology. www.birdtrack.net.

Burgess, H. K., L. B. DeBey, H. E. Froehlich, et al. 2017. "The Science of Citizen Science: Exploring Barriers to Use as a Primary Research Tool." *Biological Conservation* 208: 113–120. https://doi.org/10.1016/j.biocon.2016.05.014.

Callaghan, C. T., A. G. B. Poore, M. Hofmann, C. J. Roberts, and H. M. Pereira. 2021. "Large-Bodied Birds Are Over-Represented in Unstructured Citizen Science Data." *Scientific Reports* 11, no. 1: 19073. https://doi.org/10.1038/s41598-021-98584-7.

Cornell Lab of Ornithology. 2024. "eBird Basic Dataset. Version: EBD_relDec-2024." Cornell Lab of Ornithology. https://ebird.org/data/download.

Dickinson, J. L., B. Zuckerberg, and D. N. Bonter. 2010. "Citizen Science as an Ecological Research Tool: Challenges and Benefits." *Annual Review of Ecology, Evolution, and Systematics* 41, no. 1: 149–172. https://doi.org/10.1146/annurev-ecolsys-102209-144636.

Fiske, I. J., and R. B. Chandler. 2011. "Unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance." *Journal of Statistical Software* 43, no. 10: 1–23. https://doi.org/10.18637/jss.v043.i10.

Goldstein, B. R., and S. Stoudt. 2025. "Evidence of Novelty and Specialization Behavior in Participatory Science Reporting." *Oikos* 2025, no. 6: e10938. https://doi.org/10.1111/oik.10938.

Johnston, A., W. M. Hochachka, M. E. Strimas-Mackey, et al. 2021. "Analytical Guidelines to Increase the Value of Community Science Data: An Example Using eBird Data to Estimate Species Distributions." *Diversity and Distributions* 27, no. 7: 1265–1277. https://doi.org/10.1111/ddi.13271.

Johnston, A., E. Matechou, and E. B. Dennis. 2023. "Outstanding Challenges and Future Directions for Biodiversity Monitoring Using Citizen Science Data." *Methods in Ecology and Evolution* 14, no. 1: 103–116. https://doi.org/10.1111/2041-210x.13834.

Kelling, S., A. Johnston, A. Bonn, et al. 2019. "Using Semistructured Surveys to Improve Citizen Science Data for Monitoring Biodiversity." *Bioscience* 69, no. 3: 170–179. https://doi.org/10.1093/biosci/biz010.

McFarlane, B. L. 1994. "Specialization and Motivations of Birdwatchers." *Wildlife Society Bulletin* 22, no. 3: 361–370.

Pocock, M. J. O., M. Chandler, R. Bonney, et al. 2018. "A Vision for Global Biodiversity Monitoring With Citizen Science." In *Advances in Ecological Research*, edited by D. A. Bohan, A. J. Dumbrell, G. Woodward, and M. Jackson, vol. 59, 169–223. Academic Press. https://doi.org/10.1016/bs.aecr.2018.06.003.

Randler, C., N. Staller, N. Kalb, and P. Tryjanowski. 2023. "Charismatic Species and Birdwatching: Advanced Birders Prefer Small, Shy, Dull, and Rare Species." *Anthrozoös* 36, no. 3: 427–445. https://doi.org/10.1080/08927936.2023.2182030.

Rigby, R. A., and D. M. Stasinopoulos. 2005. "Generalized Additive Models for Location, Scale and Shape." *Journal of the Royal Statistical Society: Series C: Applied Statistics* 54, no. 3: 507–544. https://doi.org/10.1111/j.1467-9876.2005.00510.x.

Schuetz, J. G., and A. Johnston. 2019. "Characterizing the Cultural Niches of North American Birds." *Proceedings of the National Academy of Sciences* 116, no. 22: 10868–10873. https://doi.org/10.1073/pnas.1820670116.

Snäll, T., O. Kindvall, J. Nilsson, and T. Pärt. 2011. "Evaluating Citizen-Based Presence Data for Bird Monitoring." *Biological Conservation* 144, no. 2: 804–810. https://doi.org/10.1016/j.biocon.2010.11.010.

Stoudt, S., B. R. Goldstein, and P. de Valpine. 2022. "Identifying Engaging Bird Species and Traits With Community Science Observations." *Proceedings of the National Academy of Sciences of the United States of America* 119, no. 16: e2110156119. https://doi.org/10.1073/pnas.2110156119.

Stuber, E. F., O. J. Robinson, E. R. Bjerre, et al. 2022. "The Potential of Semi-Structured Citizen Science Data as a Supplement for Conservation Decision-Making: Validating the Performance of eBird Against Targeted Avian Monitoring Efforts." *Biological Conservation* 270: 109556. https://doi.org/10.1016/j.biocon.2022.109556.

van Strien, A. J., C. A. M. van Swaay, T. Termaat, and V. Devictor. 2013. "Opportunistic Citizen Science Data of Animal Species Produce Reliable Estimates of Distribution Trends if Analysed With Occupancy Models." *Journal of Applied Ecology* 50, no. 6: 1450–1458. https://doi.org/10.1111/1365-2664.12158.

van Tongeren, E., G. Sistri, M. Bonifacino, et al. 2023. "Unstructured Citizen Science Reduces the Perception of Butterfly Local Extinctions: The Interplay Between Species Traits and User Effort." *Biodiversity and Conservation* 32, no. 14: 4701–4718. https://doi.org/10.1007/s10531-023-02721-9.

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** Datasets. **Appendix S2:** Modelling with eBird. **Figure S1:** Calculation of checklist rarity scores for eBird checklists. **Figure S2:** Randomisation of eBird checklists. **Figure S3:** Schematic diagram outlining key steps in the analysis of eBird data to calculate preference ratios. **Appendix S3:** Comparison of incomplete and complete data. **Figure S4:** Relationship between species reporting rate (a proxy for rarity) and the ratio of reporting rates on incomplete and complete checklists (following the method presented in Johnston et al. 2023). **Figure S5:** Four hypotheses for the relationship between the eBird preference ratio and the incomplete: complete checklist ratio (following the method presented in Johnston et al. 2023). **Figure S6:** Correlation between the eBird preference ratio and the incomplete: complete checklist ratio (following the method presented in Johnston et al. 2023). **Appendix S4:** First-of-year effects in BirdTrack. **Figure S7:** Comparison of BirdTrack preference ratios (a measure of observer preference) for each species when observations are classified according to whether or not they are an observer's first observation for the year of that species or not. **Appendix S5:** Full results. **Table S1:** eBird and BirdTrack preference ratios and associated 95% bootstrap confidence intervals for all 429 species in the British eBird or BirdTrack datasets. **Appendix S6:** Example occupancy model plots. **Figure S8:** Comparison of estimated occupancy probability across Britain for three occupancy models of Eurasian Skylark (*Alauda arvensis*), an observer-neutral species (eBird and BirdTrack preference ratios ≈1). **Figure S9:** Difference between full (all eBird data) and long-only (short-duration checklists excluded), and between full and short-only (long-duration checklists excluded) occupancy models of Eurasian Skylark (*Alauda arvensis*), an observer-neutral species (eBird and BirdTrack preference ratios ≈1). **Figure S10:** Comparison of estimated occupancy probability across Britain for three occupancy models of Common Pochard (*Aythya ferina*), an observer-neglected species (eBird and BirdTrack preference ratios <1). **Figure S11:** Difference between full (all eBird data) and long-only (short-duration checklists excluded), and between full and short-only (long-duration checklists excluded) occupancy models of Common Pochard (*Aythya ferina*), an observer-neglected species (eBird and BirdTrack preference ratios <1). **Figure S12:** Comparison of estimated occupancy probability across Britain for three occupancy models of Red Kite (*Milvus milvus*), an observer-preferred species (eBird and BirdTrack preference ratios >1). **Figure S13:** Difference between full (all eBird data) and long-only (short-duration checklists excluded), and between full and short-only (long-duration checklists excluded) occupancy models of Red Kite (*Milvus milvus*), an observer-preferred species (eBird and BirdTrack preference ratios >1).

14724642, 2025, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ddi.70103 by University Of St Andrews University, Wiley Online Library on [13/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License