The Effect of Birdsong on Self-Reported Mental Well-Being

By ANTHONY HIGNEY, CLAIRE BUCHAN, SIMON BUTLER, LIAM CROWTHER, MELISSA MARSELLE, ELEANOR RATCLIFFE, KONRAD UEBEL, AND NICK HANLEY.*

Does birdsong increase mental well-being? Green spaces can provide important ecosystem services such as improved mental well-being, but which characteristics of green space drive this effect remains unclear. One candidate is the natural soundscape. While prior studies link birdsong and other natural sounds to short-term improvements in well-being, little is known about effects beyond the short-term. We provide the first causal estimates to move beyond short-term mental well-being effects of natural soundscapes, using a pre-registered analysis plan. We use a unique dataset combining granular acoustic data across Great Britain with panel survey data on mental well-being. Exploiting plausibly exogenous seasonal changes in birdsong and a difference-in-differences design, we find precise null effects on mental well-being. Our results suggest seasonal increases in natural soundscapes do not increase self-reported well-being, in contrast to short-term experimental findings. Our findings highlight the need to identify which specific green space attributes drive mental well-being improvements.

JEL: Q57; I31
Keywords: Ecological Economics; Ecosystem Services; Happiness, Well-Being

Teaser Birdsong may improve short-term mental well-being, but we provide the first causal evidence that birdsong does not increase mental well-being over extended periods.

I. Introduction

Depression and anxiety have increased sharply among young people in many countries over the last decade (Blanchflower et al., 2024), with some suggesting this is due to smartphones and social media (Burn-Murdoch, John, 2023). Nature is viewed as a source of stress reduction and restoration, and many studies have reported these positive effects (Twohig-Bennett and Jones, 2018). Environmental economists have long been interested

* Anthony Higney: University of Glasgow, Anthony.Higney@glasgow.ac.uk, corresponding author. Claire Buchan: University of East Anglia, c.buchan@uea.ac.uk. Simon Butler: University of East Anglia, simon.j.butler@uea.ac.uk. Liam Crowther: University of East Anglia, L.Crowther@uea.ac.uk. Melissa Marselle: University of Surrey, m.marselle@surrey.ac.uk. Eleanor Ratcliffe: University of Surrey, eleanor.ratcliffe@surrey.ac.uk. Konrad Uebel: University of Surrey, k.uebel@surrey.ac.uk. Nick Hanley: University of Glasgow, Nicholas.Hanley@glasgow.ac.uk.

in quantifying such benefits, or *ecosystem services* (Krutilla, 1967; Costanza et al., 1998; Turner and Daily, 2008). Ecosystem services that increase mental well-being may help offset the negative trend in mental well-being. However, a consistent finding in reviews of the literature is that we know little about which *specific features* of natural environments drive these effects and which causal pathways from green space to mental health are the most important (Markevych et al. (2017), Frumkin et al. (2017), Hartig et al. (2014)). Furthermore, in a review of 215 studies Marselle et al. (2019) found that while there was a wealth of evidence indicating green space was good for mental well-being, almost all studies were for *short-term* exposure, *short-term* mental health outcomes.

We move beyond short-term measures of mental well-being to seasonal changes, and focus on a single characteristic of green spaces: the natural sounds, or *soundscape*. In particular, birdsong. In temperate regions, birdsong is the main contributor to diurnal natural soundscapes. A wealth of qualitative studies indicate people find natural sounds, particularly birdsong, pleasant (see Ratcliffe (2021) for a review). In one meta-analysis, birdsong had the highest effect on short-term stress reduction of all natural sounds (Buxton et al. (2021)). Recent systematic reviews have also shown that many quantitative studies find birdsong is beneficial, at least in the short-term, across a range of psychological outcomes (Aletta, Oberman and Kang (2018); Ratcliffe (2021); Beute et al. (2023)). This would suggest bird vocalisations provide a valuable ecosystem service.

Our contribution is to provide novel, *causal* estimates that move beyond *short-term* impacts to *seasonal* impacts of one aspect of green space, natural soundscapes, on mental well-being using a pre-registered analysis plan. Our study combines new and unique acoustic data at a granular level across Great Britain with survey data on self-reported mental well-being. We use a data source of local soundscapes for every part of Great Britain in both winter and spring, since we know that bird populations vary between these two periods (in terms of which species are typically present in any location), and their vocal activity also changes. In a two-wave online panel survey of 727 British households, respondents gave mental well-being scores (using the WHO5 index) in both winter and spring. They also gave us location data. This allowed us to combine the survey and soundscapes data to analyse how local soundscapes affect mental well-being scores. We argue that measuring mental well-being over two different seasons, and using simulated natural soundscapes around each respondent's residence, allows us to explore the effects of soundscapes on well-being over a period of months, in contrast to most studies which only measure short-run effects (e.g. through playing sounds to respondents in a lab). Finally, we test both correlations and, in order to get a causal estimate of the effect of natural soundscapes, provided the assumptions are true as we discuss in section VI, we use up-to-date methods for difference-in-differences with continuous data (de Chaisemartin, D'Haultfœuille and Vazquez-Bare, 2024).

In all of our estimates, we do not find evidence that seasonal cumulative exposure to increased birdsong metrics increases mental well-being. This result holds for three acoustic indices of treatment variables. For two of them (acoustic complexity and the bio-acoustic index) we find precise nulls. For the remaining variable (acoustic entropy) we also do not reject the null of no effect, but the intervals are extremely wide. We carry out a variety of robustness checks and find similar results in all cases.

Our results suggest that, while previous research indicates that birdsong positively affects short-term mental well-being, it may not affect seasonal mental well-being. This means, as an ecosystem service, bird vocalisation may only provide short term value for mental well-being. For research that aims to decompose green space attributes to see why it is beneficial, it may be that other attributes are more important than the sounds, or that attributes must be present in combination.

II. Theoretical Background

There are two main theories in the psychological literature that detail how natural sounds can influence mental well-being. These are Stress Reduction Theory (SRT, Ulrich (1983)) and Attention Restoration Theory (ART, Kaplan and Kaplan (1989)). ART deals with reducing fatigue and restoring attention and concentration to cognitively demanding tasks. While an important mechanism, we believe for mental outcomes beyond the short-term, such as those we measure, SRT is more pertinent. SRT holds that experiences with natural sound can restore mental well-being and reduce longer term stress. This comes about through attached meanings and memories associated with the sounds (Ratcliffe, Gatersleben and Sowden, 2016). This mechanism may be mediated by other elements in the soundscape, which could interfere with this semantic connection. Uebel et al. (2021) conducted a lab experiment with 162 participants in Australia and found that perceived restorativeness increases with bird species richness, but the effect is lowered as traffic volume increases. A similar effect was found in Uebel et al. (2025*b*), who use the same treatment metrics as we use here, where they also found higher soundscape complexity drove higher perceptions of bird species richness. This may be crucial as Uebel et al. (2025*a*) found, in an online experiment of 1529 participants in the UK, that the relationship between short-term well-being and short-term birdsong exposure was entirely mediated by perceived bird species diversity. That is, the objective level of bird diversity the sounds played to participants represented only affected well-being through their influence on perceptions of bird diversity. If a highly diverse soundscape was played to someone, but they perceived it as low diversity, there would be no effect on well-being. This is the biodiversity-health hypothesis.

We expect any seasonal, rather than short-term, effect of birdsong to be primarily through the SRT mechanism. If birdsong does reduce stress then we would expect that increased, cumulative exposure would lead to higher mental well-being outcomes. Following the results of Uebel et al. (2025*a*) we would expect this to be caused by higher perceptions of local biodiversity. Although we cannot distinguish between the two mechanisms in our study or the exact causal pathway.

III. Data

Our design combines two sources of data: a panel survey for the outcome and control variables, combined with acoustic indices from simulated soundscapes calculated from

UK Bird Atlas data (Balmer et al., 2013) for the treatment variables.

A. Treatment Variables

We use three, related treatment variables to represent birdsong. All three are acoustic indices taken from modelled soundscapes. These soundscapes use UK Bird Atlas data combined with sound files from Xeno Canto www.xeno-canto.org. These acoustic indices only measure the bird song and bird calls for that local area, not other parts of the natural or anthropogenic soundscape. Therefore we are isolating the change in acoustic metric caused by local bird spceies alone.

Species presence and abundance data come from the UK Bird Atlas 2007–11 (Balmer et al., 2013), which used around 40,000 expert volunteers to sample 2km × 2km tetrads in both winter and breeding seasons over 4 years. All bird visual and auditory detections were recorded and the species noted. This gave a detailed survey of both species presence and abundance for 216 million individual birds and 520 species[1].

For each species and site, we used the maximum seasonal count and converted this to individuals per km². These were adjusted using species-specific aural detectability estimates from BBS data (2014–2019), ensuring that only likely vocalizing individuals were included.

Species-specific recordings were sourced from Xeno-Canto. The Xeno Canto site has around 900,000 recordings of bird vocalisations from all over the world. We selected only high-quality files from the site (Quality A) under 60 minutes, with vocalization types matched to season. Each soundscape was built by probabilistically selecting and inserting recordings into a 60-second audio file, using randomised timing and volume to simulate distance. Sound files were bandpass filtered (300–12,000 Hz) and standardised in format. Species with no available recordings (0.003% of records) were dropped.

We created a 60-second sound file initially containing only low-volume (vol 0.0005) white noise, which we populated with species recordings determined by the site-season densities. For each Atlas record (a species density at a given site in a given season), the density was probabilistically rounded up or down to an integer count value of individuals. For each such individual, we then used a second-order application of the BBS-derived aural detection probability – as visually detected birds may subsequently vocalise – to determine whether to include it in the final construction. For individuals carried forward for inclusion, we randomly selected a downloaded recording for the relevant species and inserted this into the sound file at a starting point randomly drawn from 0 s to 35 s (allowing all 25-second recordings to complete within the 60-second soundscape). The volume was randomly determined, drawn from a uniform distribution (Morrison et al., 2021). The construction process was repeated for each record at a site, thereby overlaying species into the same soundscape.

As the random elements of the construction process introduce a degree of stochastic-

[1] See https://www.bto.org/our-science/projects/birdatlas for more information about the survey.

ity, we repeated this process 50 times for each soundscape. All recordings were inserted relative to a calibration tone removed prior to acoustic analysis. All audio processing was carried out using the open-source software Sound eXchange (SoX; `https://sourceforge.net/projects/sox/`).

From these soundscapes, three acoustic metrics were measured: Acoustic Complexity Index (ACI), Bioacoustic Index (Bio), and Acoustic entropy (H). These are "best-practice" metrics that characterize natural soundscapes over the long term into numeric values, as described in Abrahams et al. (2023), and we follow their convention. See Ratcliffe, Gatersleben and Sowden (2016) and Ratcliffe (2021) for more on why we would expect such variables to positively affect our outcome variable. See Pieretti, Farina and Morri (2011), Boelman et al. (2007), and Sueur, Aubin and Simonis (2008) for more information on each metric.

B. Survey Data

We carried out an online panel survey using a Qualtrics sample frame. It was designed to be nationally representative for the UK in age, gender and regional distribution. A pilot was carried out in late 2023, and the first wave of 3394 respondents was sampled in early 2024, during the winter. As we said in the pre-analysis plan "We expect high rates of attrition so 3,500 will be sampled in the first wave, so that a minimum of 700 can be sampled in the second wave (with 1000 expected)" (Higney et al., 2024). The second wave was carried out during peak birdsong activity in the UK, in early May. A total of 974 respondents from the first wave answered the second wave, giving an attrition rate of 71% as expected in our pre-analysis plan. Of these 974, only 727 gave consistent location data for both waves. For some of the respondents who did not give consistent data, they may have moved, but without matching location data across both waves respondents cannot be linked to a single soundscape for that period and so were excluded. Our effective sample across both waves then is the 727 remaining.

To test for differential attrition, we carry out a covariate balance test, as specified in our pre-analysis plan. This is a check on whether the sample has differential attrition between waves 1 and 2. We carry out this test using the (Hansen and Bowers, 2008) Bonferroni corrected joint covariate balance test. We find, as shown in table 1, that there has been differential attrition. This is mainly due to young people dropping out of the sample in wave 2. We can see that all 3 of the z scores that are at least statistically significant at the 10% level, after Bonferroni correction, are age group indicators. The leads to the overall χ^2 test to have a large value and therefore a very low p-value.

This does not affect the internal validity of our estimates, as we estimate the WAOSS, which is analogous to the average effect of treatment on the treated, the ATT. Therefore, we still get a causal estimate for our sample. However, it can mean that, if the treatment effect for young people is very different, then our estimates cannot be extrapolated to groups of younger people.

TABLE 1—BALANCE TEST RESULTS

Variable	Treatment	Control	Adj. Diff	Std. Diff	Z-Score
WHOfive	14	14	-0.25	-0.04	-1.2
Noise Sensitivity	17	17	-0.14	-0.02	-0.56
Education (Some Secondary)	0.030	0.023	0.0071	0.04	1.2
Education (Vocational)	0.21	0.20	0.0085	0.02	0.56
Education (Some University)	0.073	0.084	-0.011	-0.04	-1.1
Education (Bachelor's Degree)	0.26	0.28	-0.015	-0.03	-0.89
Education (Graduate)	0.13	0.12	0.0026	0.01	0.21
Age 25-34	0.13	0.20	-0.068	-0.18	-4.7***
Age 35-44	0.19	0.16	0.03	0.08	2.1
Age 45-54	0.23	0.17	0.058	0.15	3.9***
Age 55-64	0.20	0.15	0.044	0.12	3.1*
Age 65+	0.20	0.19	0.013	0.03	0.90
Income (<£20k)	0.22	0.22	0.0001	0.00	0.01
Income (£40k-£60k)	0.23	0.22	0.01	0.02	0.63
Income (£60k-£100k)	0.15	0.14	0.0076	0.02	0.57
Income (>£100k)	0.045	0.041	0.0041	0.02	0.54
Local Greenspace (Daily)	0.036	0.037	-0.0006	0.00	-0.09
Local Greenspace (Never)	0.077	0.055	0.023	0.09	2.5
Local Greenspace (Once a Week)	0.19	0.19	0.0036	0.01	0.24
Vehicle (Yes)	0.77	0.78	-0.012	-0.03	-0.80
Urban (Town)	0.34	0.31	0.025	0.05	1.4
Urban (Village)	0.15	0.13	0.026	0.08	2.0
Urban (Inner City)	0.13	0.16	-0.032	-0.09	-2.3
Urban (Rural)	0.045	0.057	-0.012	-0.05	-1.4
Country (Scotland)	0.096	0.080	0.016	0.06	1.5
Country (Wales)	0.045	0.050	-0.0046	-0.02	-0.56
Overall Test: $\chi^2 = 126$, df = 41, $p = 1.6 \times 10^{-10}$					

Bonferroni corrected joint covariate balance test of (Hansen and Bowers, 2008)
WHOfive is the total score on the WHO mental well-being questionnaire and NS is the Noise Sensitivity score
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

OUTCOMES VARIABLE

Our outcome variable is the change between winter and spring of an individual's WHO-5 Index. This is a set of five questions developed by the World Health Organisation (WHO) to assess self-reported mental well-being. Responses are on a 6-point Likert scale ranging from "All of the Time" to "Not at All". Participants are asked to rate, based on their lived experiences in the last two weeks, the following five questions using the scale:

1) I have felt cheerful and in good spirits.
2) I have felt calm and relaxed.
3) I have felt active and vigorous.
4) I woke up feeling fresh and rested.
5) My daily life has been filled with things that interest me.

The responses are given a score, where ”All of the Time” = 5, and ”Not at All” = 0. These are summed into an index score with the values 0-25. This is the WHO-5 Index score for that respondent. The values are normally then multiplied by four but this step is superfluous and we do not do that.

The WHO-5 index has high validity, sensitivity and specificity as well as being only five quick and non-invasive questions (Topp et al., 2015). Therefore, we judge it is a suitable metric for gauging self-reported mental well-being through a survey.

Control Variables

Finally, we include a number of control variables at the level of the individual respondent in some model specifications. These variables are: income, age, education, self-reported visits to local green spaces, access to a garden, self-reported visits to other green spaces, access to private transport, and is the area urban, rural, a village or suburban, as well as which country/region of the UK they live in. We also include a measure of noise sensitivity. This is measured using the 5-item test of Benfield et al. (2014), and gives a score of 0-30, with 30 being the most sensitive to noise. See online appendices.

IV. Empirical Framework

Our main hypotheses are that increases in soundscape metrics, represented by our three acoustic indices, will increase self-reported mental well-being, as measured by the WHO5 questions. The difficulty in measuring this relationship directly, with a selection on observables strategy, is that this may be insufficient to identify the underlying relationship. For instance, if people sort into areas of higher or lower soundscape metrics based on unobservable characteristics, such as preferences, it is possible we would not see any relationship in a simple, cross-sectional regression.

To illustrate this, see figure 1. This is a causal diagram in the form of a directed acyclic graph (DAG). This shows our assumed model of the path through which birdsong affects mental well-being (Y). We assume that the birdsong actually heard by the people in our sample is a product of the characteristics of the area (urban/rural, near roads/isolated etc) and the characteristics of the local bird population (which species, and the abundance of individual species). These bird characteristics are themselves the result of local area characteristics (some species prefer wooded areas for example) and the season (different birds in spring compared to winter). The area characteristics are in turn affected by the season (e.g. deciduous trees lose their leaves in winter). All of these variables may affect mental health, Y.

We can see that time (as in the seasonality of birdsong) and the area characteristics (which are themselves changing with the seasons) if uncontrolled for will open a backdoor path to our outcome Y. For example, areas with higher levels of birdsong may be better for mental well-being for other reasons, or people with higher mental well-being may sort into areas with higher levels of birdsong. Therefore, our research design must take into account these threats to identification of a causal effect.

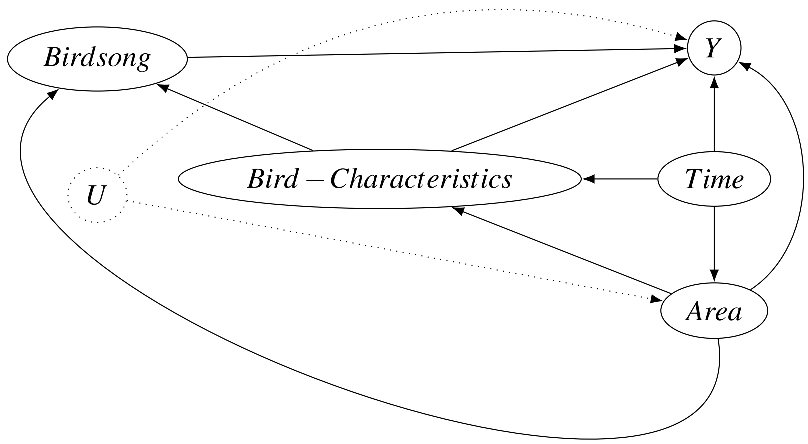

FIGURE 1. DAG MODEL OF THE EFFECT OF BIRDSONG ON MENTAL WELL-BEING

A. *Identification Strategy*

Our identification strategy relies on the plausibly exogenous change in birdsong between winter and spring, after conditioning on area characteristics and the time or season. We cannot separate bird characteristics from birdsong, so our treatment effect is better thought of as the total effect of birds on mental well-being. To estimate this, following our pre-analysis plan (Higney et al., 2024), we carry out a difference-in-differences design. The use of panel data allows us to take into account individual and area fixed effects, therefore controlling for time-invariant unobserved factors that could impede the cross-sectional analysis. The time fixed effects allow us to control for area-invariant effects of moving from winter to spring. Further information about the area is also incorporated into the design by the use of additional controls, as mentioned in section III.

However, given we have continuous treatment variables (here, treatment is equivalent to exposure to bird song), and no untreated units (everyone in the sample may hear at least some bird song), as well as no units who do not change treatment levels (since bird populations vary between the two sampling periods), we cannot estimate the causal impact of birdsong on mental well-being using two-way fixed effects. Two-way fixed effects is not robust to treatment effect heterogeneity in the presence of continuous treatment, without stronger assumptions than we make here, and we may not be able to identify the effect of treatment at all (Callaway, Goodman-Bacon and Sant'Anna, 2024). As Callaway, Goodman-Bacon and Sant'Anna (2024) show, two-way fixed effects with continuous treatment can put negative weights on units with below average doses of treatment, and positive for above average doses, meaning it likely does not estimate an estimand of any interest to the researcher. Instead then, we will use the method in de Chaisemartin, D'Haultfœuille and Vazquez-Bare (2024). This allows for treatment heterogeneity, a continuous treatment variable, and, importantly, no untreated units, or units that do not change treatment. This method relies on weaker assumptions than using two-way fixed effects, but requires the existence of "quasi-stayers". That is, units who change treatment an arbitrary, small amount. Within our sample there are areas with far less change in birdsong between winter and spring than others. For all three acoustic indices the minimum absolute value of change is less than 0.01 standard deviations. These "quasi-stayers" show our use of de Chaisemartin, D'Haultfœuille and Vazquez-Bare (2024) may be reasonable, and they allow us to estimate the counter-factual trend if our assumptions hold.

The crucial assumption, the one that is different from standard difference-in-difference designs, is that, conditional on our covariates, units that experience the same level of treatment in period one, **on average** would have had the same changes in outcome **but for the change in treatment intensity**. This is a conditional parallel trends assumption.

Formally:

$$E[\Delta Y(d) \mid \mathbf{x}, D_1 = d, D_2] = E[\Delta Y(d) \mid \mathbf{x}, D_1 = d],\ \forall d \in D_1. \tag{1}$$

Where ΔY is the change in the outcome variable between the two periods, D_1 is the treatment dose level in period 1, D_2 the dose in period 2, $\mathbf{x}$ is the value of the covariates we use and assume are needed for this conditional parallel trends assumption to hold. Other standard difference-in-difference design assumptions, such as overlap and the stable unit treatment value, are also required as normal.

To illustrate what this parallel trends assumption means, imagine there are many people in different areas, but they have similar levels of income, education, age, they all live in a town, have a garden and a car, they all access greenspace a similar amount of times, all are similarly sensitive to noise, and *in winter*, all experience a similar level of birdsong according to our chosen soundscape metric. There are differences in mental well-being between these people. Some have higher well-being, some lower. For reasons uncorrelated with our soundscape metrics, well-being will change, but if, and this is our counterfactual, all units kept the same level of acoustic metric in spring as in winter, then the *average* change in well-being would only depend on their winter level, and their characteristics x, but not on what birdsong level they actually receive in spring. There are ways this could be violated, but many of these would result in a positive bias (for example, people in high birdsong areas plan to do more leisure activity in spring, which increases their mood). Given we have a null result in a one-tailed test, positive bias is less of an issue, However, there could also be negative bias. For example, high spring birdsong areas could have more pollen, and this is not captured by our covariates on greenspace usage and access, which leads to lower mental well-being for hay-fever sufferers. In this case our causal estimate would be biased downwards.

These types of unobservable differences are the main threat to identification. We cannot test directly this parallel trends assumption. Nor can we view the trends over time, as we have only two periods and no untreated units. What we do instead is examine, in section V-C, whether units with the same level of winter acoustic metric, but different changes in acoustic metric from winter to spring, have different patterns of changes in the WHO5 index. If so, it would imply that they may not be a good counterfactual trend for each other. This would mean our parallel trends assumption is likely violated. however, we show that there appears to be no difference in pattern of change in WHO5 scores.

Our main estimand is the weighted average marginal effect of treatment. It is called the Weighted Average of Switchers' Slopes (WAOSS). This is the average effect of moving from the treatment in period one to the treatment in period two, scaled by the average change in treatment intensity. It is analogous to the Average Treatment on the Treated, but more suitable for continuous treatment. We estimate the WAOSS by:

$$\hat{\theta} = \frac{\sum_{i=1}^{N} S_i(\Delta Y_i - g_{\hat{\lambda}}(D_{1i}, 0))}{\sum_{i=1}^{N} |\Delta D_i|} \tag{2}$$

Where S_i is the sign of the change in treatment intensity for unit i (if positive then the acoustic metric increased and vice versa), ΔY_i is the change in the outcome variable between the two periods for unit i, $|\Delta D_i|$ is the absolute change in treatment intensity, and $g_{\hat{\lambda}}(D_{1i}, 0))$ is the imputed counterfactual trend the unit would have experienced if they had not changed treatment intensity. That is, $g_{\hat{\lambda}}(D_{1i}, 0))$ is the imputed value of Y_{it} when treatment in period 1 is D_{1i} and there is no change in treatment in period 2. $g_{\hat{\lambda}}(D_1, \delta))$ means treatment in period 1 is D_1 and treatment in period 2 is $D_1 + \delta$.

For a given level of treatment the function can be separated into the non-treatment change (trend) and the change due to the change in treatment, given the parallel trend assumption (de Chaisemartin, D'Haultfœuille and Vazquez-Bare, 2024):

$$\begin{aligned} g(d_1, \delta) = {} & \underbrace{E[Y_2(d_1) - Y_1(d_1) | \mathbf{x}, D_1 = d_1]}_{\text{Trend Effect}} \\ & \underbrace{+ \delta E\left[\frac{Y_2(d_1+\delta) - Y_2(d_1)}{\delta} \middle| \mathbf{x}, D_1 = d_1, \Delta D = \delta \right]}_{\text{Change in Treatment Effect}} \end{aligned} \tag{3}$$

We assume the form of the $g_{\hat{\lambda}}(d_1, \delta)$ function with a parametric linear model:

$$E[Y_2(d_1) - Y_1(d_1)|\mathbf{x}, D_1 = d_1] = \lambda_1 + \lambda_2 d_1 + \mathbf{x}_i \boldsymbol{\psi_1} + d_1 \cdot \mathbf{x}_i \boldsymbol{\psi_2}$$

$$\delta E\left[\frac{Y_2(d_1+\delta) - Y_2(d_1)}{\delta}\middle|\mathbf{x}, D_1 = d_1, \Delta D = \delta\right] = \lambda_3\delta + \lambda_4 d_1 \cdot \delta + \lambda_5\delta^2 + \delta \cdot \mathbf{x}_i\boldsymbol{\psi_3} + d_1 \cdot \delta \cdot \mathbf{x}_i \boldsymbol{\psi_4} + \delta^2 \cdot \mathbf{x}_i \boldsymbol{\psi_5} \tag{4}$$

We will estimate these terms by a regression in the form:

$$\Delta Y_i = \lambda_1 + \lambda_2 d_1 + \mathbf{x}_i \boldsymbol{\psi_1} + d_1 \cdot \mathbf{x}_i \boldsymbol{\psi_2} + \lambda_3\delta + \lambda_4 d_1 \cdot \delta + \lambda_5\delta^2 + \delta \cdot \mathbf{x}_i\boldsymbol{\psi_3} + d_1 \cdot \delta \cdot \mathbf{x}_i \boldsymbol{\psi_4} + \delta^2 \cdot \mathbf{x}_i \boldsymbol{\psi_5} \tag{5}$$

Essentially, this is a two-step estimator. First we estimate the regression in (5). The trend without a change in treatment (i.e. $g_{\hat{\lambda}}(D_1, 0))$ is represented by the first four terms in that equation. In the second step, we use those four terms as the counterfactual trend the unit would have experienced but for the change in treatment. We plug those values into equation (2), as our $g_{\hat{\lambda}}(D_1, 0)$ function. Given our assumptions, this allows us to plausibly estimate the weighted average treatment effect of bird song on mental well-being. In all cases we estimate standard errors with the influence function as in de Chaisemartin, D'Haultfœuille and Vazquez-Bare (2024). For each acoustic index, we test the null hypothesis with a one-tailed test, as specified in our analysis plan.

Our hypotheses are:

1) **H1**: An increase in the acoustic complexity index increases self-reported mental well-being
   H0: An increase in the acoustic complexity index does not increase self-reported mental well-being
2) **H1**: An increase in bio acoustic index increases self-reported mental well-being
   H0: An increase in bio acoustic index does not increase self-reported mental well-being
3) **H1**: An increase in acoustic entropy score increases self-reported mental well-being[2]
   H0: An increase in acoustic entropy score does not increase self-reported mental

[2] The pre-analysis plan stated a different H1, namely that the effect would be non-linear for acoustic entropy. However, due to the small amount of variation in the acoustic entropy data, addition of non-linear terms led to collinearity, so we have dropped those terms and instated the same H1 as the other two indices.

well-being

V. Results

The treatment is the change in our soundscape metrics between winter and spring. We create a dataset of local soundscape metrics purely based on the local bird species presence (i.e. no other natural or anthropogenic sound is included). The three treatment variables we use are explained in more detail in section III, but we also summarise them here:

- **Acoustic Complexity Index**: This is a measure of the variation in acoustic intensity. Birdsong will tend to produce higher values of the ACI, whereas constant noise produces low values. We expect there to be a positive relationship between this variable and mental well-being.
- **Bioacoustic Index**: High values of this index occur when there is a large disparity in volume between the quietest and loudest parts of a recording. Therefore increases in this index tend to be associated with increases in bird vocalizations, or bird species richness. We expect there to be a positive relationship between this variable and mental well-being.
- **Acoustic Entropy**: High values can either mean extremely noisy soundscapes or very quiet soundscapes. Low values mean that noise is concentrated in a small band of frequencies. We expect that mental well-being will initially increase as acoustic entropy increases, as more birdsong is heard from more individuals, but this relationship will reverse at higher values of entropy. This is because some birdsong is preferred to complete silence.

In figure 2 we see the change in each soundscape variable between winter and spring. Each acoustic metric has a different pattern of change. The bioacoustic index shows increases across highland areas but decreases across more lowland areas. This likely picks up the migration of larger winter birds such as geese, which will winter in Britain, before heading north in spring, and have loud calls. The acoustic complexity index, in contrast, is more sensitive to birdsong, rather than all bird vocalizations, and so tends to show increases across all of Britain, and this is highest in highland areas. Acoustic entropy also tends to show increases but there is far less variation than in the other two indices. In fact, many areas show almost no change at all. This may show that entropy is less sensitive to change in bird species vocalization between winter and spring.

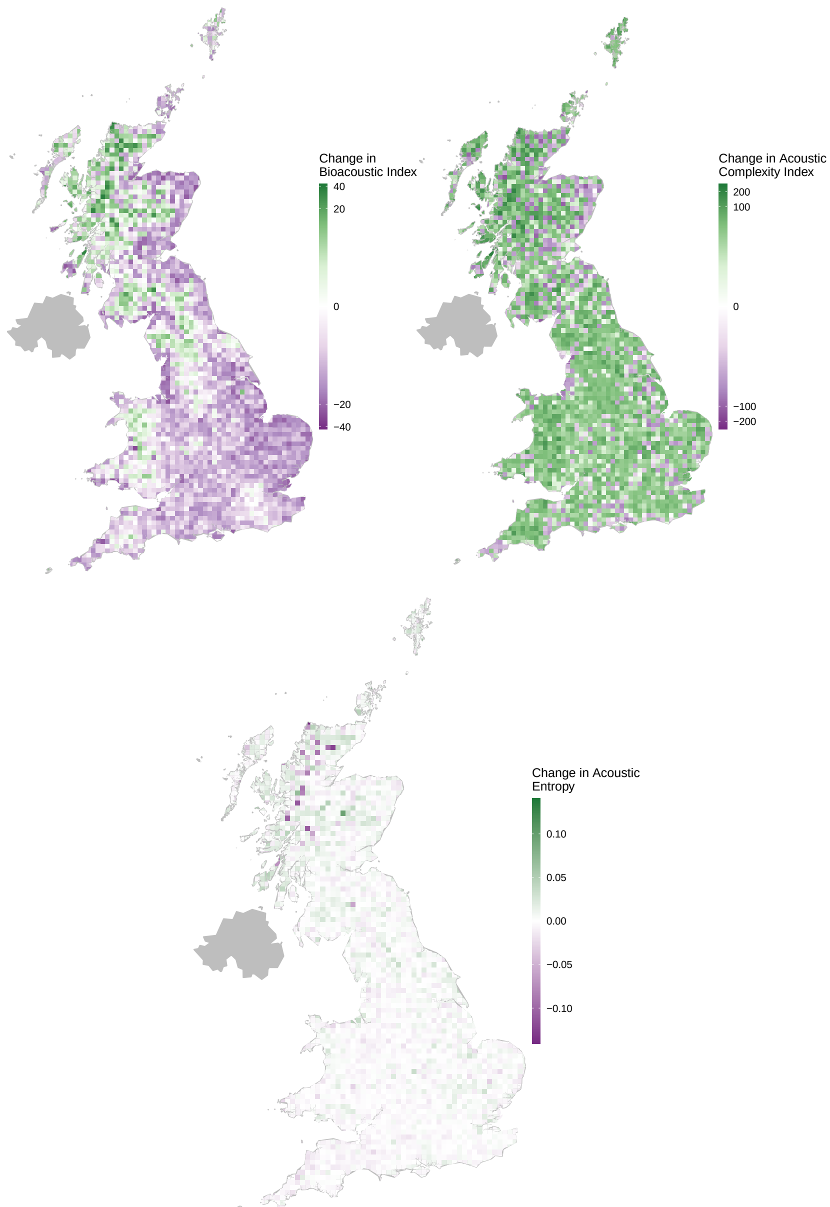

FIGURE 2. CHANGE BETWEEN WINTER AND SPRING IN ACOUSTIC INDICES

Note: Areas with no data are grey. Variables are pseudo-log transformed before plotting.

A. *Baseline Correlations*

Before our main, pre-registered results, we check if there is a simple correlation between our treatment and outcome variable. In figure 3, we include a heat map of our dependent variable, the change in WHO5 mental well-being score, against the change in each of the three acoustic variables. We can see for all three acoustic variables that there appears to be little to no relationship between their change and the change in mental well-being score. This is further confirmed in Appendix B, where we find no correlation between any of our metrics and mental well-being.

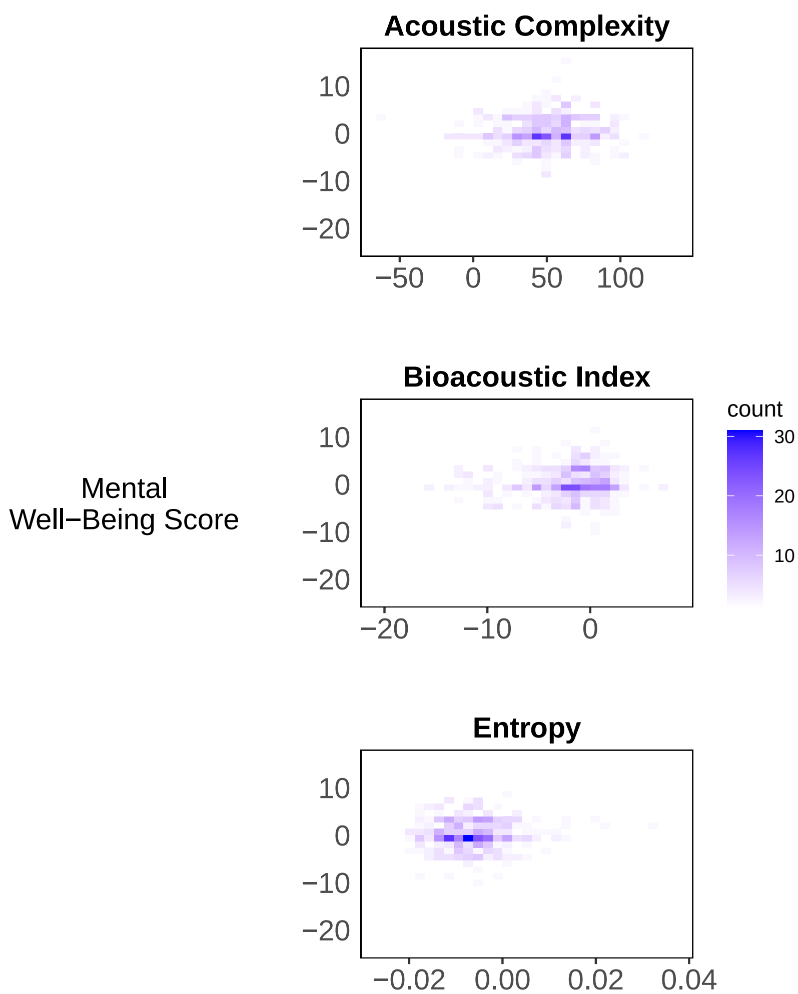

FIGURE 3. CHANGE BETWEEN WINTER AND SPRING IN ACOUSTIC INDICES AGAINST CHANGE IN MENTAL WELL-BEING SCORE

B. Main Results

Our main results use a difference-in-differences design estimated with the two step method in de Chaisemartin, D'Haultfœuille and Vazquez-Bare (2024).

Table 2 and figure 4 show the estimates from this process. We first regress without covariates and then with for each of the three treatment variables. Our $\hat{\theta}$ is an estimate of the Weighted Average of Switchers' Slopes (WAOSS) which is analogous to the Average Effect of Treatment on the Treated (ATT), but used in settings where there are no untreated units and treatment is continuous (de Chaisemartin et al., 2022).

In all cases, the $\hat{\theta}$ estimates are not statistically significant. The ACI and BIO estimates are extremely precise null results. Acoustic entropy (H) gives implausibly large point estimates with extremely wide intervals. We believe this is due to the lack of variation in H, but we include it as it was part of our pre-analysis plan. The p-value column shows p-values for a one-tailed test of $\hat{\theta} \geq 0$, as specified in our pre-analysis plan. As a sense check of the power of our estimates, we present a column showing the estimated minimal detectable effect (MDE) following Rainey (2024). For our one-tailed test this is calculated as $2.5 \times SE$. Here we confirm that we could detect very small changes for ACI and BIO, but not for H.

TABLE 2—EFFECT OF ACOUSTIC CHANGES ON WHO5 SCORE

Variable	$\hat{\theta}$	SE	Z-score	P-Value	MDE	N
ACI, No Covariates	0.00	(0.01)	-0.04	0.51	0.03	727
ACI, With Covariates	-0.03	(0.02)	-1.75	0.96	0.05	727
BIO, No Covariates	-0.06	(0.08)	-0.78	0.78	0.20	727
BIO, With Covariates	0.01	(0.10)	0.12	0.45	0.25	727
H No Covariates	10.64	(45.02)	0.24	0.41	113	727
H With Covariates	20.75	(59.57)	0.35	0.36	149	727

Notes: ACI = Acoustic Complexity Index.
Bio = Bioacoustic Index.
H = Entropy. SE = standard error. $\hat{\theta}$ is the WAOSS estimate (see section IV)
MDE is minimal detectable effect, calculated as 2.5*SE.
P-value shows one-tailed p-values.
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Overall, our results do not provide evidence that birdsong affects mental well-being, beyond the short-term effects other papers have found.

C. Parallel Trends Check

As stated in section IV, we cannot directly test our parallel trends assumption due to the two period design. However, we perform a check to see if units with similar levels of acoustic metric in winter, but different changes in acoustic metric, are a reasonable counterfactual for each other. We plot this in 5. We divided baseline winter soundscape exposure into deciles and, within each decile, classified changes in exposure from

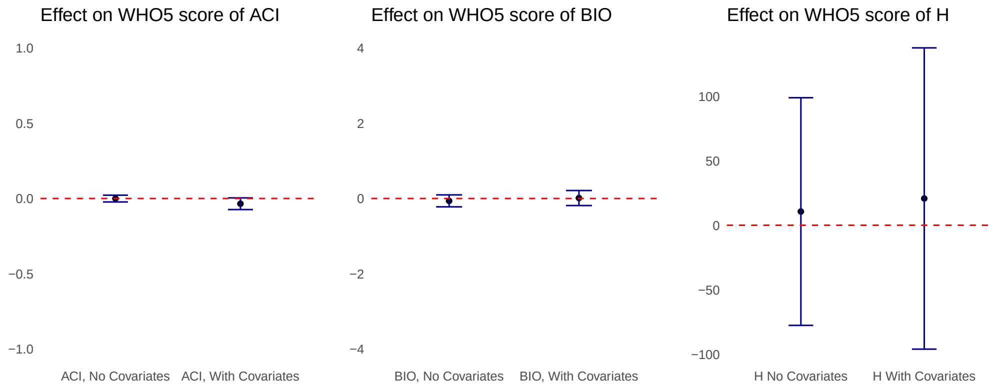

FIGURE 4. EFFECT OF ACOUSTIC CHANGES ON WHO5 SCORE

winter to spring into low, medium, and high tertiles. For each combination of baseline decile and change category, we calculated the mean change in self-reported well-being (WHO-5) and plotted these averages with 95% confidence intervals. This is a simple non-parametric check. If groups with similar levels of winter acoustic metrics but different spring metrics had large changes differences in how wellbeing changed, yet our main results show no treatment effect, it could be because of violations in the parallel trends assumption, but we do not see this. Across all three indices the group means overlap, confidence intervals are wide, and no consistent pattern by change category emerges, indicating little evidence that differences in seasonal shifts in soundscapes are linked to changes in well-being. This, combined with no correlation between treatment and mental well-being, indicates the results are not due to a failure in the assumptions we use to test for a causal effect, but simply due to almost no overlapping variation in treatment and outcome.

d1deciles1.pdf

FIGURE 5. WINTER ACOUSTIC LEVEL BY DECILE, CHANGE IN WHO5 SCORE, AND HIGH, MEDIUM OR LOW ACOUSTIC CHANGE

D. Robustness Checks

In the appendix, we carry out a number of robustness results. These include additional estimates of the WAOSS including covariates we did not pre-specify, such as the northing and easting of where our survey respondents live. We estimate with more fine-grained soundscape data, which reduces our sample, but may be more representative of the actual local soundscape for each individual. We also subset to a rural only sample. Some people do not have consistent location data and were excluded from our main results., We call these "movers" and we look at only people who moved between wave 1 and wave 2, to see if the movement to a higher acoustic metric area has any effect. In all cases, our results are similar to our main results.

VI. Discussion

Evidence from environmental psychology suggests green space offers valuable ecosystem services for mental well-being, but less is known about which aspects of green space are most important, or the causal pathways involved. Previous work has shown exposure to birdsong can have positive effects on short-term mental well-being. In order to investigate the effects of one aspect of greenspace, the soundscape, on seasonal mental well-being outcomes, we used a panel survey of 727 respondents in Great Britain who self-report an index of mental health, combined with a new, unique data set of local soundscapes, to estimate these effects for a sample of respondents in Great Britain. We used this data in both a simple correlational analysis and a difference-in-differences design, according to a pre-registered analysis plan, to estimate a casual effect of natural soundscapes on mental well-being.

We did not find evidence of an effect for any of the three acoustic indices we use on self-reported mental well-being, whether we look at correlations or at our plausibly causal estimates with a difference-in-difference design. Our results suggest that, while previous research indicates that short-term exposure to birdsong positively affects short-term mental well-being, it may not affect it in the medium term. Note that our study only investigates one facet of natural soundscapes. There are many ways birdsong and natural soundscapes can be valued by people beyond the ecosystem service effects on mental well-being.

Our results have implications for evidence-based environmental health policy and urban design strategies aimed at using ecosystem services to promote population mental well-being. Other attributes of green spaces may be more important than the soundscape for mental well-being, or attributes must be present in combinations. We suggest future research continues to examine the different attributes of green space to see which are most beneficial, and moves beyond short-term exposure, short-term outcome settings.

REFERENCES

Abrahams, Carlos, Bob Ashington, Ed Baker, Tom Bradfer-Lawrence, Ella Browning, Jonathan Carruthers-Jones, Jennifer Darby, Jan Dick, Alice Eldridge, David Elliott, Becky Heath, Paul Howden-Leach, Alison Johnston, Alexander Lees, Christoph Meyer, Usue Ruiz Arana, Siobhan Smyth, and Oliver Metcalf. 2023. *Good practice guidelines for long-term ecoacoustic monitoring in the UK.*

Aletta, Francesco, Tin Oberman, and Jian Kang. 2018. "Associations between Positive Health-Related Effects and Soundscapes Perceptual Constructs: A Systematic Review." *International Journal of Environmental Research and Public Health*, 15(11): 2392. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

Balmer, D. E., S. Gillings, B. J. Caffrey, R. L. Swann, I. S. Downie, and R. J. Fuller. 2013. *Bird Atlas 2007–11: The Breeding and Wintering Birds of Britain and Ireland.* Thetford, UK:British Trust for Ornithology.

Benfield, Jacob A., Gretchen A. Nurse, Robert Jakubowski, Adam W. Gibson, B. Derrick Taff, Peter Newman, and Paul A. Bell. 2014. "Testing Noise in the Field: A Brief Measure of Individual Noise Sensitivity." *Environment and Behavior*, 46(3): 353–372.

Beute, F., M. R. Marselle, A. Olszewska-Guizzo, M. B. Andreucci, A. Lammel, Z. G. Davies, J. Glanville, H. Keune, L. O'Brien, R. Remmen, A. Russo, and S. de Vries. 2023. "How do different types and characteristics of green space impact mental health? A scoping review." *People and Nature*, 5(6): 1839–1876. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pan3.10529.

Blanchflower, David G, Alex Bryson, Anthony Lepinteur, and Alan Piper. 2024. "Further Evidence on the Global Decline in the Mental Health of the Young." National Bureau of Economic Research Working Paper 32500.

Boelman, Natalie T., Gregory P. Asner, Patrick J. Hart, and Robert E. Martin. 2007. "Multi-trophic invasion resistance in Hawaii: bioacoustics, field surveys, and airborne remote sensing." *Ecological Applications*, 17(8): 2137–2144.

Burn-Murdoch, John. 2023. "Smartphones and social media are destroying children's mental health." *Financial Times*.

Buxton, Rachel T., Amber L. Pearson, Claudia Allou, Kurt Fristrup, and George Wittemyer. 2021. "A synthesis of health benefits of natural sounds and their distribution in national parks." *Proceedings of the National Academy of Sciences*, 118(14): e2013097118.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024.

"Difference-in-differences with a Continuous Treatment."

Costanza, Robert, Ralph d'Arge, Rudolf de Groot, Stephen Farber, Monica Grasso, Bruce Hannon, Karin Limburg, Shahid Naeem, Robert V O'Neill, Jose Paruelo, Robert G Raskin, Paul Sutton, and Marjan van den Belt. 1998. "The value of ecosystem services: putting the issues in perspective." *Ecological Economics*, 25(1): 67–72.

de Chaisemartin, Clément, Xavier D'Haultfœuille, and Gonzalo Vazquez-Bare. 2024. "Difference-in-Differences Estimators with Continuous Treatments and No Stayers."

de Chaisemartin, Clément, Xavier D'Haultfœuille, François Pasquier, and Gonzalo Vazquez-Bare. 2022. "Difference-in-Differences for Continuous Treatments and Instruments with Stayers." SSRN Working Paper. Accessed: 2025-01-21.

Frumkin, Howard, Gregory N. Bratman, Sara J. Breslow, William C. Sullivan B. Lynn, Peter H. Williams, Jon F. Ogren, Danielle A. Daily, Kevin J. Green, Susan L. Lawler, Paul H. Levin, Joshua J. Tandon, Julia A. Wolf, and Thomas A. Woodcock. 2017. "Nature Contact and Human Health: A Research Agenda." *Environmental Health Perspectives*, 125(7): 075001.

Hansen, Ben B., and Jake Bowers. 2008. "Covariate Balance in Simple, Stratified and Clustered Comparative Studies." *Statistical Science*, 23(2): 219–236. Publisher: Institute of Mathematical Statistics.

Hartig, Terry, Richard Mitchell, Sjerp de Vries, and Howard Frumkin. 2014. "Nature and Health." *Annual Review of Public Health*, 35: 207–228.

Higney, Anthony, Nicholas Hanley, Claire Buchan, Simon Butler, Eleanor Ratcliffe, Melissa Marselle, and Konrad Uebel. 2024. "The Effect of Birdsong on Self-Reported Mental Health: Pre-Analysis Plan." Open Science Framework (OSF), CC-By Attribution 4.0 International, accessed January 21, 2025.

Kaplan, Rachel, and Stephen Kaplan. 1989. *The experience of nature: a psychological perspective.* Cambridge:Cambridge University Press. OCLC: 18814094.

Krutilla, John V. 1967. "Conservation reconsidered." *The American Economic Review*, 57(4): 777–786.

Markevych, Iana, Julia Schoierer, Terry Hartig, Alexandra Chudnovsky, Perry Hystad, Angel M. Dzhambov, Sjerp de Vries, Margarita Triguero-Mas, Michael Brauer, Mark J. Nieuwenhuijsen, Gerd Lupp, Elizabeth A. Richardson, Thomas Astell-Burt, Donka Dimitrova, Xiaoqi Feng, Maya Sadeh, Marie Standl, Joachim Heinrich, and Elaine Fuertes. 2017. "Exploring pathways linking greenspace to health: Theoretical and methodological guidance." *Environmental Research*, 158: 301–317.

Marselle, Melissa R., Jutta Stadler, Horst Korn, Katherine N. Irvine, and Aletta Bonn, ed. 2019. *Biodiversity and Health in the Face of Climate Change*. Springer Nature. Accepted: 2020-03-18 13:36:15.

Morrison, C. A., A. Auniņš, Z. Benkő, L. Brotons, T. Chodkiewicz, P. Chylarecki, V. Escandell, D. P. Eskildsen, A. Gamero, S. Herrando, F. Jiguet, J. A. Kålås, J. Kamp, A. Klvaňová, P. Kmecl, A. Lehikoinen, Å. Lindström, C. Moshøj, D. G. Noble, I. J. Øien, J.-Y. Paquet, J. Reif, T. Sattler, B. S. Seaman, N. Teufelbauer, S. Trautmann, C. A. M. van Turnhout, P. Voříšek, and S. J. Butler. 2021. "Bird population declines and species turnover are changing the acoustic properties of spring soundscapes." *Nature Communications*, 12(1): 6217.

Pieretti, Nadia, Almo Farina, and Davide Morri. 2011. "A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI)." *Ecological Indicators*, 11(3): 868–873.

Rainey, Carlisle. 2024. "Power Rules: Constructing Theories and Models with More Useful Predictions." Accessed: 2025-01-21.

Ratcliffe, Eleanor. 2021. "Sound and Soundscape in Restorative Natural Environments: A Narrative Literature Review." *Frontiers in Psychology*, 12. Publisher: Frontiers.

Ratcliffe, Eleanor, Birgitta Gatersleben, and Paul T. Sowden. 2016. "Associations with bird sounds: How do they relate to perceived restorative potential?" *Journal of Environmental Psychology*, 47: 136–144.

Sueur, Jérôme, Thierry Aubin, and Christophe Simonis. 2008. "Seewave, a free modular tool for sound analysis and synthesis." *Bioacoustics*, 18(2): 213–226.

Topp, Christian Winther, Søren Dinesen Østergaard, Susan Søndergaard, and Per Bech. 2015. "The WHO-5 Well-Being Index: A Systematic Review of the Literature." *Psychotherapy and Psychosomatics*, 84(3): 167–176.

Turner, R. Kerry, and Gretchen C. Daily. 2008. "The Ecosystem Services Framework and Natural Capital Conservation." *Environmental and Resource Economics*, 39(1): 25–35.

Twohig-Bennett, Caoimhe, and Andy Jones. 2018. "The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes." *Environmental Research*, 166: 628–637.

Uebel, Konrad, Eleanor Ratcliffe, Claire Buchan, Simon J. Butler, Nicholas Hanley, Anthony Higney, and Melissa Marselle. 2025*a*. "Natural soundscapes are associated with mental well-being via capacity-building and capacity-restoring pathways." *OSF Preprints*. Preprint.

Uebel, Konrad, Melissa Marselle, Angela J. Dean, Jonathan R. Rhodes, and Aletta Bonn. 2021. "Urban green space soundscapes and their perceived restorativeness." *People and Nature*, 3(3): 756–769. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pan3.10215.

Uebel, Konrad, Melissa Marselle, Claire Buchan, Simon J. Butler, Nicholas Hanley, Anthony Higney, and Eleanor Ratcliffe. 2025*b*. "The influence of acoustic characteristics and anthropogenic noise on restorative perceptions of natural soundscapes." *OSF Preprints*. Preprint.

Ulrich, Roger S. 1983. "Aesthetic and Affective Response to Natural Environment." In *Behavior and the Natural Environment*. , ed. Irwin Altman and Joachim F. Wohlwill, 85–125. Boston, MA:Springer US.