



Counting butterflies—are old-fashioned ways of recording data obsolete?

Elisabeth Kühn^{1,2} · Alexander Harpke^{1,3} · Thomas Schmitt^{2,4} · Josef Settele^{1,5,6} · Ingolf Kühn^{3,5,6}

Received: 1 November 2023 / Accepted: 24 March 2024 / Published online: 22 April 2024
© The Author(s) 2024

Abstract

Citizen Science projects aim to make data entry as easy as possible and often provide online data recording or data recording with an App. However, many participants cannot or do not want to use these possibilities and record their data the “old-fashioned” way with pen on paper. We ask whether the quality of data recorded in the “old-fashioned” way (transect walkers record their data with pen on paper) is of the same, better or worse quality than data recorded “online” (transect walkers enter their data via an online tool). We use the project “Butterfly Monitoring Germany” as an example, where we identify three different types of volunteers: those who enter their data online, those who send their data to the project coordination via email in different formats and those who send their data to the project coordination via ordinary mail. We observed minor quantitative differences for transect walkers not entering their data online but significant qualitative differences. Transect walkers who send their data via email record significantly more data for some rare or difficult to determine species. This is essential to properly calculate these species’ trends. In addition, the results of a questionnaire showed that “old fashioned” transect walkers did not use the online data entry because (i) data entry takes too long, (ii) is too cumbersome, (iii) they have bad or no internet connection or (iv) lack of technical capabilities. Accounting for different preferences of Citizen Scientists, different ways of data-submission should be made available (e.g. online, via app, or the old-fashioned way on paper). For the future, projects that collect large amounts of Citizen Science data should further develop low-threshold input data pipelines. **Implications for insect conservation** Our results show that data recorded in the old-fashioned way contributes significantly to increasing data quality. It is therefore very important to continue to enable different forms of data recording in the future. Furthermore, it is crucial to keep in mind that Citizen Science projects are only partly for the sake of science, but also volunteers should benefit by being part of a community and having access to information about (butterfly) biodiversity.

Keywords Citizen science · Butterfly monitoring · Data entry tools · Loss of species knowledge

✉ Elisabeth Kühn
elisabeth.kuehn@ufz.de

- ¹ Department of Conservation Biology and Social-Ecological Systems, Helmholtz Centre for Environmental Research-UFZ, Theodor-Lieser-Straße 4, 06120 Halle (Saale), Germany
- ² Entomology and Biogeography, Institute of Biochemistry and Biology, Faculty of Science, University of Potsdam, Potsdam, Germany
- ³ Department of Community Ecology, Helmholtz Centre for Environmental Research-UFZ, Halle, Germany
- ⁴ Senckenberg German Entomological Institute Müncheberg, Müncheberg, Germany
- ⁵ German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany
- ⁶ Institute of Biology, Martin Luther University Halle-Wittenberg, Halle, Germany

Introduction

One of the most pressing questions of our time is how to stop the widely documented loss of biodiversity (Hallmann et al. 2017; Cowie et al. 2022). In order to tackle this problem effectively, it is important to get information on the occurrence and abundance of species, how both change over time, and which drivers are behind these changes. Only with such information, we can effectively protect species. In recent decades, Citizen Science projects, i.e. the collection of data by volunteers, have been proven as an excellent way for generating such occurrence and abundance data as a basis for further analyses (Schmeller et al. 2009; Chandler et al. 2017). This is particularly true for popular groups, especially if many of their common species are easy to identify, such as birds or butterflies.

Along with the growing popularity of Citizen Science (CS) projects, the technical possibilities for recording, processing and analyses of the resulting data are constantly increasing. While 15–20 years ago, data were mostly recorded in the traditional way by noting observations and records with paper and pencil, increasingly modern recording tools have emerged since then. They range from websites with corresponding data entry sheets, simple recording apps to advanced apps with integrated AI based image recognition, species identification and geolocation of the individuals. The progress in this field is fast and the resulting possibilities inspiring (Wäldchen et al. 2018; Høye et al. 2021; Mäder et al. 2021). Alongside, data management has also evolved. In recent decades, for example, data recording has evolved from index cards via Excel spreadsheets to powerful databases.

At the same time, however, there is a completely different development in the field of Citizen Science, which at the first glance is independent from the technological development of data recording. The older generation, i.e. numerous people with good knowledge of species, and some of them have even specialised in specific species groups, is so far not replaced by the next generation adequately. For example the median age of participants of “Butterfly Monitoring Germany” is 62 with only few young people participating (Kühn et al. 2022a). The ability to identify animal and plant species in the field or even to re-identify them with the help of identification books has declined sharply within the younger generation; claims also include that taxonomist are nearing extinction alongside the species they identify (Miller 2005; Hooykaas et al. 2019; Gerl et al. 2021; European Commission. Directorate General for Environment. 2022). This applies to professional as well as to amateur species experts, both have profound knowledge in special fields. Many of these “old-fashioned” species experts have not changed their data collection approaches for decades and hardly use modern technologies. This development goes along with the frequent opinion that “old-fashioned” data collection with pen and paper in the field is obsolete. The result is that it is increasingly difficult to make observations available to science when not collected by individuals using a mobile phone, laptop or computer.

It is often argued that it is important to increasingly and particularly attract young people to Citizen Science projects because presently many such projects mainly attract older people (Füchslin et al. 2019; Aristeidou et al. 2021; Kühn et al. 2022a). This might be due to several reasons. On the one hand, people at the beginning of their professional career and/or with small children have less time for volunteer activities, while retired people have more time available to participate in CS projects. At the same time, however, “old-fashioned” data collection is not as attractive to young people as data collection via an App on a mobile phone, where

it may also be possible to determine species automatically via image recognition and at the same time still having a lot of additional information at hand. It remains to be seen how successful these modern methods of data collection attract more young people, as well as, less experienced groups.

A CS project in which both types of data recorders (i.e. traditional and modern ones) are active is the Butterfly Monitoring in Germany (= Tagfalter-Monitoring Deutschland—TMD, www.tagfalter-monitoring.de). In this project, volunteer butterfly enthusiasts count butterflies several times a year and over many years along defined transect routes (Kühn et al. 2008). The project started in 2005 and participants have three options to submit their data. They can send their data by (i) ordinary mail on paper or (ii) by email via attachments in many formats (Excel files, pdf files, photos of completed data entry forms) to central coordination. These options are often used by “old fashioned” recorders who prefer to count butterflies in the field with paper and pen. Such data are entered manually into the online data entry tool at the headquarters of TMD. (iii) The most commonly chosen option is that recorders enter their data via the same online data entry tool (as used by central coordination) themselves (“online recorders”). The additional work of manually entering such data collected using the “old fashioned” method (sent by letter or email) is not only time-consuming and cost-intensive but also prone to errors, as transmission errors can easily occur. Consequently, the project coordinators discuss every year whether this form of data entry is worth the effort. The advantage, however, is that when “old fashioned” data is manually transferred to the central database, all data is available in a standardised format.

Therefore, we sent a questionnaire to all recorders not entering their data themselves via the online data entry tool, asking them why they are not using this method. We also asked what hurdles exist that makes the use of the online data collection problematic. In addition, we inquired what needed to be done (or implemented) to convince recorders, who use the “old fashioned” method, to enter their data online.

In this publication, we address the following questions related to this topic.

Is the quality of data collected by “old-fashioned” data recorders of the same, better or worse quality than data collected by “online recorders”? Specifically, we would like to know:

1. Do “old-fashioned” recorders generate data that we cannot get via online data entry?
2. Is “old-fashioned” data collection outdated and can we do without it in favour of online data collection on computers or mobile phones?

3. Can we derive recommendations for the development of online data collection tools that are used by as many participants in Citizen Science projects as possible?

Methods

Background

Butterfly Monitoring Germany is a Citizen Science project in which volunteers participate all over Germany. Whoever is interested in counting butterflies can join, and the site for a butterfly transect is freely chosen by the transect walkers. Most transect walkers decide to walk a transect near their homestead (Kühn et al. 2008). The methodology of how to count is highly standardised and counts are performed each year from April to September, ideally once a week (Kühn et al. 2014). Transect walkers enter their results for each transect walk on recording sheets, noting the butterfly species they saw and the number of individuals per species (Fig. 1). Later, these results can be entered online via a data entry tool for the project. Alternatively, the recording sheets can be sent to the project's central coordination via email or letter and the data is entered by project staff into a central database. The data is then available in a uniformly standardised format.

For this paper, we used data from the German Butterfly Monitoring from the year 2021. In this year, 593 transects were walked, 367 transect walkers were active with many counting butterflies on more than one transect. 151 different butterfly and burnet species and 77 (mostly day active) moth species were counted. We used Reinhardt and Bolz (2011) to identify red listed species, considering all species that are listed in categories 1–3 (i.e. threatened with extinction, highly endangered and endangered), the others are not endangered.

Comparison of data collected by different groups of data entry

To answer the above-mentioned questions, we compare the data of “old-fashioned recorders”, who do not enter their data online themselves, with those of the “online recorders” who enter their data themselves via our online data entry tool. The “old-fashioned recorders” are divided into those who send us their data via email and those who send us their data via letter. We assume that the participants sending their data by email would theoretically also have the possibility to enter data online themselves. The participants sending their data by letter will most likely not be able to enter data online themselves in the future.

The groups are abbreviated as follows: group “online” = Data entry online, group “email” = Data submission via email, group “letter” = Data submission via letter.

Questionnaire “old fashioned” data entry

In the context of this project, a survey was developed consisting of three multiple-choice questions (full questionnaire as supplementary material S4 [English] and S5 [German]). Here, we analyse the answers on question 2 and 3 as these are relevant for the content of this paper. A total of 93 questionnaires were sent to all transect walkers sending their data for the year 2021 via letter or email (but were not entering their data themselves). They were asked to answer the questions and to send us the questionnaire back, using the stamped envelope included.

Statistical analyses

All statistical analyses were performed in R (R Core Team 2022) v. 4.2.2 with RStudio (RStudio Team 2023) v. 2023.09.1 +494 as user interface.

We compared recorder activity, the individuals and species richness of all butterflies and moths (jointly) and those on the Red List (Reinhardt and Bolz 2011; butterflies only), respectively, using Analysis of Variance based on glm with quasipoisson distribution to account for the, on cases, substantial overdispersion. In case of significant differences (at $\alpha = 0.05$), we performed a Tukey's all-pair comparisons post-hoc test using the function `glht()` in the package `multcomp` (Hothorn et al. 2008; Bretz et al. 2010).

We identified associations between species and recorder groups using the R package `indicspecies` v. 1.7.13 (Cáceres and Legendre 2009), using a multilevel pattern analysis (function `multipatt()`). Here, three metrics are reported, A, B, and stat. A is the mean abundance of the species in the target site group divided by the sum of the mean abundance values over all groups. B is the relative frequency of occurrence of the species inside the target site group. Stat refers to the statistics of association (a special form of correlation coefficient, for details see Cáceres and Legendre (2009)). Proportions of transects with these indicator species present were then compared in a mixed-effects ANOVA using `aov()` and species identity as random factor. These two analyses were restricted to butterflies only.

All test statistics are provided as supplementary material in supplement S1 and S2.

Erfassungsbogen Tagfalter-Monitoring Deutschland 2022/09

Transekt: ST-4137-01 Lödderitz	Bemerkungen: BLÜTE!
Datum: 15.05.2022	ROT-WEISS-HORN- WUNDKLEE.
Zähler: Dr. Jürgen Ziegeler	ROTE NACHTNELKE.
Uhrzeit von: 10.00 UHR	WIEGEN-WUCHERBLUME.
Uhrzeit bis: 13.00 UHR	GEMEINER BEIKWELL.
Bewölkung (%): FAST WOLKENLOS	GEM. HAHNENFUSS.
Wind: OST BIS NORDOST DER STÄRKE 1	WIESEN-HERBEL.
Temperatur: 22°C.	

Art/Abschnitt	1	2	3	4	5	
Zitronenfalter (<i>Gonepteryx rhamni</i>)						
Grober Kohlwelbling (<i>Pieris brassicae</i>)				1		01
Kl. Kohl-/Grünader-Welbling (<i>P. rapae/napl</i>)	IIII	IIII	III	IIII	IIII	31
Kleiner Kohl-Welbling (<i>Pieris rapae</i>)						
Grünader-Welbling (<i>Pieris napl</i>)						
Aurorafalter (<i>Anthocharis cardamines</i>)	1♂ 1♀	1♂	1♂ 1♀	1♂ 1♀		10
Kleiner Feuerfalter (<i>Lycaena phlaeas</i>)	1 (BELEGFOTO)				1	02
Kauleckel-Bläuling (<i>Polyommatus leucus</i>)	1♂			III♂	III♂	07
Kaisermantel (<i>Argynnis paphia</i>)						
Kl. Pflüchterfalter (<i>Issoria lathonia</i>)						
Admetal (<i>Vonessa atalanta</i>)						
Distelfalter (<i>Vonessa cardui</i>)						
Tagpfauenauge (<i>Aglais io</i>)					1	01
Kleiner Fuchs (<i>Nymphalis urticae</i>)						
C-Falter (<i>Nymphalis e-album</i>)			1			01
Landkärtchen (<i>Araschnia levana</i>)			II			02
Waldbrettspiel (<i>Pararge aegeria</i>)					1	01
Kl. Wiesenvägelchen (<i>Coenonympha pamphilus</i>)	IIII					07
KL. SONNENRÖSCHENBLÄULING: 1					1	02
						TAGFALTERSUMME: 64
						ARTENSUMME: 11



**AURORAFALTER
MÄNNCHEN
(ANTHOCHARIS
CARDAMINES)**

Fig. 1 Recording sheet from an “old-fashioned” transect walker, filled in with a pen on paper and sent to the coordination of Butterfly Monitoring Germany via letter

Results

In 2021, butterflies were counted on 593 transects. The vast majority of these transect data were entered online; little more than 15% of the data had to be entered manually by the project coordination (Fig. 2).

To measure the effort invested in butterfly counting by transect walker groups, we first compared the number of transect surveys conducted per person per year in each group (Fig. 3). There were significant differences ($p=0.007$) among the different groups of recorders. The Posthoc-Test yielded significantly more walks for the group “email” compared to “online” and marginally more walks than in group “letter”. The latter two did not differ significantly (please see supplementary material S1 and S2 for a full overview of all relevant test statistics).

As a next step, we compared the number of butterfly species recorded in each group in 2021 (Fig. 4). No significant differences ($p=0.19$) in numbers of butterfly species were recorded per transect among groups. Further, we compared the number of butterfly individuals recorded per year in each group in 2021 (Fig. 5), but no significant differences were obtained either ($p=0.76$).

To assess differences in the value and ecological importance of the data, we compared the number of red-listed butterfly species (threat status 1–3) recorded per year in each group (Fig. 6). We could only find marginally significant differences among groups ($p=0.051$), but the Tukey posthoc test did not show any significant pairwise differences. Further, we compared the number of individuals of red-listed butterfly species recorded per year in each group

Fig. 2 Type of data entry in the data repository of Butterfly Monitoring Germany for the recording year 2021

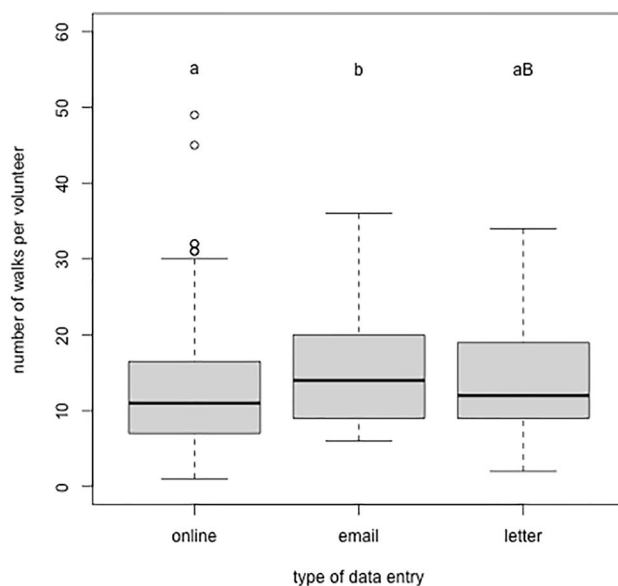
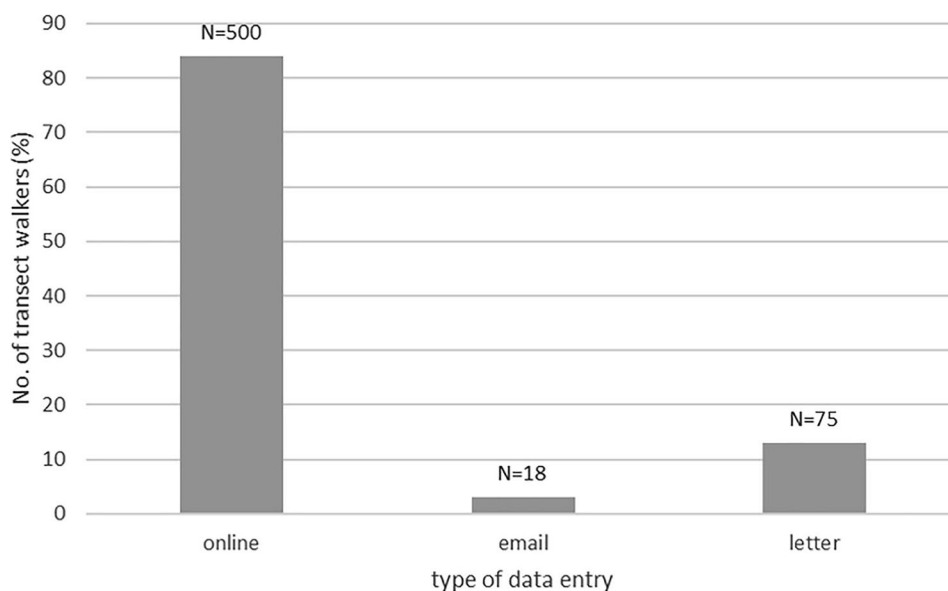


Fig. 3 Activity of volunteers (No. of walks per volunteer) for the 2021 monitoring season. Significantly different groups ($p<0.05$) are marked by different letters, marginally significantly different groups (here $p=0.08$) are indicated by different capitalisations of the same letter

(Fig. 7). The overall test yielded no significant differences ($p=0.41$).

Restricting the analyses to those transects with the presence of red-listed species (to avoid zero-inflation) resulted in significant differences ($p=0.005$), with group “email” data delivery recording significantly more species than the

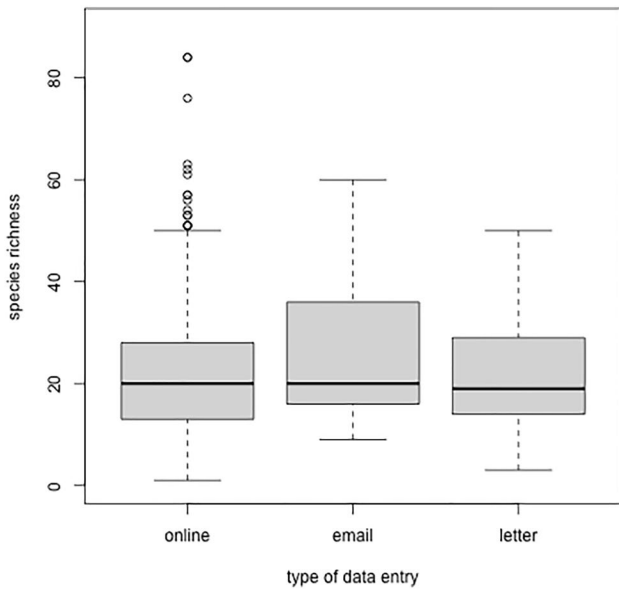


Fig. 4 Number of butterfly species recorded per groups of data entry per transect in 2021

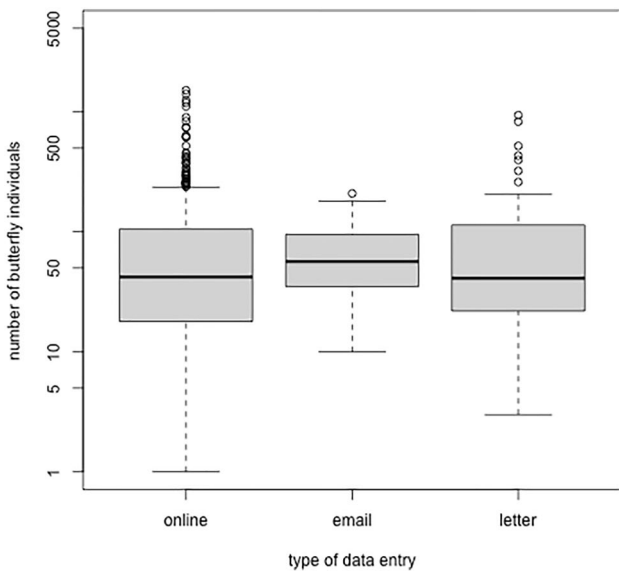


Fig. 5 Number of butterfly individuals per transect recorded per group of data entry in 2021 (note the log scale on the y axis!)

two other groups, while group “online” records marginally more species than group “email” (Fig. 8). The differences are even more pronounced for abundances ($p < 0.001$) (Fig. 9), but here the records submitted by group “letter” are significantly lower than the two others, while the number of online records was just marginally lower than those submitted via email.

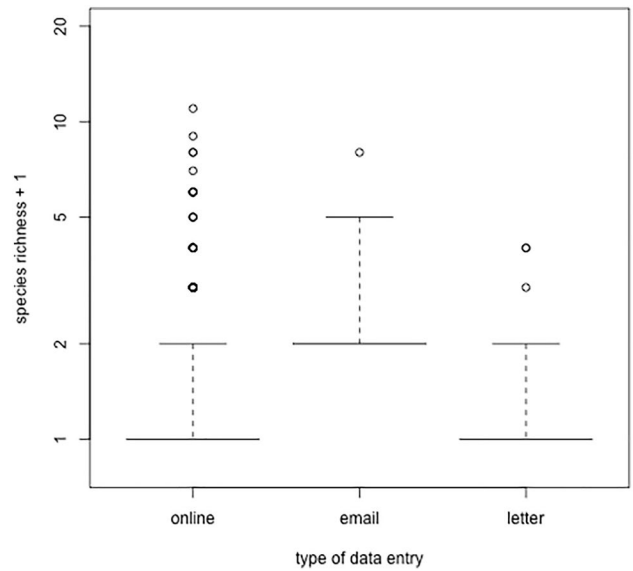


Fig. 6 Number of red list species recorded per transect recorded per groups of data entry (note the log scale on the y axis; therefore 1 needed to be added to species richness!). Number of individuals of all nationally red-listed butterfly species per transect for the entire year 2021 recorded per groups of data entry in 2021 (note the log scale on the y axis; therefore 1 needed to be added to species richness!).

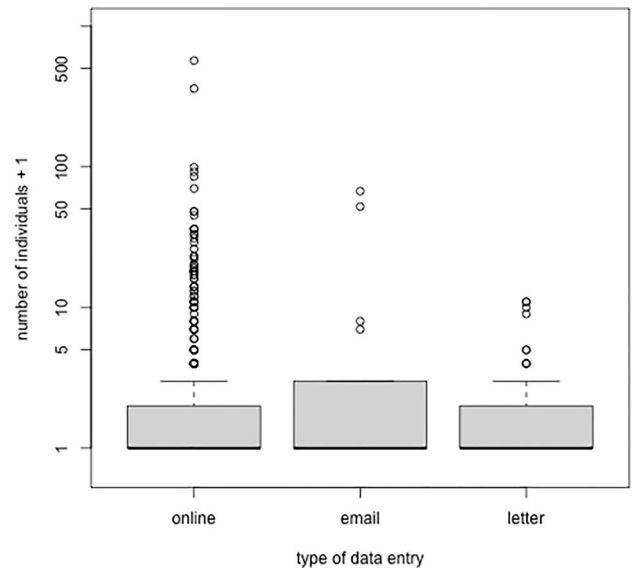


Fig. 7 Number of individuals of all nationally red-listed butterfly species per transect for the entire year 2021 recorded per groups of data entry in 2021 (note the log scale on the y axis; therefore 1 needed to be added to species richness!).

Indicator species analysis

To assess differences in the value and ecological importance of the data on species level, we looked for butterfly species

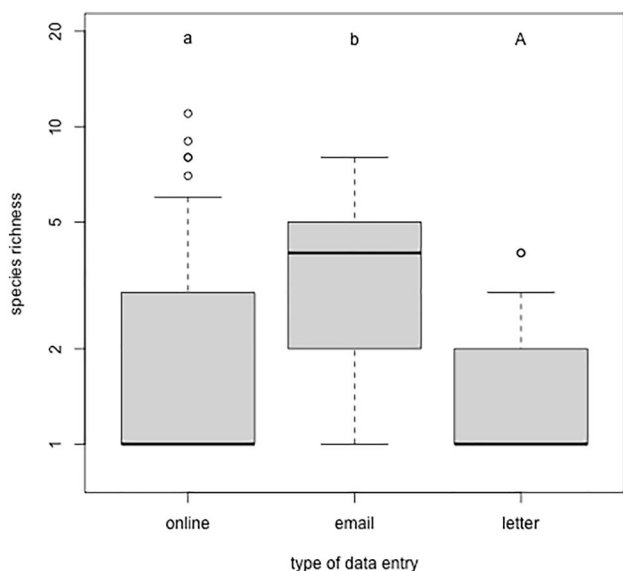


Fig. 8 Number of red-listed butterfly species per transect recorded per groups of data entry in 2021, restricted to those transects with red-listed species occurrences (note the log scale on the y axis). Significantly different groups ($p < 0.05$) are marked by different letters, marginal significance (here $p = 0.09$) is marked by capitalization of the same letter

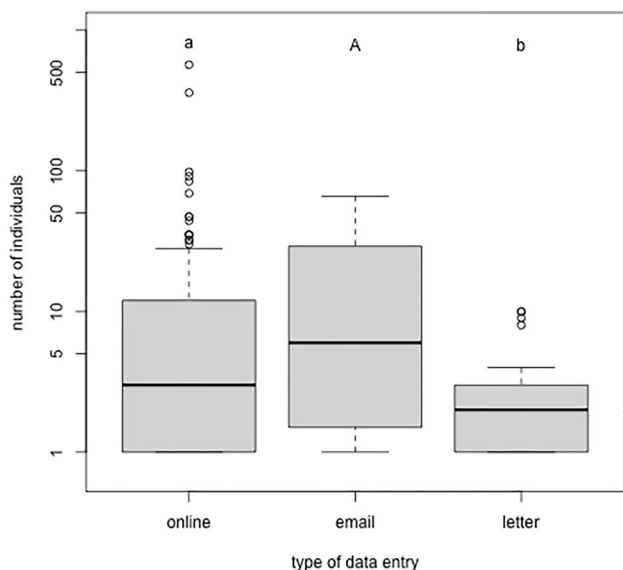


Fig. 9 Number of individuals of red-listed butterfly species per transect recorded per groups of data entry in 2021, restricted to those transects with red-listed species occurrences (note the log scale on the y axis). Significantly different groups ($p < 0.05$) are marked by different letters, marginal significance (here $p = 0.09$) is marked by capitalization of the same letter

that “indicate” one group of data entry, i.e. are significantly more frequent/abundant in one group than in any other. 25 species were associated with a special group of data entry.

The species complex *Thymelicus lineola/sylvestris* is associated to the group “online” and a total of 24 species is associated to the group “email” (Table 1). If we look at these “indicator species” in detail (see also Supplementary material S3), 14 of these are red-listed species (seven vulnerable, two endangered and five critically endangered). The frequency of “indicator species” differed significantly among groups ($p < 0.0001$) and was higher in transects of the group “email” compared to the other two (Fig. 10).

Table 1 List of species significantly associated to one of the two groups “online” and “email. A is the mean abundance of the species in the target site group divided by the sum of the mean abundance values over all groups. B is the relative frequency of occurrence of the species inside the target site group. Stat refers to the statistics of association (a special form of correlation coefficient, for details see Cáceres and Legendre (2009))

Species	A	B	stat	p-value	Red List status
Group “online”					
<i>Thymelicus lineola/sylvestris</i>	0.947	0.184	0.418	0.046 *	*
Group “email”					
<i>Araschnia levana</i>	0.626	0.389	0.493	0.039 *	*
<i>Brenthis ino</i>	0.917	0.167	0.391	0.013 **	*
<i>Brintesia circe</i>	0.86	0.111	0.309	0.013 *	3
<i>Callophrys rubi</i>	0.69	0.167	0.339	0.026 *	V
<i>Coenonympha glycerion</i>	0.812	0.111	0.307	0.027 *	V
<i>Coenonympha tullia</i>	0.992	0.056	0.235	0.031 *	2
<i>Colias alfariensis</i>	0.723	0.167	0.347	0.004 **	*
<i>Colias crocea</i>	0.748	0.111	0.288	0.029 *	*
<i>Gonepteryx rhamni</i>	0.56	0.833	0.683	0.025 *	*
<i>Limnitis populi</i>	0.949	0.111	0.325	0.007 **	2
<i>Lycaena phlaeas</i>	0.513	0.722	0.609	0.027 *	*
<i>Lycaena virgaureae</i>	0.851	0.111	0.307	0.027 *	V
<i>Minois dryas</i>	0.94	0.167	0.396	0.001 ***	2
<i>Phengaris alcon</i>	1.0	0.056	0.236	0.034 *	2
<i>Phengaris nausithous</i>	0.731	0.167	0.349	0.011 *	V
<i>Phengaris teleius</i>	0.965	0.056	0.232	0.032 *	2
<i>Plebejus argyrognomon</i>	0.86	0.111	0.309	0.004 **	*
<i>Polyommatus amandus</i>	0.872	0.111	0.311	0.01 **	*
<i>Pyrgus malvae</i>	0.661	0.222	0.383	0.019 *	V
<i>Satyrium spini</i>	0.985	0.111	0.331	0.007 **	3
<i>Speyeria aglaja</i>	0.819	0.222	0.427	0.001 **	V
<i>Zygaena ephialtes</i>	0.78	0.111	0.294	0.022 *	*
<i>Zygaena lonicerae</i>	0.96	0.056	0.231	0.017 *	V
<i>Zygaena viciae</i>	0.858	0.111	0.309	0.036 *	*

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’ 1
 Red List status (Reinhardt and Bolz 2011): * = least concern, V = vulnerable, 3 = endangered, 2 = critically endangered

Questionnaire

Of 93 questionnaires sent to the transect walkers of the groups “email” and “letter”, 50 were returned (i.e. a response rate of 54%). The main reasons transect walkers did not use the online data entry were that data entry takes too long and is too cumbersome, but also technical reasons such as bad or no internet connection and lack of technical capabilities were stated (Fig. 11). Almost half of transect walkers would not use the online data entry under any circumstances, but

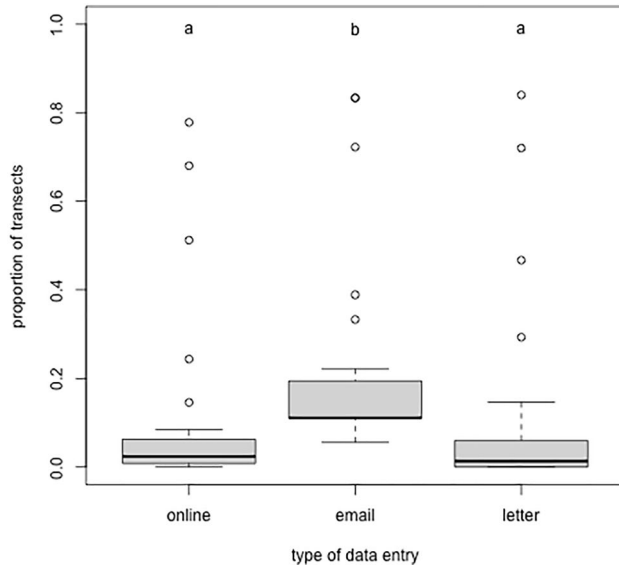


Fig. 10 Proportion of indicator species per group of data entry in 2021

about a quarter would use the online data entry if it was faster, and a fifth if it was easier (Fig. 12).

Discussion

Transect walkers not entering their data online invested on average more time in the field, and performed more transect walks. However, this did not result in a higher number of butterfly species or individuals recorded (also not of those on the Red List). Restricting the analyses to transects with the presence of red-listed species, transect walkers submitting their data via email obtained significantly higher proportions of individuals and numbers of butterfly species included in the Red Lists than the other groups of transect walkers. Most importantly, 24 species were significantly more frequent on transects of recorders submitting their data by email.

In general, data collected by “old fashioned” recorders were of comparable but not better quantity than those collected by “online recorders”. Taking species identity (on taxon level) into account revealed a qualitative difference. Thus, the Alcon Blue *Phengaris alcon* (critically endangered in Germany) was exclusively counted on one transect by a transect walker submitting data via email; the Large Heath *Coenonympha tullia* (critically endangered) was only counted on two transects, one by a transect walker entering data online and one by a transect walker submitting data via email. Data on endangered species (i.e. Red List categories 1–3) are generally underrepresented in “Butterfly Monitoring Germany” because recording these species usually requires more effort: Transect walkers have to visit specific habitats often not in the immediate vicinity of their place

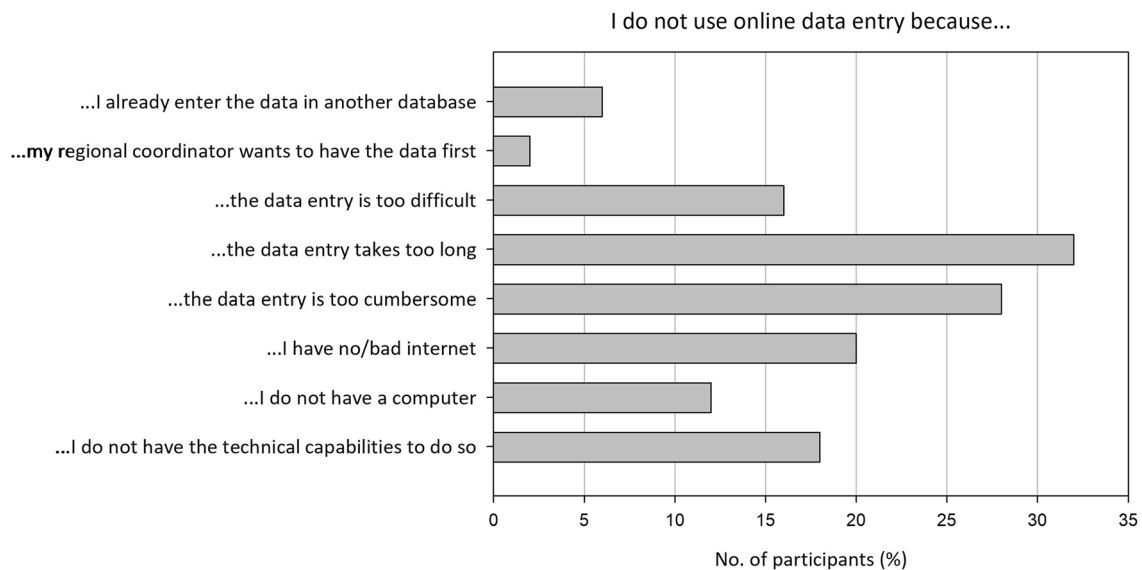


Fig. 11 Answers to a questionnaire sent out to the project participants returning their record via email or letter in 2021

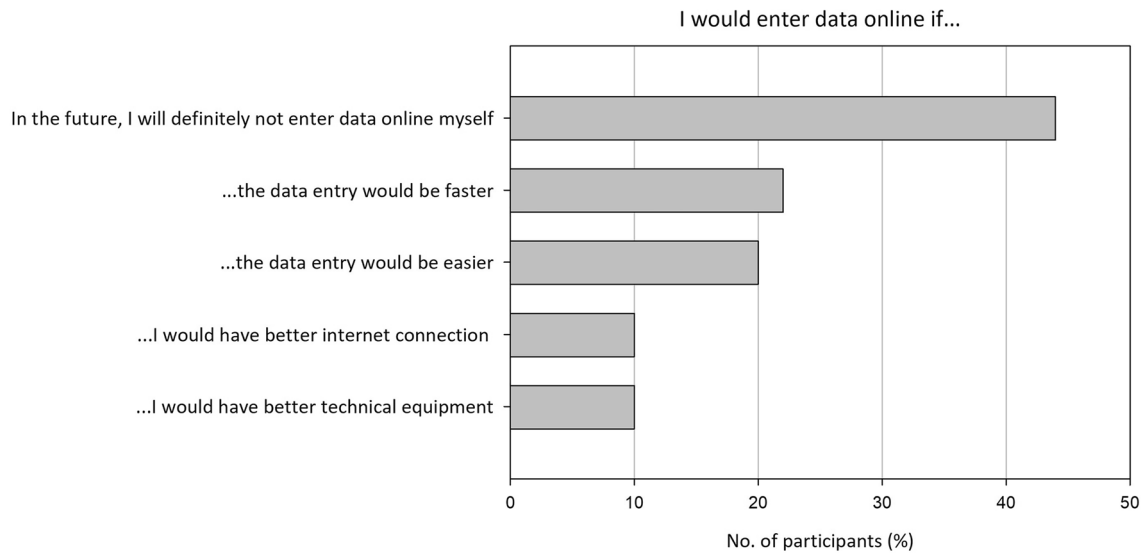


Fig. 12 Answers to a questionnaire sent out to the project participants returning their record via email or letter in 2021

of residence. The list of species more frequently recorded by transect walkers of the “Email” group showed that rare and endangered habitats were visited by them for numerous species, such as bogs (with *Coenonympha tullia* and *Minois dryas*) or wet meadows (with *Phengaris alcon* and *Ph. teleius*). In addition, the list reported by the “old fashioned” recorders also included species difficult to distinguish from closely related taxa (e.g. *Colias alfacariensis*, *Plebejus argyrognomon*, *Zygaena viciae*). For these, in depth taxonomic knowledge is necessary to address these species correctly. In summary, many of the transect walkers submitting their data via email are volunteers with very good, and in some cases specialised, species knowledge specifically selecting their transects in species-rich habitats.

One main output of “Butterfly Monitoring Germany” are population trends of butterfly species, which are analysed and reported every year (Kühn et al. 2022b). For analysing these trends, having sufficient data of the respective species is crucial. Especially for rare species (which in most cases are also endangered), each individual data set is of importance for trend calculations. Many studies focus on widespread species because data for rarer species are missing or not sufficient (Staats and Regan 2014; Plummer et al. 2023). In our study, for example, *Phengaris nausithous*, a species listed as “vulnerable”, is counted on 14 transects by transect walkers entering their data online and on three transects by transect walkers submitting their data via email. However, a trend could not be calculated without these three transects of “old-fashioned” transect walkers (Fig. 13). For most of the “indicator” species also listed on the Red List, we cannot calculate species trends because of still insufficient data, underlining the high importance of every data on these taxa.

Consequently, data collected by transect walkers submitting their data via email are of high if not even crucial importance for “Butterfly Monitoring Germany”. They include data on species that are either rare and/or endangered or difficult to determine. Without these data, we would have considerably less information about these special species. Therefore, these data should be entered into the respective database under all circumstances even if this includes a lot of extra-work. Unfortunately, this is associated with considerable additional effort in the case of the email submitters because they already had entered their data in some electronic form of data sheet (often in an Excel sheet) which then has to be entered again manually into the online tool by someone that did not perform the transect walks. This is rather time-consuming and is prone to errors as someone of “Butterfly Monitoring Germany” staff has to enter 20 to 40 recording forms in a short time what needs a lot of attention.

(Modern) technology/exchange with other databases

An important question for “Butterfly Monitoring Germany” is how to minimise this extra-work without losing important data. In our questionnaire, many of the “old fashioned” transect walkers stated that they did not use the online data entry because it is too time consuming and/or too cumbersome. For this group, an easy and simple to use online entry tool offering additional benefits for the transect walkers would be highly beneficial. Most of the highly skilled transect walkers with very good knowledge of species also work in various other projects. In our questionnaire, some transect walkers stated that they already had entered their data into a different

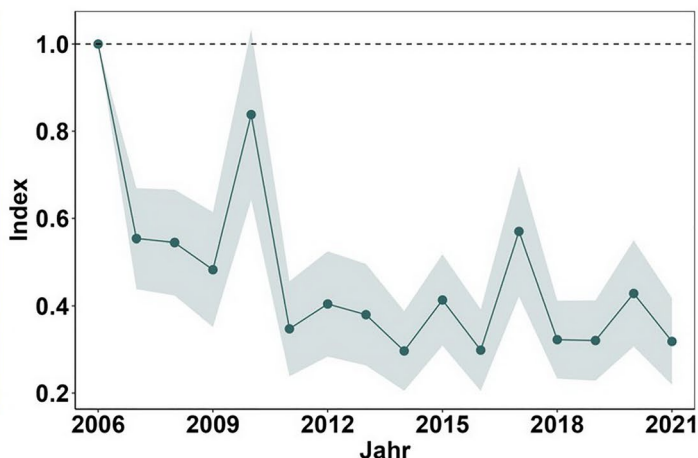


Fig 13 Dusky Large Blue (*Phengaris nausithous*), Foto: Josef Settele Trend of *Phengaris nausithous* in Germany, Source “Butterfly Monitoring Germany”

database and therefore were not willing to perform the double work to enter it in our online tool again. This group might be convinced to change their habits if being offered a good and easy to use online data entry offering some additional benefits, such as AI driven image recognition or the possibility to carry out own statistical evaluations.

Another important point for transect walkers might be the possibility to enter opportunistic data in addition to the highly standardised TMD transect data. This would allow participants to record single sightings of interesting species besides butterfly monitoring. In Germany, ObsIdentify (Schermer and Hogeweg 2018) and iNaturalist (Nugent 2018) are the most popular tools to capture, archive and manage own data. In addition, databases open to everyone also exist to record data on the state level (e. g. Bavaria, Brandenburg). For these tools, good interfaces are important because transect walkers would like to have the possibility to export/import their data from one tool to another.

The different forms of data recording and increased data quality delivered by “old fashioned” transect walkers comes with a trade-off: In the head office staff needs to spend considerable effort to enter the data manually into the central data base. Still, this results in harmonised data ready for analyses, so that scientists using the data in their studies do not need to worry about the different ways of recording, but receiving high-quality data. This once again emphasises the importance of transferring the data from the various data collection groups to a common database. Even if this means additional organisational effort, it pays off in the long term and can ultimately improve nature conservation.

Different people—different ways to collect data

In our questionnaire, 44% of the “old fashioned” transect walkers stated that they will not enter their data online under any circumstances. This astonishingly high percentage underlines that some participants always will reject such technologies. Some of these recorders used the comment section in our questionnaire to explain why they do not and will not use online data entry tools. They explained that they do not want to spend their time at the computer but rather like to go out to enjoy nature or that they generally have a negative attitude towards computer technology. In this context, it is crucial to keep in mind that Citizen Science projects are only partly for the sake of science, but also volunteers should benefit by being part of a community and having access to information about (butterfly) biodiversity. Therefore, also people not willing to use advanced online tools should be given the possibility to be part of such projects, even if their integration is labour-intensive for these projects.

Species knowledge

Our analyses underline the importance to capture the data of “old-fashioned” transect walkers submitting their records via email, even if this means a significant additional effort. Their records were filling important gaps and contributed to a substantially better data situation, especially on rare and/or endangered species. On the other hand, data collected by “old-fashioned” transect walkers submitting their records via letter did not contribute to more information about these species and did not fill knowledge gaps.

Nevertheless, involving all types of transect walkers is important for “Butterfly Monitoring Germany” on a completely different level. Numerous studies show that the knowledge about species or the recognition of species is getting worse, especially among children and young people (Frobel and Schlumprecht 2016; Sturm et al. 2020; Gerl et al. 2021). To be able to continue citizen science projects such as “Butterfly Monitoring Germany” on the long run, it is important to get young people interested in this subject and to train them. Hence, the young can learn from the old, and only when the knowledge available in the older generation is passed on, we can hope for a next generation that will continue such projects. This will not only be about the continuation of “Butterfly Monitoring Germany”, but also for the actions deriving from it. Previous research showed that Citizen Science projects can contribute educating the public about biodiversity (Peter et al. 2021) and that species literacy is an important factor for further actions on conserving biodiversity (Hooykaas et al. 2019, 2022).

Outlook and conclusion

Taking purely a data quality perspective, data of transect walkers submitting their records via letter do not provide additional information (i.e. increase data quality). In contrast if we excluded the data of transect walkers submitting their data via email, we would miss important information about specific butterfly species and with potential downstream impacts on trends analyses. For this group of transect walkers, it is worth trying to design a data entry tool that suits their needs to such an extent that using the tool piqued their interest. Main points in this case are that data entry has to be as easy as possible and that datasets can be up- or downloaded and im- and exported in different formats. However, if scientists make use of data collected by citizen scientists, they should always keep in mind that these types of projects should not only benefit science but also the people investing their time and effort. If this is considered, people are more likely to participate long-term in a project and thus help to collect long-term data important for the conservation of biodiversity (Kühn et al. 2022a).

Finally, people are different and have different interests and motivations to join a Citizen Science project. Many people joining a CS project in the field of biodiversity have an affinity for nature with a corresponding low interest in technology. Therefore, there will always be a certain amount of people who will not use online data entry tools, even if they are specially designed to their needs. Therefore, we advocate a pledge to “leave no one behind”, as stated in the agenda 2030 for Citizen Science (Montanari et al. 2021). Every participant should have the possibility to submit data in a way

that suits him or her best and should be assured that the data will be entered into the database and analysed.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10841-024-00577-0>.

Acknowledgements We would like to thank all participants of the Butterfly Monitoring Germany project for collecting butterfly data and filling out the questionnaire. We are very grateful to Sarah Passonneau for linguistic improvements.

Author contributions Elisabeth Kühn conceived the idea; Elisabeth Kühn and Ingolf Kühn developed the research questions and carried out the analyses, Alexander Harpke led the curation of the data. Elisabeth Kühn led the writing with contributions by all authors.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was funded by the Helmholtz-Centre for Environmental Research—UFZ, Halle (Germany).

Declarations

Competing interests No competing interests to declare.

Ethical approval Participation in the survey was voluntary, anonymous and questions could be skipped. The approval of an ethics committee is not required for such research in Germany. The study was conducted in line with the German data protection law. Results do not allow any conclusions to be drawn about the individuals.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aristeidou M, Herodotou C, Ballard HL et al (2021) How do young community and citizen science volunteers support scientific research on biodiversity? The Case of iNaturalist. *Diversity* 13:318. <https://doi.org/10.3390/d13070318>
- Bretz F, Hothorn T, Westfall P (2011) *Multiple Comparisons using R*. Chapman and Hall/CRC, New York
- Cáceres MD, Legendre P (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology* 90:3566–3574. <https://doi.org/10.1890/08-1823.1>
- Chandler M, See L, Copas K et al (2017) Contribution of citizen science towards international biodiversity monitoring. *Biol Cons* 213:280–294. <https://doi.org/10.1016/j.biocon.2016.09.004>
- Cowie RH, Bouchet P, Fontaine B (2022) The sixth mass extinction: fact, fiction or speculation? *Biol Rev* 97:640–663. <https://doi.org/10.1111/brv.12816>
- European Commission Directorate General for Environment (2022) *European Red List of insect taxonomists*. Publications Office, LU

- Frobel K, Schlumprecht H (2016) Ergebnisse einer Befragung und notwendige Reaktionen. *Naturschutz und Landschaftsplanung*
- Füchslin T, Schäfer MS, Metag J (2019) Who wants to be a citizen scientist? Identifying the potential of citizen science and target segments in Switzerland. *Public Underst Sci* 28:652–668. <https://doi.org/10.1177/0963662519852020>
- Gerl T, Randler C, Jana Neuhaus B (2021) Vertebrate species knowledge: an important skill is threatened by extinction. *Int J Sci Educ* 43:928–948. <https://doi.org/10.1080/09500693.2021.1892232>
- Hallmann CA, Sorg M, Jongejans E et al (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE* 12:e0185809. <https://doi.org/10.1371/journal.pone.0185809>
- Hooykaas MJD, Schilthuizen M, Albers CJ, Smeets I (2022) Species identification skills predict in-depth knowledge about species. *PLoS ONE* 17:e0266972. <https://doi.org/10.1371/journal.pone.0266972>
- Hooykaas MJD, Schilthuizen M, Aten C et al (2019) Identification skills in biodiversity professionals and laypeople: a gap in species literacy. *Biol Cons* 238:108202. <https://doi.org/10.1016/j.biocon.2019.108202>
- Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. *Biometrical J* 50:346–363. <https://doi.org/10.1002/bimj.200810425>
- Høye TT, Ärje J, Bjerger K et al (2021) Deep learning and computer vision will transform entomology. *Proc Natl Acad Sci USA* 118:e2002545117. <https://doi.org/10.1073/pnas.2002545117>
- Kühn E, Becker M, Harpke A et al (2022a) The benefits of counting butterflies: recommendations for a successful citizen science project. *Ecol Soc* 27:38
- Kühn E, Feldmann R, Harpke A et al (2008) Getting the public involved in butterfly conservation: lessons learned from a new monitoring scheme in Germany. *Israel Journal of Ecology & Evolution* 54:89–103. <https://doi.org/10.1560/IJEE.54.1.89>
- Kühn E, Musche M, Harpke A et al (2014) Tagfalter-Monitoring Deutschland-Anleitung. *Oedippus* 27:50
- Kühn E, Musche M, Harpke AA et al (2022b) Tagfalter-Monitoring Deutschland: Jahresauswertung 2021. *Oedippus* 40:57
- Mäder P, Boho D, Rzanny M et al (2021) The Flora Incognita app—Interactive plant species identification. *Methods Ecol Evol* 12:1335–1342. <https://doi.org/10.1111/2041-210X.13611>
- Miller JR (2005) Biodiversity conservation and the extinction of experience. *Trends Ecol Evol* 20:430–434. <https://doi.org/10.1016/j.tree.2005.05.013>
- Montanari M, Jacobs L, Haklay M et al (2021) Agenda 2030's, “Leave no one behind”, in citizen science? *JCOM* 20:A07. <https://doi.org/10.22323/2.20060207>
- Nugent J (2018) Citizen Science: iNaturalist. *Sci Scope* 041:12–13. https://doi.org/10.2505/4/ss18_041_07_12
- Peter M, Diekötter T, Kremer K, Höffler T (2021) Citizen science project characteristics: connection to participants' gains in knowledge and skills. *PLoS ONE* 16:e0253692. <https://doi.org/10.1371/journal.pone.0253692>
- Plummer KE, Dadam D, Brereton T et al (2023) Trends in butterfly populations in UK gardens—new evidence from citizen science monitoring. *Insect Conserv Diversity Icad* 17:12645. <https://doi.org/10.1111/icad.12645>
- R Core Team (2022) R: A language and environment for statistical computing
- Reinhardt R, Bolz R (2011) Rote Liste und Gesamtartenliste der Tagfalter (Rhopalocera) (Lepidoptera: Papilionoidea et Hesperioidea) Deutschlands. *Naturschutz und Biologische Vielfalt*
- RRStudio Team (2023) RStudio: Integrated Development for R
- Schermer M, Hogeweg L (2018) Supporting citizen scientists with automatic species identification using deep learning image recognition models. *BISS* 2:e25268. <https://doi.org/10.3897/biss.2.25268>
- Schmeller DS, Henry P-Y, Julliard R et al (2009) Advantages of volunteer-based biodiversity monitoring in Europe. *Conserv Biol* 23:307–316. <https://doi.org/10.1111/j.1523-1739.2008.01125.x>
- Staats WT, Regan EC (2014) Initial population trends from a 5-year butterfly monitoring scheme. *J Insect Conserv* 18:365–371. <https://doi.org/10.1007/s10841-014-9644-6>
- Sturm U, Voigt-Heucke S, Mortega KG, Moormann A (2020) Die Artenkenntnis von Berliner Schüler_innen am Beispiel einheimischer Vögel. *ZfDN* 26:143–155. <https://doi.org/10.1007/s40573-020-00117-8>
- Wäldchen J, Rzanny M, Seeland M, Mäder P (2018) Automated plant species identification—Trends and future directions. *PLoS Comput Biol* 14:e1005993. <https://doi.org/10.1371/journal.pcbi.1005993>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.