Simplified and still meaningful: assessing butterfly habitat quality in grasslands with data collected by pupils

Johannes Rüdisser1, Erich Tasser2, Janette Walde3, Peter Huemer4, Kurt Lechner1, Alois Ortner1, Ulrike Tappeiner1,2

Abstract Public participation in scientific research, now commonly referred to as citizen science, is increasingly promoted as a possibility to overcome the large-scale data limitations related to biodiversity and conservation research. Furthermore, public data-collection projects can stimulate public engagement and provide transformative learning situations. However, biodiversity monitoring depends on sound data collection and warranted data quality. Therefore, we investigated if and how trained and supervised pupils are able to systematically collect data about the occurrence of diurnal butterflies, and how this data could contribute to a permanent butterfly monitoring system. We developed a specific assessment scheme suitable for laypeople and applied it at 35 sampling sites in Tyrol, Austria. Data quality and its explanatory power to predict butterfly habitat quality was investigated comparing data collected by pupils with independent assessments of professional butterfly experts. Despite substantial identification uncertainties for some species or species groups, the data collected by pupils was successfully used to predict the general habitat quality for butterflies using a linear regression model ($r^2 = 0.73$, $p < 0.001$). Applying the proposed method in a citizen science context with laypeople could support both the long term monitoring of butterfly habitat quality, as well as the efficient selection of sites for professional in-depth assessments.

Keywords Lepidoptera · Citizen science · Monitoring · Indicator · Data quality · Volunteer

Introduction

Biodiversity assessments and continuous monitoring schemes are important and generally recognized prerequisites to efficiently address the ongoing biodiversity crises (Schmeller et al. 2015). Nevertheless, resources for collecting data on biodiversity are—and will always be—limited (Baan et al. 2013; Rüdisser 2015). Public participation seems to be an interesting option to support the collection and processing of biodiversity data (Domroese and Johnson 2017; Amano et al. 2016), and, at the same time, to generate authentic opportunities for environmental education (Chen and Cowie 2013). Such projects can lead to public engagement and can have transformative learning potential (Bela et al. 2016; Dickinson et al. 2012; Hobbs and White 2012; Lewandowski and Oberhauser 2017). The involvement and contribution of volunteers to academic research by collecting and submitting observation data has a long tradition in ecology and especially ornithology. For decades—the longest-running bird counting project in the US has collected data for over 110 years so far—thousands of amateur and professional ornithologists worldwide have been contributing to an otherwise unfeasible database (Tulloch et al. 2013). Public participation in scientific research, now often called ‘citizen science’ (CS) (cf. Irwin 1995), was for many years rarely reported in scientific publications. However,
this changed dramatically around the beginning of the current decade (Follett and Strezov 2015). Out of 1935 CS-related publications published before 2015-12-15 retrieved from the Web of Science by Kullenberg and Kasperowski (2016), 98% were published after 2000 and 75% after 2010. The observed boom in CS publications might still underestimate the potential for citizen science, because only a small portion of CS projects contribute data to peer-reviewed scientific articles (Theobald et al. 2015). The fast development of the internet, digital databases, and mobile devices has stimulated the development of innumerable CS projects with almost limitless thematic applications (Kobori et al. 2016; Tulloch et al. 2013), but the majority of studies applying CS can still be found in biology, ecology and conservation research (Kullenberg and Kasperowski 2016).

CS appears to be a powerful possibility to gather data about biodiversity over large temporal and spatial scales, and consequently make valuable contributions to monitoring activities. Today birds are the species most commonly observed in CS projects followed by terrestrial invertebrates, with the vast majority of projects focusing on butterflies (Follett and Strezov 2015). Both species groups are already used for the calculation of global and European biodiversity indicators covering species population and community composition trends (Butchart et al. 2010; van Swaay et al. 2008; van Swaay et al. 2015). Butterfly monitoring schemes have been established in many European countries with the help of highly dedicated volunteers (EEA-European Environment Agency 2013).

Biodiversity monitoring depends on sound data collection, appropriate sampling designs and data quality (Loos et al. 2015). While data quality issues are not distinctive to CS projects, sampling schemes that incorporate data from many observers—with different levels of knowledge and experience—can lead to varying levels of error and specific bias that must be addressed (Dickinson et al. 2010; McDonough MacKenzie et al. 2017; Ward 2014). When dealing with quality issues of biodiversity data collected by laypeople two main topics must be distinguished: (1) error and bias related to observation and identification accuracy and (2) sampling effort in respect to time, space, and taxonomic identification level (Lewandowski and Specht 2015). While the second aspect is strongly determined by the applied sampling scheme, both aspects are influenced by the skills and motivation of the involved observers. Many traditional CS projects were built upon the gratuitous contribution of experienced and skilled volunteers (Lukyanenko et al. 2016), now more and more projects involve laypeople. While the terms laypeople, layperson, volunteer, and citizen scientist are often used synonymously, one should be aware that the competence of volunteers contributing in citizen science projects can range from very experienced (often highly qualified and with specialised formal education) to totally unexperienced. In this article we focus on unexperienced volunteers hereinafter being referred to as ‘laypeople’. To involve laypeople in biodiversity monitoring in an inclusive way—as proposed by Lukyanenko et al. (2016)—simplified assessment protocols with flexible identification levels need to be developed and must be evaluated in regards to explanatory power and achievable data quality.

The citizen science project Viel-Falter (http://www.viel-falter.at), which was launched in 2013 aims to investigate if and how trained and supervised pupils can systematically collect data about the occurrence of butterflies, and how the data could be used to support a permanent biodiversity monitoring program as requested by international conventions (Convention on Biological Diversity (CBD) 2010; Han et al. 2014). While international experience has shown that involving qualified volunteers can substantially contribute to the implementation of such monitoring tasks (Amano et al. 2016; Burgess et al. 2016), only few studies have focused on laypeople or pupils (cf. Abadie et al. 2008; Ballard et al. 2017; Olivier et al. 2015). Therefore, we developed and applied a specific assessment scheme suitable for pupils or any other layperson, to systematically collect data on butterflies. By comparing with independent assessments conducted by professional butterfly experts, we investigated if the achieved data quality is sufficient to support a permanent butterfly monitoring system. Additionally, we investigated how the pupil’s motivation to engage in butterfly observation activities develops during the course of the project and what project factors might be crucial to support a continuous engagement. In this article, we focus on the question of whether the proposed simplified butterfly assessment scheme, which can be executed by any trained layperson, could be used to determine butterfly habitat quality of different sites. We investigated if we can predict a butterfly habitat quality index, based on comprehensive expert assessments on species level, with data derived from a simplified assessment scheme. Using data from both experts and laypeople we tried to separate the effects of the information reduction caused by the simplified assessment scheme from the effects related to inaccurate data collection.

Methods

Participating schools and study sites

From 2013 to 2015 548 pupils, of ages ranging from 6 to 20 years and from 14 different schools, collected data at 35 sampling sites in Tyrol, Austria. The sites which are distributed all over Tyrol were chosen to represent characteristic butterfly grassland habitats in and around settlement areas.
Thus, the sites, ranging in altitude from 505 to 1485 m above sea level, included hay meadows (n = 27), pastures (n = 4), and fallow land or ruderal sites (n = 4). Land use intensity ranged from none or extensive (one cut per year) to very intensive (three or more cuts per year).

### Simplified assessment scheme

Butterfly surveys were performed by supervised school classes using a predefined and visualised list of 13 characteristic species and ten species groups (Fig. 1). To facilitate the assessments morphologically similar species were grouped. The species groups consisted of species from the same genus (*Coenonympha, Erebia*), from a group of genera (*Apatura & Limenitis*), from the same subfamily or tribe (*Coliadinae, Heliconiinae, Melitaeini, Pyrginae*), and from the same family (*Lycaenidae, Hesperiidae, Pieridae*). The selection of species and species groups was based on habitat preferences (Spiss 2014). The species and species groups used in Viel-Falter were similar to those used by Olivier et al. (2015) for butterfly monitoring in private gardens in France.

In order to ensure survey quality, pupils and their teachers were intensively trained during two half-day workshops in butterfly identification and on how to perform the survey. Additionally they received butterfly identification keys and learning materials. During the surveys, pupils scattered themselves across the study site, remaining at least 10 m apart. All pupils used the graphic butterfly list (Fig. 1) to support identification and counted all the butterflies within a semi-circle of five meters and for a period of 5 min. If the pupils observed a species that was not on the list or they were unsure about, they were asked to record it in the aggregated groups indicated with rectangles in Fig. 1. Schools were asked to conduct at least three surveys in May, June, and July and—following the recommendations of Pollard & Yates (1993)—between 10:00 and 17:00, during sunny or warm days with no or only low wind speed.

---

**Fig. 1** The simplified butterfly assessment scheme used by pupils for identification and systematic counting of butterflies. The original version used by the schools was printed with German names and in A3 format so that the size of the illustrated butterflies corresponded to their natural size. Rectangles indicate the aggregated species groups used to predict butterfly habitat quality.
After conducting the surveys the pupils transferred the data into the project database using an online application developed for this purpose (http://www.viel-falter.at). All analyses presented in this article are based on data collected in the years 2013–2015.

**Expert butterfly surveys**

At all study sites \(n=35\) a butterfly expert conducted four comprehensive butterfly assessments on species level. These expert assessments were executed in the same season as by the schools (May–July) for at least 2 years between 2013 and 2015. Consecutive assessments at the same site were separated by at least three weeks. Assessments were conducted, as recommended by Pollard & Yates (1993), between 10:00 and 17:00, during sunny or warm days, and with no or only low wind speed. To guarantee a similar sampling effort at each study site, diurnal butterflies were counted for 30 min from an area of approximately 2500 m².

**Estimating butterfly habitat quality**

Although butterflies are widely accepted as a meaningful and applicable indicator to investigate and monitor biodiversity changes (Pearman and Weber 2007; van Swaay et al. 2008, EEA—European Environment Agency 2012; Feest et al. 2011), there is no commonly applied indicator that measures butterfly habitat quality. The two most commonly applied measures for habitat quality are species richness and abundance. While species richness—if measured properly—is a natural measure of biodiversity (Gotelli and Colwell 2001), abundance incorporates aspects of ecosystem functions and services (Schwartz et al. 2000). Therefore, we combined both species richness and abundance to a butterfly habitat quality index (BHQ). BHQ was calculated for each site by multiplying the rarefied species richness \(S_{\text{est}}\) with butterfly abundance \(A\) estimated on the basis of the pooled data from four expert surveys at each site. Gotelli and Colwell (2001) clearly illustrated that the number of observed species at different sites using a standardized sampling effort depends not only on species richness, but also on the number of individuals, or the mean density of individuals. Thus the observed number of species per sampling represents species density—the number of species per sampling area—and not species richness, which is the number of observed species in relation to the number of potential species. In order to account for the effect of abundance in the estimation of species richness, we used individual-based rarefied species richness for 50 individuals for the calculation of the BHQ (Chao et al. 2014). Rarefaction is a technique computing the expected number of species \(S_{\text{est}}\) from a so-called individual-based rarefaction curve that provides the estimated dependence of the number of species on the accumulated number of individuals (Gotelli and Colwell 2001). This curve is constructed by drawing repeatedly random subsamples from the sampled at a site and plotting the mean number of obtained species against the subsample size (i.e., the number of sampled individuals). \(S_{\text{est}}\) was computed using EstimateS (Version 9, R. K. Colwell, http://purl.oclc.org/estimates). Both \(S_{\text{est}}\) and \(A\) were normalized with the min–max method. Hence BHQ was calculated as:

\[
\text{BHQ}_i = \left( \frac{S_{\text{est},i} - S_{\text{est},\text{Min}}}{S_{\text{est},\text{Max}} - S_{\text{est},\text{Min}}} \right) \ast \left( \frac{A_i - A_{\text{Min}}}{A_{\text{Max}} - A_{\text{Min}}} \right)
\]

where \(S_{\text{est},i}\) is the rarefied species richness at site \(i\) and \(A_i\) the abundance at the corresponding site.

BHQ, which theoretically ranges from 0 to 1, was used as dependent variable in all of the following analyses. However, before using BHQ, we validated its relevance for biodiversity and conservation related aspects by comparing BHQ values at all sites with the number of butterfly species on the national Red List and with the number of individuals from Red List species (cf. Höttinger and Pennerstorfer 2005). For this we estimated Spearman’s rank correlation coefficient, because this relationship might not be linear.

**Analysis of the explanatory power of the simplified assessment scheme**

To investigate the validity of the simplified assessment scheme, we explored if the BHQ which was obtained on the basis of detailed expert surveys could be assessed with data from a simplified assessment scheme. To separate the effects of the information reduction caused by the aggregation into species groups, from effects related to possible biases or errors resulting from wrong butterfly identification by laypeople, the analysis was split in two different steps. First, the explanatory power of the simplified assessment scheme to predict butterfly habitat quality was analysed using only the data collected by experts, and second the same analysis was conducted with the data collected by schools (Fig. 2): to investigate the explanatory power of the reduced assessment scheme to predict BHQ we aggregated the individual species abundance data from the expert surveys to the following nine species groups: Pieridae (excl. Coliadinae), Coliadinae, Nymphalinae (excl. Melitaeini), Lycaenidae, Satyrinae, *Apatura & Limenitis*, Heliconiinae plus Melitaeini, Hesperiidae (without Pyrginae), and Pyrginae (Species groups are indicated with rectangles in Fig. 1). These nine species groups complemented with the three very characteristic and large species (*Parthenos sylvia*, *Papilio machaon*, and *Nymphalis antiopa*) were then used as independent variables to predict BHQ with a linear regression model. *Iphiclides podalirius* was
excluded from analysis, because it occurs in only small parts of the study region. We used linear regression models because BHQ is based equally on \( S_{\text{est}} \) and A and therefore a direct recalculation on the basis of aggregated species data does not make sense, because the aggregation to species groups impairs the calculation of \( S_{\text{est}} \). Because the dataset comprised only 35 sites but 12 independent variables, the degrees of freedom were small, consequently the power of the statistical tests weak, and the risk for overfitting is high (Babyak 2004). Therefore, we opted to reduce the number of variables using principal component analysis (PCA) with Varimax as rotation method. A PCA extracts orthogonal components, each component comprises all variables highly correlated with each other but separating these variables form the ones not highly correlated with them. Components with an eigenvalue larger than one were employed to identify groups of variables containing almost the same information. In our case the variables used in PCA were the pooled abundance data from the four expert surveys aggregated to the nine species groups plus the three characteristic species listed above. Based on the PCA results, we selected a species or species group as representative for each component. Hence, the selection of these variables was done completely independent from the BHQ which was used (as dependent variable) in the regression step: estimating the explanatory power of the simplified assessment scheme to predict BHQ, just using the selected representatives as explanatory variables in the linear regression model (Fig. 2).

**Results**

**Data collection and commitment of schools**

Between 2013 and 2015, the 14 schools involved conducted 159 site visits and 2616 individual butterfly assessments. Although most schools participated with a very high level of commitment, two schools, due to internal organizational and motivational reasons, did not manage to conduct at least three independent visits per site following the outlined assessment protocol. After exclusion of the data from those schools, the final dataset used for the validation of the simplified assessment scheme consisted of 2456 individual assessments during 151 site visits at 30 different sites. The
number of school visits per site ranged thereby from 3 to 9 resulting in 30–163 individual assessments per site.

Despite some organizational problems with a few schools, 83% of the pupils stated in an anonymous online survey that they would like to participate in a similar project again. 87% of the pupils reported a high degree of interest and enjoyment in participating.

During 140 site visits the butterfly experts identified 73 different butterfly species at 35 sites. The number of recorded species per site ranged from 5 to 28 and the number of individuals ranged from 17 to 191. *Aglais urticae, Coenonympha pamphilus* and *Polyommatus icarus* were the species with the highest frequency, occurring at 91, 80, and 77% of the investigated sites respectively. The most abundant species was *Maniola jurtina* with 308 observed individuals, followed by *Polyommatus icarus* with 182, *Coenonympha pamphilus* with 178, *Aglais urticae* with 154, and *Pierisrapae* with 149 observed individuals. Fourteen species occurred only at two sites and 13 species only at one site.

### Data accuracy and quality

Comparing the data collected by pupils with the data from expert surveys revealed that the degree of accordance varied substantially between different species or species groups (Table 1). Pearson’s correlation coefficient between abundance data from experts and schools ranged from not significant for *Nymphalis antiopa, Apatura & Limenitis, Coliadinae & Pyrginae* to above 0.9 for *Parnassius apollo* and *Hesperidae* (Table 1).

### Butterfly habitat quality (BHQ) and endangered species

An important aspect of butterfly habitat conservation is the protection of endangered species. To investigate the indicator capacity of BHQ for rare or endangered species, we compared it with the occurrence of Red List species (Fig. 3). Both Red List species richness and Red List species abundance were significantly correlated with BHQ (Spearman $R=0.73$, $p<0.001$, and $R=0.71$, $p<0.001$, respectively). Red List species richness and Red List species abundance were also correlated (but with lower correlation coefficients) with the two components of BHQ.
namely rarefied species richness (Spearman $R=0.53$, $p<0.05$, and $R=0.37$, $p<0.05$, respectively) and abundance (Spearman $R=0.60$, $p<0.001$, and $R=0.65$, $p<0.001$, respectively). This indicates that BHQ is a sensitive proxy for the occurrence of rare species and has more predictive power than each of its components.

Explanatory power of the simplified assessment scheme

The linear regression model to predict BHQ using all 12 variables of the aggregated species groups based on expert data provided a very good fit ($r^2=0.92$, adjusted $r^2=0.87$, $p<0.001$, $n=35$). Although all variance inflation factors (VIFs) were smaller than four we opted to reduce the number of independent variables to increase the power of the statistical tests and to reduce collinearities in order to obtain appropriate standard errors and therefore $p$ values. This was appropriate due to the relatively low number of sampling sites ($n=35$). Thus, we applied PCA with Varimax rotation using all 12 variables of the aggregated species groups from the expert’s data and selected a representative variable for each component. Based on the PCA results we selected *Parnassius apollo*, Nymphalinae, Lycaenidae, Heliconiinae & Melitaeini and Hesperiidae as representative variables. All selected variables—with the exception of Hesperiidae—were those with the highest component loadings. Hesperiidae were selected instead of *Apatura* & *Limenitis*, because they occur more frequently. *Nymphalis antiopa*, a typical species in woody habitat, would have been representative for a 6th component, but we excluded the species from further analysis due to its rarity in the study area and our focus on grasslands. The linear regression model to predict BHQ using only the five selected variables still provided a very appropriate fit ($r^2=0.86$, adjusted $r^2=0.84$, $p<0.001$, $n=35$) with all variables being significant ($p<0.05$) and all VIFs being smaller than 1.3. In the next step, we assessed the explanatory power of the data collected by schools. We used the same five selected variables and estimated the model again via linear regression analysis. The model based on the data collected by schools still provided a good fit ($r^2=0.73$, adjusted $r^2=0.67$, $p<0.001$, $n=30$), although the variables Lycaenidae and Nymphalinae were not significant (Fig. 4).

Discussion

Estimating butterfly habitat quality

While inconsistent or low data quality can be a severe obstacle to the scientific use of citizen science data (Burgess et al. 2016), a systematic review of peer-reviewed articles dealing with the quality of data collected by volunteers revealed that many citizen science projects already proved its potential to collect high-quality data (Lewandowski and Specht 2015). Although well-designed and executed citizen science projects with adequate volunteer training can provide biodiversity data comparable to data collected by professionals (Holt et al. 2013; Kremen et al. 2011; Lewandowski and Specht 2015; Lovell et al. 2009) and support the conservation of endangered species (Jue and Daniels 2015), new survey schemes involving laypeople should be rigorously evaluated regarding its adequacy and applicability for the involved volunteers and its final data quality. The characteristics of survey protocols including preliminary training, survey duration, or number of surveyors can have important implication on detection probability and identification quality (Albergoni et al. 2016). Furthermore, the desirable inclusion of a broad array of citizens in the collection of biodiversity data (Lukyanenko et al. 2016) may make it necessary to use higher level classification instead of species identification, and the resulting loss of information must be evaluated regarding its remaining informative value for the desired purpose (cf. Le Féon et al. 2016).

To compare butterfly diversity of different sites (or for analysing trends in time) one needs applicable and meaningful measures (MacDonald et al. 2017). There is a wide consensus that diversity of an ecosystem consists of two components: richness (the number of species or traits) and evenness (the proportional abundance of species or traits among a community) (Hillebrand and Matthiessen 2009; Tuomisto 2012). While species richness is a commonly applied and straightforward measure of diversity there
is much debate about adequate measures for community abundance composition (Tuomisto 2012). The increasing emphasis on the question how biodiversity affects ecosystem functions and related services, by both research and politics, has further boosted this debate. Although some authors reported effects of community abundance composition on ecosystem functions (Hillebrand et al. 2008; Tilman et al. 2014), the effects of changing evenness on ecosystem properties cannot be generalised and might be more influenced by specific arrangements of dominant species than by evenness per se (Winfree et al. 2015; Wohlgemuth et al. 2016). MacDonald et al. (2017) analysed data from a long-term butterfly monitoring program and revealed that species richness and evenness of the observed butterfly communities were negatively related. They conclude that evenness might be an important and helpful additional measure to investigate differences in butterfly communities but should not be used in a compound index (e.g. Shannon-Wiener index or Simpson’s index). Therefore we deliberately avoided using evenness indicators as a measure for habitat quality in our study but opted to consider overall butterfly abundance as an additional measure related to ecosystem functions and services. It is well known that population sizes of most butterfly species fluctuate from year to year due to changing weather conditions for example (Pollard and Yates 1993; Thomas 2005). Although we aggregated data from four expert surveys in two consecutive years for the calculation of BHQ to reduce the effect of short term population fluctuations, the relatively short time period of our survey weakens its informative value regarding absolute habitat quality. However, population fluctuations of butterfly species are generally similar over large areas and therefore relative values of population index comparing different sites remain fairly stable (Pollard and Yates 1993; Roy et al. 2015). Thus we are confident that BHQ is a suitable metric to compare different sites and to estimate the explanatory power of the simplified butterfly assessment conducted by laypeople.

Butterfly data collected by laypeople should be used and analysed thoughtfully. Our systematic and planned comparison of expert butterfly assessments with data collected by pupils using a simplified assessment scheme illustrates both potentials and limitations of such data. Comparison of abundance data from species and species groups collected by pupils with the data from experts revealed substantial differences. While some of these differences might be caused by time lags between the surveys conducted by pupils and experts, the varying degree of accordance (Table 1) indicates substantial identification uncertainties. This is in line with Vantieghem et al. (2016) who encountered substantial species identification errors among morphologically similar skipper butterflies examining photographs uploaded with observations from citizen scientists. Data collected by laypeople using aggregated species groups should definitely not be seen as a substitute to butterfly assessment on species level. Such detailed and high quality data are irreplaceable for species distribution and population estimations in the context of nature conservation and environmental impact assessments. However, data from simplified butterfly assessment schemes applicable for laypeople could complement the professional collection and processing of biodiversity data in the context of long term biodiversity monitoring (Rüdisser 2015; Theobald et al. 2015).

The abundance data of *Parnassius apollo* and the four species groups Nymphalinae, Lycaenidae, Heliconiinae & Melitaeini, and Hesperiidae collected by pupils could explain 73% of the variance of the butterfly habitat quality measured by the BHQ index. The BHQ index was obtained on the basis of completely independent and detailed expert surveys at species level. Considering the heterogeneous character of the involved school classes as well as the observed and reported identification uncertainties, the predictive power of the data collected by pupils is surprisingly high. As BHQ is calculated on the basis of species richness and total butterfly abundance assessed by experts one could speculate that the high explanatory power of the data collected by pupils could be driven mainly by butterfly abundance differences and therefore just counting butterflies without any further identification would lead to similar results. However Pearson correlation between BSQ and the overall butterfly abundance assessed by pupils was low ($r=0.23$) and not significant, indicating that species identification—at least at the coarse level proposed—in fact is important for adequately predicting the BHQ. As we had only 30 sites with parallel butterfly assessments of pupils and experts we opted to reduce the number of independent variables (observed species groups by pupils) to increase statistical power. The selection of these variables was based on a PCA including all independent variables and not on the explanatory power in the subsequent regression model. It is important to notice that the decision to use only five explanatory variables to predict BSQ instead of the whole dataset has only methodological reasons. Therefore further studies leading to a larger sample size would permit the inclusion of the data of additional species groups as independent variables, and likely lead to an even better predictability of the BHQ. Although our results should be evaluated with further investigations (including additional sites), the results are already very good news for citizen science initiatives that aim to foster biodiversity assessments with the help of laypeople. The proposed observation and reporting method for diurnal butterflies can be relatively easily applied by any layperson after elementary training. Although these assessments might not serve to draw conclusions about the occurrence of individual species, they

© Springer
could permit meaningful predictions about the general habitat quality for butterflies on the investigated sites and its surroundings. As adult (nectarivorous) butterflies—on which we are focusing—might use different habitats than their larval stages, results should be interpreted from a landscape perspective rather than focusing on individual site characteristic.

Another option to use simplified butterfly surveys is in the context of result-oriented measures for biodiversity protection as done by Stolze et al. (2015). Especially, if considering ecosystem services such as pollination or the potential for resilience in ecosystems, rough but affordable estimates based on volunteer surveys could complement biodiversity monitoring programs as proposed by Obrist and Duelli (2010). However, simplified butterfly surveys conducted by laypeople should not be seen as a substitute for expert surveys. The use of species groups impedes conclusions about distinct species and it is impossible to calculate estimates on beta or gamma diversity at the landscape scale (Obrist and Duelli 2010). For inventories, biographical studies, or prioritization of sites for conservation, expert surveys on species level are needed (Krell 2004).

Citizen science and biodiversity monitoring

The principal aim of the presented study was to investigate whether trained and supervised pupils together with their teachers are able to systematically collect data about the occurrence of diurnal butterflies. While international experiences showed that involving adult (and often experienced) volunteers can substantially contribute to the successful implementation of biodiversity monitoring programmes (Miller-Rushing et al. 2012; Schmeller et al. 2009; Pocock et al. 2015), until now only few projects have focused on juvenile or unexperienced volunteers. Our findings support the assumption that even data from laypeople (such as pupils) may contribute to biodiversity monitoring tasks. While relying on volunteers might seem a cost-effective approach to collect large scale biodiversity data (Gardiner et al. 2012; Theobald et al. 2015), one should be aware that CS projects are not cheap or even free of charge (Chandler et al. 2016). The development and maintenance of a well-functioning framework for effective data collection, storage and analysis, as well as the training and support of volunteers to achieve the desired data quality are often demanding, time consuming and hence costly tasks (Tulloch et al. 2013). However, cost analysis of CS projects should consider that CS often provides many benefits beyond its merely scientific aim (McKinley et al. 2017). Involving volunteers in the collection of biodiversity data can increase both scientific and environmental literacy as well as engagement. In our case the cooperation with schools enabled us to reach many and different pupils, but also resulted in additional challenges. Communication and organisational aspects were often demanding and fluctuation of responsible teachers due to school change or leave was surprisingly high in some schools. While very young pupils clearly needed more support from their teachers, organisational aspects and the motivation of the involved pupils were better in primary schools than in middle and high schools. However, the high degree of interest and enjoyment in participating in the project, as well as several grassroots projects autonomously initiated by the participating schools including peer teaching, environmental protection, and awareness rising activities confirmed the potential to promote education and empowerment with well-designed CS projects.

Conclusions

Summarizing, data collected by laypeople using the proposed survey method could support both the long-term monitoring of butterfly habitat quality, as well as the efficient selection of sites for further investigation conducted by experts. In practice, monitoring-data from laypeople could serve for ‘surveillance’ purposes as proposed by Forrester et al. (2015) to facilitate early protection of environmental changes. Furthermore spatially comprehensive observation by volunteer laypeople could be the basis for a cost efficient site selection for focused and complementary expert assessments. However, for an effective biodiversity monitoring a long term perspective including the necessary funding for a continuous collection of data by both experts and laypeople is needed. Promoting an easy way of butterfly observation could broaden the addressed target groups and enthuse more individuals for nature observation and biodiversity issues. In successful CS projects (some) laypeople gradually gain experience, resulting in positive effects on data quality and meaningfulness. They sometimes even turn into experts. Butterfly observation conducted by laypeople should not substitute expert assessments, but rather complement them. More data—collected by both professionals and laypeople—are needed to develop a robust forecasting tool for butterfly habitat quality based on the simplified survey scheme, to establish benchmarks, and to strengthen the results of this study.

Acknowledgements Open access funding provided by University of Innsbruck and Medical University of Innsbruck. Our special thanks go to 24 teachers and 548 pupils of the participating schools (BRG in der Au, HLW Kufstein, NMS Fließ, NMS Längenfeld, NMS Umhausen, NMS Weer, PHT/PMS der Pädagogischen Hochschule Tirol, Reithmannsgymnasium, VS Brandberg, VS Innervillgraten, VS Obsteig, VS Schwendt, VS Steinach and, VS Tux). The project Viel-Falter was financially supported by the Federal Ministry of Science, Research, and Economy (Sparkling Science Project Viel-Falter, SPA 04/018 and Top Citizen Science Initiative; TCS 12 Viel-Falter TCS).
Ulrike Tappeiner and Johannes Rüdisser are a member of the research area ‘Alpine Space— Man and Environment’ at the University of Innsbruck. Finally, we wish to thank Hans-Peter Wymann for providing butterfly pictures and Jacob Dein and Julliane de Oliveira Rüdisser for proof reading.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Abadie J-C, Andrade C, MacHon N, Porcher E (2008) On the use of parataxonomy in biodiversity monitoring: a case study on wild flora. Biodivers Conserv 17(14):3485–3500. doi:10.1007/s10531-008-9354-z

Albergoni A, Bride I, Scialfa CT, Jocque M, Green S (2016) How appropriate are volunteers for visual biodiversity surveys? An evaluation of skill level and group size during a conservation expedition. Biodivers Conserv 25(1):133–149. doi:10.1007/s10531-015-1039-9

Amano T, Lammers JL, Sutherland WJ (2016) Spatial gaps in global biodiversity information and the role of citizen science. Bioscience 66(5):393–400. doi:10.1093/biosci/biw022

Baan L de, Alkemade R, Koellner T (2013) Land use impacts on biodiversity in LCA: a global approach. Int J Life Cycle Assess 18(6):1216–1230. doi:10.1016/j.ijslc.2012-04-012-0

Babyak MA (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosomat Med 66(3)

Ballard HL, Dixon CG, Harris EM (2017) Youth-focused citizen science: examining the role of environmental science learning and agency for conservation. Biol Conserv 208:65–75. doi:10.1016/j.biocon.2016.05.024

Bela G, Peltola T, Young JC, Alkemade R, Koellner T (2013) Land use impacts on biodiversity in LCA: a global approach. Int J Life Cycle Assess 18(6):1216–1230. doi:10.1016/j.ijslc.2012-04-012-0

Butchart SHM, Walpole P, Colwell R, Young BE, Brooks TM, Sánchez de Lozada A, Bubb P, Butchart SHM, Larsen FW, Hamilton H, Hansen MC, Turner WR (2014) A biodiversity indicators dashboard: addressing challenges to monitoring progress towards the aichi biodiversity targets using disaggregated global data. PLoS ONE 9(11):e101185. doi:10.1371/journal.pone.0101185

Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecol Monogr 84(1):45–67. doi:10.1890/13-0133.1

Chen J, Cowie B (2013) Developing ‘butterfly warriors’: a case study of science for citizenship. Res Sci Educ 43(6):2153–2177. doi:10.1007/s11165-013-9349-y

Convention on Biological Diversity (CBD) (2010) COP 10 Decision X/2: Strategic Plan for Biodiversity 2011–2020. http://www.cbd.int/decision/cop/?id=12268

Dickinson JL, Zuckerberg B, Bonter DN (2010) Citizen science as an ecological research tool: challenges and benefits. Annu Rev Ecol Evol Syst 41:149–172. doi:10.1146/annurev-ecolsys-102009-104339

Dickinson JL, Shirk J, Bonter D, Bonney R, Crain RL, Martin J, Phillips T, Purcell K (2012) The current state of citizen science as a tool for ecological research and public engagement. Front Ecol Environ 10(6):291–297. doi:10.1890/110236

Domroese MC, Johnson EA (2017) Why watch bees? motivations of citizen science volunteers in the great pollinator project. Biol Conserv 208:40–47. doi:10.1016/j.biocon.2016.08.020

EEA—European Environment Agency (2012) Streamlining European biodiversity indicators 2020: Building a future on lessons learnt from the SEBI 2010 process.

EEA—European Environment Agency (2013) The European Grassland Butterfly Indicator: 1990–2011. EEA Technical report, vol 11, Luxembourg

Feest A, van Swaay C, Aldred TD, Jedamzik K (2011) The biodiversity quality of butterfly sites: a metadata assessment. Ecol Indic 11(2):669–675

Follett R, Strezov V (2015) An analysis of citizen science based research: usage and publication patterns. PLoS ONE 10(11):e0143687. doi:10.1371/journal.pone.0143687

Forrester G, Baily P, Conetta D, Forrester L, Kintzing E, Jarecki L (2015) Comparing monitoring data collected by volunteers and professionals shows that citizen scientists can detect long-term change on coral reefs. J Nat Conserv 24(0):1–9. doi:10.1016/j.jnc.2015.01.002

Gardiner MM, Allee LL, Brown PMJ, Losey JE, Roy HE, Smyth RR (2012) Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs. Front Ecol Environ 10(9):471–476. doi:10.1890/110185

Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol Lett 4(4):379–391. doi:10.1046/j.1461-0248.2001.00230.x

Han X, Smyth RL, Young BE, Brooks TM, Sánchez de Lozada A, Bubb P, Butchart SHM, Larsen FW, Hamilton H, Hansen MC, Turner WR (2014) A biodiversity indicators dashboard: addressing challenges to monitoring progress towards the aichi biodiversity targets using disaggregated global data. PLoS ONE 9(11):e10246. doi:10.1371/journal.pone.010246

Hillebrand H, Matthiessen B (2009) Biodiversity in a complex world: consolidation and progress in functional biodiversity research. Ecol Lett 12(12):1405–1419. doi:10.1111/j.1461-0248.2009.01388.x

Hillebrand H, Bennett DM, Cadotte MW (2008) Consequences of dominance: a review of evenness effects on local and regional ecosystem processes. Ecology 89(6):1510–1520. doi:10.1890/10-1053.1

Hobs SJ, White PCL (2012) Motivations and barriers in relation to community participation in biodiversity recording. J Nat Conserv 20(6):364–373. doi:10.1016/j.jnc.2012.08.002
MacDonald ZG, Nielsen SE, Acorn JH (2017) Negative relationships. J Insect Conserv 21:677–688
Tasser E, Vincent A, Wezel A (2015) Result-oriented Measures for Biodiversity in Mountain Farming: A Policy Handbook
Theobald EJ, Ettinger AK, Burgess HK, DeBey LB, Schmidt NR, Froehlich HE, Wagner C, HilleRisLambers J, Tewksbury J, Harsch MA, Parrish JK (2015) Global change and local solutions: tapping the unrealized potential of citizen science for biodiversity research. Biol Conserv 181(0):236–244. doi:10.1016/j.biocon.2014.10.021

Thomas JA (2005) Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. Phil Trans R Soc B 360(1454):339–357

Tilman D, Isbell F, Cowles JM (2014) Biodiversity and ecosystem functioning. Annu Rev Ecol Evol Syst 45(1):471–493. doi:10.1146/annurev-ecolsys-120213-091917

Tulloch AIT, Possingham HP, Joseph LN, Szabo J, Martin TG (2013) Realising the full potential of citizen science monitoring programs. Biol Conserv 165:128–138. doi:10.1016/j.biocon.2013.05.025

Tuomisto H (2012) An updated consumer’s guide to evenness and related indices. OIKOS 121(8):1203–1218. doi:10.1111/j.1600-0706.2011.19897.x

van Swaay CA, Nowicki P, Settele J, van Strien AJ (2008) Butterfly monitoring in Europe: methods, applications and perspectives. Biodivers Conserv 17(14):3455–3469

van Swaay CA, van Strien A, Aghababayan K, Astrom S, Botham M, Brereton T, Chambers P, Collins S, Ferre MD, Escobes R, Feldmann R, Fernandez-Garcia JM, Fontaine B, Goloshchapova S, Gracianteparaluceta A, Harpke A, Heliola J, Khanamirian G, Juliard R, Kuhn E, Lang A, Leopold P, Loos J, Maes D, Mestdagh X, Monasterio Y, Munguira ML, Murray T, Musche M, Ounap E, Pettersson L, Popoff S, Prokofev I, Roth T, Roy D, Settele J, Stefanescu C, Svirta G, Teixeira SM, Tittsaar A, Verovnik R, Warren M (2015) The european butterfly indicator for grassland species 1990–2013. http://nora.nerc.ac.uk/511714/

Ward DF (2014) Understanding sampling and taxonomic biases recorded by citizen scientists. J Insect Conserv 18(4):753–756. doi:10.1007/s10841-014-9676-y

Winfree R, Fox JW, Williams NM, Reilly JR, Cariveau DP (2015) Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. Ecol Lett 18(7):626–635. doi:10.1111/ele.12424

Wohlgemuth D, Solan M, Godbold JA (2016) Specific arrangements of species dominance can be more influential than evenness in maintaining ecosystem process and function. Scientific Rep 6:39325