



HAL
open science

Science learning in biodiversity citizen science: Inputs from the analysis of online social interactions within a contributory project for pollinators' monitoring

Baptiste Bedessem, Ana-Cristina Torres, Colin Fontaine, Nicolas Deguines

► To cite this version:

Baptiste Bedessem, Ana-Cristina Torres, Colin Fontaine, Nicolas Deguines. Science learning in biodiversity citizen science: Inputs from the analysis of online social interactions within a contributory project for pollinators' monitoring. *Biological Conservation*, 2022, 276, pp.109807. 10.1016/j.biocon.2022.109807. hal-03874485

HAL Id: hal-03874485

<https://hal.science/hal-03874485v1>

Submitted on 28 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Science learning in biodiversity citizen science: Inputs from the analysis of online social interactions within a contributory project for pollinators' monitoring

Baptiste Bedessem^{a,*}, Ana-Cristina Torres^b, Colin Fontaine^a, Nicolas Deguines^c

^a Centre d'écologie et des sciences de la conservation, 43 rue Buffon, 75005 Paris, France

^b Lab'Urba, 14-20, Boulevard Newton, Champs-sur-Marne, 77454 Marne la Vallée, France

^c Université de Poitiers, CNRS, EBI, Poitiers, France

ARTICLE INFO

Keywords:

Citizen science
Biodiversity
Science learning
Procedural skills
Seasonal cycles

ABSTRACT

Biodiversity citizen science projects are both valued for their contribution to scientific research and for their impact on participants' science learning and engagement towards the environment. In this paper, we assess the impact of participation in a biodiversity citizen science project (the Spipoll, dedicated to pollinators' monitoring) through the analysis of online interactions within the program's data sharing platform. By drawing on a previous qualitative analysis of the comments exchanged by the participants within this platform, we focus on those comments which share items on aspects of biology and ecology related to the Spipoll program. This sample gathers 2279 comments from 2010 to 2018. We first classified the different constitutive elements from these comments into seven categories following the topics they deal with. We then studied the temporal change in occurrence of each of these topics from 2010 to 2018. We show that long-term participation is associated with the growing expression of scientific procedural skills: formulation of hypothesis and explanation, proposition of new research questions. To our knowledge, our study is the first one that detects the acquisition of such procedural skills in biodiversity citizen science. We also show that long-term participation is associated with the growing attention to natural seasonal cycles. This study finally illustrates the value of the online traces of citizen scientists' activities to analyze participants' outcomes of citizen science. Consequently, it should encourage the development of such online communication spaces within contributory projects, without restricting them to online citizen science.

1. Introduction

Broadening citizen participation in knowledge production about biodiversity has become an explicit objective of national and supranational institutions (European Commission, 2013; Office of Science and Technology Policy, 2019). For instance, the development of environmental citizen science projects gives to a growing number of lay citizens the possibility to take part in real-life research processes, most often by collecting data aiming to survey or monitor biodiversity (Couvet et al., 2008). Beyond their value for scientific research about biodiversity and the environment (Roger et al., 2019), these programs are also expected to promote public engagement towards conservation issues (Turrini et al., 2018), and promote public learning or public understanding of science (Peter et al., 2021). A common assumption is indeed that this participation in the research process may foster citizens' insight of science knowledge regarding biodiversity, and of the principles,

methods and procedures grounding knowledge production through scientific research (Aristeidou and Herodotou, 2020; Kloetzer et al., 2021). Phillips et al. (2018) proposed a framework to evaluate these science learning outcomes of citizen science. They distinguish between two potential objects of learning for participants: (i) the “Content, process and nature of science knowledge” — that is, the “Knowledge of science content and the nature of science” (p. 7): in other words, science as a corpus of formal knowledge (theories, laws, explanations, classifications) and an ensemble of formal methods and principles (hereafter, science content knowledge); (ii) the “Skills of science inquiry” — that is, “Procedural skills such as asking questions, designing studies, collecting, analyzing and interpreting data, experimenting, argumentation, synthesis, technology use, communication, and critical thinking” (p. 7; hereafter, science inquiry knowledge or procedural skills).

* Corresponding author.

E-mail address: baptiste.bedessem@gmail.com (B. Bedessem).

<https://doi.org/10.1016/j.biocon.2022.109807>

Received 1 July 2022; Received in revised form 11 October 2022; Accepted 2 November 2022

0006-3207/© 20XX

Various studies have been conducted which document an increase in science content knowledge after participation in citizen science programs (see Peter et al., 2019 for a review). For instance, participants improve their species identification skills (e.g. Jordan et al., 2011; Deguines et al., 2018), or their ecological knowledge regarding species behaviors (e.g. Druschke and Seltzer, 2012) or local ecosystems (e.g. Leong and Kyle, 2014). More sparsely, some studies also indicate an increase of the understanding of science methods and processes (Haywood, 2016), but this point is still strongly debated (Aristeidou and Herodotou, 2020). However, regarding science inquiry or procedural skills, data are still lacking, even if science inquiry learning is a long lasting promise of citizen science (Trumbull et al., 2005). A reason could be the difficulty to detect empirically this kind of learning (Trumbull et al., 2000). A pending question is then the following: to what extent, and under which conditions, participation in biodiversity citizen science may contribute to enhancing such procedural skills? This paper proposes to tackle this question by using as research material the virtual interactions between participants from a French biodiversity citizen science project — the Photographic Survey of Flower Visitors (the Spipoll, see Deguines et al., 2012) — aiming at monitoring plant-pollinator interactions through time across France. Participants share their observations (pictures and identifications of insects visiting plants, see Methods section) on an online platform allowing them to write comments on the webpages displaying the observations of each other. No particular topic was expected for these comments, and threads were possible. This online interaction platform offers a long-term (8 years) and large-scale (1404 participants, 78,568 comments posted by 402 contributors) dataset to study participant interactions and their changes within a biodiversity citizen science program.

Up to now, citizen scientists' online interactions have been mainly studied in so-called “virtual citizen science” programs which engage citizens in online analysis or screening of data (e.g. Galaxy Zoo, <https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/>). In these last cases indeed, virtual spaces of communication are often designed within the platforms, where participants can comment on others' contributions and exchange ideas or suggestions (Baudry et al., 2021). Some studies have been conducted which analyze the dynamics of these social interactions, mostly in the case of the Zooniverse citizen science platform. Among others, Luczak-Roesch et al. (2014) proposed a notable contribution: by analyzing the content of the “Talk” function of the Galaxy Zoo program, they show that online exchanges are characterized by an increasing use of a specialized vocabulary in astronomy with time. In other words, participants' interactions reflect participants' acquisition of scientific knowledge.

Out of virtual citizen science, the characteristics and impact of online interactions within a biodiversity field citizen science project (the Spipoll, <https://www.spipoll.org/>) were recently studied. Deguines et al. (2018) first show that participants' online interactions allowed within the Spipoll plant pollinator contributory monitoring scheme is associated with the acquisition of insects' identification skills. This acquisition of science content knowledge may be linked to the comments exchanged by participants in the online data sharing platform. Torres et al. (2022) extracted all the comments posted by the participants in Spipoll's platform from 2010 to 2018, and performed a content analysis to classify them depending on their topic. In this paper, we explore these comments in order to tackle the following research question: do the contents of these exchanges reflect a scientific learning on the behalf of the participants? In particular, we investigate the traces of procedural skills learning, making the hypothesis that this form of learning may be detectable through the content of participants' interactions. The paper is organized as follows. First (Section 2) we present our dataset and the qualitative (content analysis) and quantitative (Bayesian multinomial mixed-effects model) methods we used to analyze it. Then, we describe the results we obtained (Section 3), and we finally discuss participants' outcomes of biodiversity citizen science projects and the value

of online devices allowing participants' interactions to assess learning outcomes of citizen science (Section 4).

2. Methods

2.1. Our data: virtual exchanges among participants of a biodiversity contributory science projects

The Spipoll citizen science program (Photographic Survey of Flower Visitors, www.spipoll.org) was created in 2010 by the French Museum of Natural History (MNHN) and a French entomological learned society (OPIE for “Office pour les insectes et leur environnement”). It was designed to survey pollinator assemblages (Deguines et al., 2012; Serret et al., 2019; Torres et al., 2022). Participants apply the following standardized protocol. First, they select a flowering plant species (possibly several individuals of the same species within a ten-meter diameter circle) of their choice and take two pictures: one of the plant's flower itself and one of the surrounding environment. Then, they take pictures of every insect visiting the flowers of the selected plant species during a 20-min period. Second, they identify insects and plants using a dedicated online identification tool. Third, they upload their pictures and associated identification, as well as date, time, and location of observations (hereafter a collection) on the Spipoll's website, where collections are shared among participants. As described in Torres et al. (2022), from the beginning of the program, anyone registered on the Spipoll platform (volunteers — any participants — and experts — professional entomologists or botanists validating the identifications) could comment on each other's collections. The online platform was indeed designed to allow for the viewing of the data collected (i.e. pictures of plants and their visiting insects) by anyone and to allow commenting on these observations. The community manager of the program used this functionality to communicate with participants, mostly about insects' identifications and to remind them of the protocol when necessary. Participants were allowed to modify their identifications according to other participants' or experts' comments. From May 1st, 2010 to April 30th, 2018 (period of the present analysis), >295,000 insect pictures were collected by 1404 participants, most of them without previous entomological or biological training. When registering to the program, participants are indeed asked to auto-evaluate their level of expertise in entomology (“novice”, “occasional”, “amateur”, or “professional”). The analysis of these declarations indicates that only 34 % of them consider themselves as “amateur” and 1 % as “professional” (vs. 17 % as “novice” and 48 % as “occasional”). Among these participants, some of them contribute to online interactions, with a highly variable frequency among participants. While there is an overall positive association between the number of collections made by the participants and the total number of comments they post, it is worth noting that the first remains a poor indicator of the second for many participants (see Fig. 1 in Torres et al., 2022). The thematic content of these comments were analyzed by Torres et al. (2022). The empirical material they analyzed represents a total of 78,500 comments (Spipoll-Vigie Nature 2010–2018). Among these 78,500 comments, they sampled 7850 comments (10 %) by following a stratified sampling strategy, per year and per participant. Then, they descriptively coded the content of each comment included in the sample, in several cycles following Saldaña (2021). This iterative process led to describing 10 comment types within the comments' corpus. These 10 types of comments, which are not mutually exclusive (one comment can belong to different types), are described in detail in Torres et al. (2022). The four most common types comprised admiration comments (a.g concerning aesthetical aspects of the picture, 36.4 %); comments about insects' and flowers' identifications (20.7 %); friendly exchanges (16.7 %); or comments presenting reflections on biology and ecology aspects (17 %). This last, so-labeled “Biology-Ecology” type contains exchanges about natural processes (e.g. “Note that the drone is incapable of feeding itself, it is maintained with porridge by

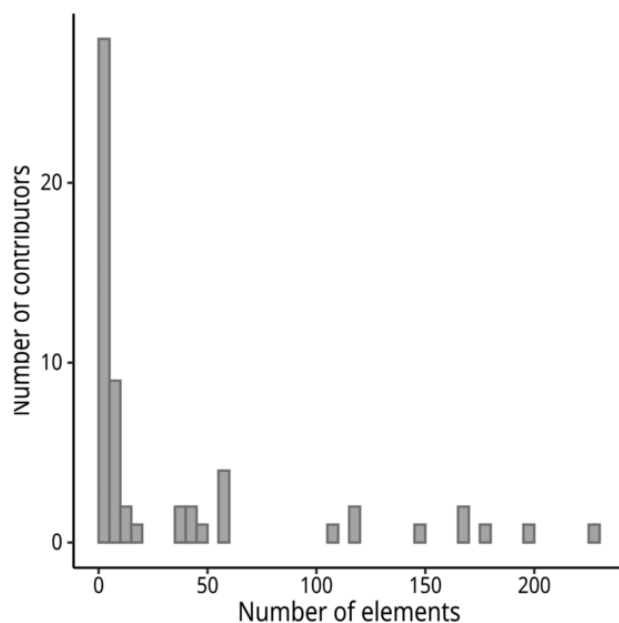


Fig. 1. Distribution of the number of elements authored by contributors in our sample set.

its sisters in the hive”), about plant properties (“*The parsnip is an excellent root, rich in proteins, starches, and pectins. It contains 30mg/100g of C vitamin*”), exchanges regarding species frequency or rarity (“*unusual and interesting species*”), or about pictures’ environmental conditions: (talking about domestic bees) “[...] *I am perplexed by their release in this fairly cool temperature 15-17 degrees, a little bit more under the sun, of course... Some small valleys that they fly over are still frozen around noon (12h)*”. Interestingly, some of the comments in this type also propose hypotheses to explain observed phenomena (“*the Onion probably has a more attractive smell...*”) or propose questions to explore (“*It would be interesting to make a summary of all the ornamental plants which attract the insects which are still active in autumn*”). This last kind of comments are particularly important regarding the goals of this paper — assessing whether or not there is evidence of science learning through participants’ interactions. Indeed, this Biology-Ecology type of comments gathers elements belonging to the two forms of science learning as defined by Phillips et al. (2018). For instance, comments referring to natural processes or plant properties may be linked to the science content knowledge. By contrast, comments which propose hypotheses or questions correspond typically to the science inquiry dimension of science learning. Exploring more thoroughly the content of these Biology-Ecology comments and its changes in time may then give important information regarding the acquisition of science knowledge through participation in biodiversity citizen science. In this paper, we thus focus on the so-labeled “Biology-Ecology” type, which represents 1641 comments. We descriptively coded the thematic content of each comment included in the sample, in several cycles following Saldaña (2021). We found seven categories of content (see Section 3). Each comment may include elements belonging to distinct categories; in those cases, every element from the same comments were classified in the appropriate category. This empirical material represents a total of 2125 elements (Spipoll-Vigie Nature 2010–2017). Hereafter, we refer to those participants which posted at least one comment as contributors.

2.2. Bayesian analysis methodology

To investigate how proportions of these seven categories of elements from the comments changed through the years (i.e. through the age of the Spipoll community), we fitted multinomial mixed-effects Bayesian models. We used R version 4.1.3 (R Core Team, 2022) and the

brms package (Bürkner, 2017), which is a high-level interface to Stan (Carpenter et al., 2017).

We first fitted a model to the complete dataset. The response variable of the model was the category of elements (seven levels), with one of the categories (labeled “Research”, see Section 3) set as the reference level. Indeed, we wanted to assess how the proportion of Research comments changed through the years, relatively to the other six categories (Section 3). Age of the Spipoll community, that is the age of the program in years (integer variable, from 1 to 8, corresponding to years 2010–2017) was included as a fixed effect. The latter was given $N(0,5)$ priors, considered non-informative as this encompasses a larger range of values than those possible for the beta parameters to be estimated (Gelman and Hill, 2007). In our dataset, many participants have authored multiple comments (Fig. 1). To account for this dependence in the data, we included the identification code of participants’ as a random effect on the intercept as well as the slope of the fixed effect age of the Spipoll community. We ran the model with four Markov chains of 5000 iterations (with the 2500 first ones being warm up draws, not used for inference) and a thinning rate of one. Parameters *adapt_delta* and *max_treedepth* were raised to 0.9999 and 90 respectively. Model convergence towards stable parameter estimates were estimated by visual inspection of the trace of the four Markov chains, and by checking R-hat values (all R-hat = 1) and effective sample size (minimum Bulk or Tail effective sample size = 3466). Model fit to the data was evaluated graphically using posterior predictive checks and numerically with cross-validation (all Pareto k diagnostic values < 0.7).

To assess the extent to which the changes in comments’ contents observed in the full dataset may arise from long-term participation, we compared results obtained from two models fitted to two separate datasets: short-term (1 to 4 years) contributors only, and long-term (5 to 8 years) contributors only. Models’ structure was the same as the one described for the full dataset, and we checked convergence and fit to the data similarly. Convergence of the short-term contributors model required setting *adapt_delta* and *max_treedepth* at 0.99999 and 100 respectively.

3. Results

3.1. Content analysis of the biology-ecology comments

Our content analysis led to distinguishing 7 categories. Below we label and describe briefly the different categories we identified.

a-Surprise (28 % of the biology-ecology comments): expression of a surprise in front of a rare or unforeseen event: “what surprised me a lot is the change of insects within a few days’ interval”; “I thought the Forsythias did not attract pollinators”.

b-Perspective (24 %): elements which put into perspective a picture with one’s own experience. Most of the elements in this category refer to other collections (“*Beautiful collection, I just uploaded one made on Rosemary, but much less rich*”) or to individuals’ gardening practices (“*I planted six Cistus last year. I look forward to them growing up*”).

c-Description (16 %): description of natural objects in the pictures, both from those who took them and from other participants who comment on them: “*Queen and worker bees were static*”; “*The first Syrphid Hoverfly has its bodywork broken*”.

d-Context (28 %): elements of contextualization of the pictures, giving information about the location or the physical environment (“*Picture taken at Bourlatier, 1350m high*”); the weather conditions (“*several invincible despite the 9°C and the strong wind*”); or the difficulties of making pictures (“*Not easy to photograph Comfrey’s hosts...It would be needed to lie on the ground*”).

e-Cycle (11 %): references to natural changes due to seasonal cycles: “*butterflies are coming back !*”; “*You already have flowering Asters ! So Autumn is not far away*”.

f-Knowledge (11 %): This category gathers items about biological or ecological knowledge in relation to the commented picture. They generally focus on medicinal properties of observed plants (“*Medicinal plant with purgative properties*”; “*old medicinal plant already used in the Roman Empire*”) or on ecological knowledge about plants or insects (“*Fringed orchid flowers' release their fragrance during the night to attract moths*”). This category may be associated with the science content kind of science learning. However, let us note here that this *Knowledge* category is not strictly equivalent to Phillips et al.' (2018) ‘Science Knowledge’ category.

g-Research (17 %): this last category gathers elements which express different forms of scientific reasoning, including the formulation of hypothesis (“*I guess this is due to the warm weather, the bees from the nests nearby have not been killed by cold*”); the expression of results obtained from an inductive process (“*It's funny: from my almost hundred observations from 2005, these ones love Teasels, lavender and Hollyhock*”); the suggestion of research questions: “*May we measure global warming with the insects collected at this period?*”; “*I asked scientists to do graphs with the average number of taxa/collections/months/years to see if this year was so bad with respect to the other ones*”. This category of comments may be linked to science inquiry knowledge or procedural skills, as discussed in Section 4.

The number of elements (that is, comments or part of comments) vary with years (Table 1) and contributors (Fig. 1). Additionally, there are different patterns of contributing comments along the years among contributors (Fig. 2). Some contributors posted comments only during a few years, whereas others are characterized by a long-term commenting activity. At the extreme ends of the spectrum, 4 contributors posted comments during the 8 years, whereas 30 contributors posted comments only during one year.

Once comments' content has been analyzed and classified, we conducted a quantitative analysis to assess the changes in time of the frequency of elements in each of the 7 categories we figured out.

3.2. Change in comments' content

Fig. 3 shows the proportions of elements from Spipoll's comments belonging to each of the seven categories we defined, from 2010 (year 1) to 2017 (year 8). These observed data indicate an increasing trend of the *Research* category (from 4 % to 24 % of the total number of comments). To assess the statistical significance of this change, we computed a Bayesian multinomial mixed-effects model as described in the Methods section. Results are shown on Fig. 4a,b. The model confirms a clear increase in the proportion of the *Research* elements relatively to the other categories. Beta estimates computed by taking *Research* as the reference category indicate the probability for a comment to have an element belonging to other categories than *Research*. We found that all

Table 1

Number of elements from the Biology-Ecology category, and long-term and short-term contributors sample sets, from 2010 (year 1) to 2017 (year 8). Years are counted from May 1st, n to April 30th, $n + 1$. The two last columns give the total number of elements, and the corresponding proportion with regards to the whole sample (in %).

Year	1	2	3	4	5	6	7	8	Total	%
Number of elements (total)	61	319	366	398	208	287	199	177	2015	100
Number of long-term contributors	7	12	13	15	15	13	14	13	102	56.7
Number of elements (long-term contributors)	58	216	228	256	188	255	185	156	1542	76.5
Number of short-term contributors	2	13	13	17	10	9	7	7	78	43.3
Number of elements (short-term contributors)	3	103	138	142	20	32	14	15	467	23.5

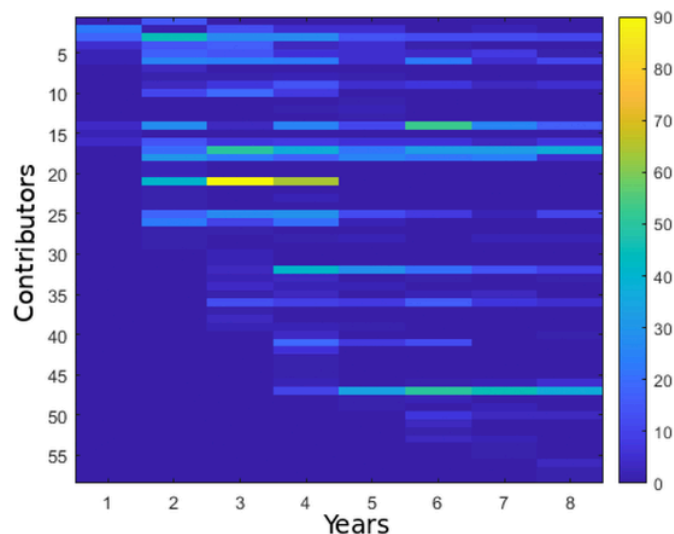


Fig. 2. Number of elements (color scale from blue [0] to yellow [> 80]) posted per contributors as a function of the year. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

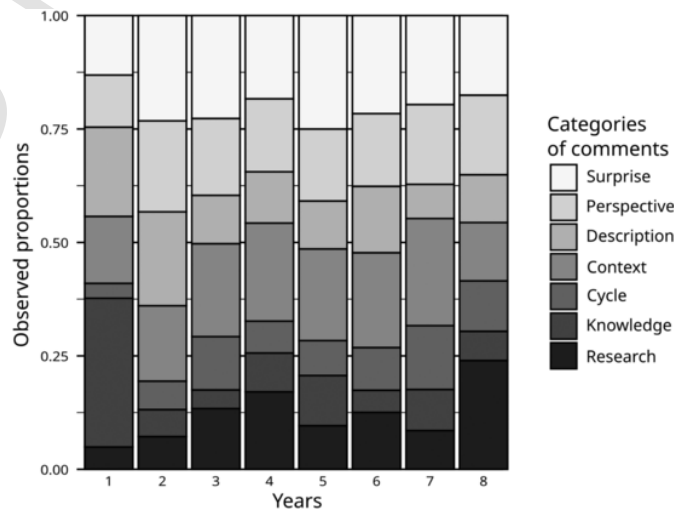


Fig. 3. Observed proportions of the different categories of elements from Spipoll's comment from year 1 (2010) to 8 (2017), from all contributors.

Beta estimates are negative, showing that *Research* is the category that increases most, followed (in decreasing order) by *Cycle*, *Context*, *Surprise*, *Perspective*, *Knowledge* and *Description*. This negative value is statistically significant for three of the six categories. The negative Beta estimate for the *Knowledge* category is not statistically different from 0 due to the specific dynamics of this category, which decreases sharply between year 1 and 2 (Fig. 3). The *Context* category also is not significantly different from the *Research* one at the 5 % threshold. However, raw data show no difference between year 1 and year 8 for this category, and only a small increase from year 1 to year 2 (15 % to 20 %). Notably, the *Cycle* category also shows an increasing trend with years (Fig. 4a). A reasonable conclusion is then that the relative proportions of *Research* and *Cycle* comments significantly increases with time (from 4 % to 24 % for *Research*, and from 3 % to 11 % for *Cycle*, following the observed data).

Given this first result, one can interrogate the drivers of the observed dynamics: is it due to a change of the comments posted only by long-term contributors, or is it a general trend of the whole contributors community? To answer this question, we computed a second Bayesian multinomial mixed-effects model by considering only the data from

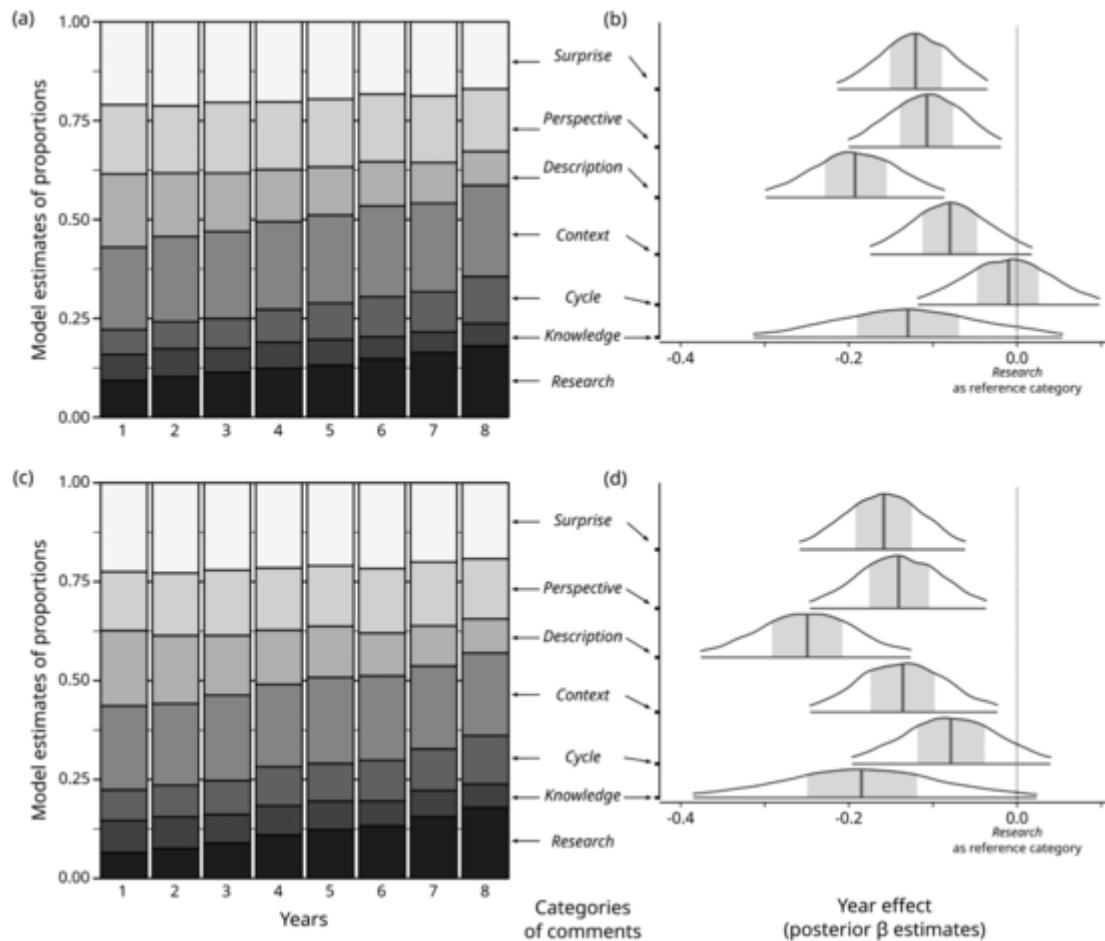


Fig. 4. Results from our Bayesian multinomial mixed-effects model, as applied to the whole (a–b) or long-term contributors only (c–d) datasets. Panels a and c show estimated proportions for each category of elements from Spipoll’s comments, from year 1 (2010) to 8 (2017). Panels c and d display the posterior β estimates of the effect of years as density plots for the different categories of comments, with Research taken as the reference category. In c and d, shaded areas represent 50 % uncertainty intervals, and the outer area represent 95 % intervals.

long-term contributors — that is, contributors who posted comments during 5 to 8 years. Let us note here that it would have been more relevant to directly include the experience of the contributors as an additional explanatory variable in the model; however, the strong correlations between the age of the Spipoll community and the experience of the contributors (Pearson’s $r = 0.82$, p -value $< 10^{-16}$) did not permit inclusion of both explanatory variables in the model (Dormann et al., 2013). Table 1 gives the size of this sub-sample from year 1 to 8, and results of the model are shown on Fig. 4c,d. The model indicates a clear increase of the proportion of comments from the *Research* category (from 2 % to 25 % following the observed data). This observation is confirmed by the model results showing that when taking *Research* as the reference category, all the Beta estimates are negative (with a 95 % certainty except for the *Knowledge* category and the *Cycle* category - for the same reasons as previously exposed regarding the whole dataset results). One can then conclude that long-term contributors tend to change the kind of comments they post over time, with a tendential increase of the frequency of elements belonging to the *Research* and (at a lower extent) *Cycle* categories compared to elements which share ecological knowledge, which describe the pictures or compare them with individual experience, or which manifest a surprise regarding one observation. We performed the same analysis for the short-term contributors (that is, the whole sample minus the long-term contributors, see Table 1 for the sample size), and the Bayesian model does not indicate that elements from the *Research* significantly increase compared to any other categories (see figure in Supplementary Materials S1). Our hypothesis is that the changes observed for the long-term contributors

does not go with a shift in the posted comment for shorter-term contributors. However, this last conclusion should be considered with caution given the small size of the short-term contributors sample.

4. Discussion

Science education is one of the often claimed goals of citizen science (Turrini et al., 2018). But what do participants to biodiversity citizen science projects precisely learn about science? To date, empirical reports of these learning outcomes have given contrasting results, and the sound evaluation of science learning through citizen science still needs much research efforts, as stated by recent literature reviews (Aristeidou and Herodotou, 2020; Kloetzer et al., 2021). These research efforts should consider both the diversity of participation practices (contributory data collection, online citizen science, community-based research) and the diversity of the objects of learning. The simplified and yet useful typology by Phillips et al. (2018) distinguishes between the learning of science content and the learning of science inquiry (or procedural skills). In this paper, we evaluate science inquiry gains in the Spipoll biodiversity citizen science project through the content analysis of the online interactions among participants.

4.1. Long-term participation may foster procedural skills learning

Our main finding is that the proportion of comments proposing new research questions, hypotheses or explanations increases with time. This result is significant for the whole sample of contributors, and even

clearer for long-term contributors (more than four years). However, it was not observed for short-term contributors. This absence of measurable effects might be due to the small size of the sample, it is then difficult to conclude firmly regarding short-term contributors. However, our hypothesis (to take with caution) is that only individuals who contribute to online interactions during a certain time tend to change the content of their comments. In other words, long-term participation in contributory citizen science projects may foster the acquisition of procedural skills. This first result is important since it demonstrates in an original way that long-term engagement in citizen science fosters the acquisition of procedural skills, such as formulating research questions, explanations or hypotheses. Qualitative and quantitative methods, including comments content analysis (see Luczak-Roesch et al., 2014), have already been used in other studies to show the acquisition of science content knowledge (see Peter et al., 2019; Kloetzer et al., 2021 for a review). To our knowledge, our study is the first one that detects learning of science inquiry skills (in the sense of Phillips et al., 2018) in biodiversity citizen science.

Out of this specific result, it then demonstrates the interest of using as research material the online traces of citizen scientists' activities, when available. Consequently, it should encourage the development of such online communication spaces within contributory projects, without restricting them to online citizen science. Out of their known interest to foster the emergence of a stable community of participants (Jackson et al., 2014; Jennett et al., 2016), our findings highlight that these spaces constitute indeed a rich tool to evaluate the impact of citizen science on participants. However, as noted by Torres et al. (2022), this value of long-term online contributory citizen science should not obliterate the importance of maintaining other communication channels: in the Spipoll program, newsletters are regularly sent to the participants, scientific results and scientific information are posted on the project's website, and physical meetings are organized every two years. The impact of online citizen science platforms might strongly depend on these facilitation efforts.

4.2. Long-term participation impacts the attention to natural cycles

A second result worthy of attention pertains to the so-labeled "Cycle" category of comments. This category gathers contributions referring to natural changes due to seasonal cycles. We show that the frequency of comments belonging to this category also grows with time for long-term contributors. This result should be interpreted by considering the fact that pollination is itself a highly seasonal phenomenon. Our hypothesis is then that the confrontation to seasonal changes in insect-flower interactions (through participation to Spipoll protocol and/or through consultations of participants' data) fosters the acquisition of a sense of natural seasonal cycles. Furthermore, it also suggests that long-term participation (and/or contributions to online comments) also increases the attention to seasonal changes in terms of weather as well as flowers and insects types (species or morphospecies, Deguines et al., 2012) and abundances. This result is of major significance since it is well established that the attention to local seasonal cycles is crucial for individuals' perceptions of the effects of climate changes (Fierros-Gonzalez and Alejandro Lopez-Feldman, 2021). Yet, these individuals' perceptions might be a driver for collective action. It has been documented, for instance, that farmers' perceptions of climate change (through an attention to variations in seasonal cycles) is important for them to implement adaptation measures (Meldrum et al., 2018). The same applies to indigenous populations in areas submitted to strong effects of climate change, such as sub-arctic areas (Takakura et al., 2021). In a distinct context, long-term participation to the Spipoll programs (and/or confrontation to others' collections) might allow citizens to directly observe the effects of climate change in their local areas, through a larger attention to and a better knowledge of seasonal cycles, and by extension, to their inter-annual variations (would they be linked to cli-

mate change or not). This idea is confirmed by the presence of numerous comments from this *Cycle* category which insist on the earliness or delay in flowering for some species ("This year they appear fifteen days earlier", "there is no more seasons") or to anomalies in temperature ("this autumn is exceptionally hot"). This indicates that after various years of data collection, participants discuss more frequently (and maybe identify more easily) inter-annual variations in seasonal cycles.

4.3. Perspectives

While our analyses reveal a clear pattern of learning procedural skills in participants to a biodiversity citizen science project, it remains to determine the exact cause(s) and distinguish between the impact of participation in data collection and the impact of contributing to online interactions. Indeed, it is not easy to evaluate which of these activities is the main driver of the measured changes in comments' contents, as there is a strong correlation between the number of posted collections and the number of years of contributions to online interactions (Pearson's $r = 0.54$, $p\text{-value} < 10^{-4}$). More importantly, when considering the 25 % of participants who posted the fewest number of collections, we only found one long-term contributor within that group. It is then impossible, with our present sample, to distinguish between the impact of participation and that of commenting observations in procedural skills acquisition.

Our hypothesis is that these changes are driven by the synergy between these two kinds of activities: participants change their attitudes and skills if they collect data *and* comment on others' contributions. To evaluate this hypothesis, the use of qualitative methods, including interviews with long-term participants and/or contributors, could be useful. Furthermore, we analyzed only a small sample of the whole comments posted in the platform between 2010 and 2018 (10 %). Even if the method used to select them (stratified sampling) ensures a good representativeness in terms of contributors, statistical analysis would be more powerful by using a larger number of comments as empirical material. First, using a larger sample of comments would allow us to distinguish between the impact of participation and that of contributions to online exchanges to foster participants' learning. Second, it might be possible to analyze the time-dependent changes of the comments content more precisely: notably, it would be interesting to determine to what extent the changes are gradual or discontinuous. Third, it would also be interesting to assess the correlations between the changes in comments from the "Research" and "Cycle" categories and the whole set of comments within the Spipoll platform as classified in Torres et al. (2022). It may be asked, indeed, if the growing expression of procedural skills and the growing perception of seasonal changes is associated with specific kinds of comments, such as aesthetic ones, or species identification related-ones. The challenge here would be to assess if there exists a specific road towards procedural skills acquisition, and towards an increase of individual attention to natural cycles.

Our analysis relies on comments that were extracted from the conversation they belonged to. An interesting and complementary analysis would consider the dynamics of discussion, in order to better understand how the different topics emerge through social interactions.

Finally, this work could be coupled with qualitative analysis both to confirm our results and to better grasp the links between the observed learning processes, the sociological characteristics of the participants, as well as their life trajectories.

5. Conclusion

In this work, we propose a qualitative and quantitative analysis of the online comments exchanged among participants of a field biodiversity citizen science project. By using the rich and original material from the Spipoll program's data sharing tool, we were able to track the long-term changes on the thematic contents of the comments exchanged by

participants. We show that long-term participation is associated with the growing expression of scientific procedural skills: formulation of hypothesis and explanation, proposition of new research questions. We also show that long-term participation is associated with the growing attention to natural seasonal cycles. Further research is now necessary to assess the mutual role of Spipoll practices and online contributions as drivers of these changes.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.biocon.2022.109807>.

Uncited references

Chase and Levine, 2018
 Deguines et al., 2020
 Jordan et al., 2012
 Mayer and McPherson Frantz, 2004
 Nisbet et al., 2009
 Schultz, 2002
 Spipoll - Vigie-Nature, 2010 - 2018
 Toomey and Domroese, 2013

Declaration of competing interest

Baptiste Bedessem reports financial support and administrative support were provided by Natural History Museum. Baptiste bedessem reports a relationship with National Museum for Natural History that includes: employment.

Data availability

Data will be made available on request.

Acknowledgements

We warmly thank the Spipoll participants, the Vigie-Nature team and the OPIE for giving us the opportunity to develop this study. We also thank two anonymous reviewers, Grégoire Lois and all the CESCO members, and Mathieu de Flores for useful comments and information. We also thank the ANR project Headwork (ANR-16-CE23-0015).

References

- Aristeidou, M., Herodotou, C., 2020. Online citizen science: a systematic review of effects on learning and scientific literacy. *Citizen Sci. Theory Pract.* 5 (1), 11. <https://doi.org/10.5334/cstp.224>.
- Baudry, J., Tancoigne, E., Strasser, B.-J., 2021. Turning crowds into communities: the collectives of online citizen science. *Soc. Stud. Sci.* 302 (3), 399–424.
- Bürkner, P.-C., 2017. Brms: an R package for bayesian multilevel models using Stan. *J. Stat. Softw.* 80, 1–28.
- Carpenter, B., Gelman, A., Hoffman, M.-D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A., 2017. Stan: a probabilistic programming language. *J. Stat. Softw.* 76, 1–32.
- Chase, S., Levine, A.-S., 2018. Citizen science: exploring the potential of natural resource monitoring programs to influence environmental attitudes and behaviors. *Conserv. Lett.* 11 (2), e12382.
- Couvet, D., Jiguet, F., Julliard, R., Levrel, H., Teysseire, A., 2008. Enhancing citizen contributions to biodiversity science and public policy. *Interdiscip. Sci. Rev.* 33 (1), 95–103.
- Deguines, N., Julliard, R., de Flores, M., Fontaine, C., 2012. The whereabouts of flower visitors: contrasting land-use preferences revealed by a country-wide survey based on citizen science. *Plos One* 7 (9), e45822.
- Deguines, N., de Flores, M., Lois, G., Julliard, R., Fontaine, C., 2018. Fostering close encounters of the entomological kind. *Front. Ecol. Environ.* 16 (4), 202–203.
- Deguines, N., Princé, K., Prévot, A.-C., Fontaine, B., 2020. Assessing the emergence of biodiversity practices in citizen scientists of a backyard butterfly survey. *Sci. Total Environ.* 136842.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, Carl, G., Carré, G., Garcia Marquez, J.R., Gruber, B., Lafourcade, B., Leitao, P.J., Münkermüller, T., McClean, C., Osborne, P.F., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36 (1), 27–46.

- Druschke, C.G., Seltzer, C.E., 2012. Failures of engagement: lessons learned from a citizen science pilot study. *Appl. Environ. Educ. Commun.* 11 (3–4), 178–188.
- European Commission, 2013. IN-DEPTH REPORT: Environmental Citizen Science.
- Fierros-Gonzalez, I., Alejandro Lopez-Feldman, A., 2021. Farmers' perception of climate change: a review of the literature for Latin America. *Front. Environ. Sci.* 9, 672399.
- Gelman, A., Hill, J., 2007. *Data Analysis Using Regression and Multilevel/hierarchical Models*. Cambridge University Press, Cambridge.
- Haywood, B.K., 2016. Beyond data points and research contributions: the personal meaning and value associated with public participation in scientific research. *Int. J. Sci. Educ.* 6 (3), 239–262.
- Jackson, C., Østerlund, C., Crowston, K., Mugar, G., De Vries Hassman, K., 2014. Motivations for sustained participation in citizen science: case studies on the role of talk. In: 17th ACM Conference on Computer Supported Cooperative Work & Social Computing. pp. 1–10.
- Jennett, C., Kloetzer, L., Schneider, D., Iacovides, I., Cox, A., Gold, M., Fuchs, B., Eveleigh, A., Methieu, K., Ajani, Z., 2016. Motivations, learning and creativity in online citizen science. *J. Sci. Commun.* 15 (3), 1–23.
- Jordan, R.C., Gray, S.A., Howe, D.V., Brooks, W.R., Ehrenfeld, J.G., 2011. Knowledge gain and behavioral change in citizen-science programs. *Conserv. Biol.* 25 (6), 1148–1154.
- Jordan, R.C., Ballard, H.L., Phillips, T.B., 2012. Key issues and new approaches for evaluating citizen-science learning outcomes. *Front. Ecol. Environ.* 10 (6), 307–309.
- Kloetzer, L., Lorke, J., Roche, J., Golumbic, Y., Winter, S., Jögeva, A., 2021. Learning in citizen science. In: Vohland, Katrin (Ed.), *The Science of Citizen Science*. Springer, Berlin, pp. 283–308.
- Leong, K.M., Kyle, G.T., 2014. Engaging park stewards through biodiversity discovery: social outcomes of participation in bioblitzes. *Park. Sci.* 31, 106–111.
- Luczak-Roesch, M., Tinati, R., Simperl, E., Van Kleek, M., Shadbolt, N., Simpson, R., 2014. Why won't aliens talk to us? Content and community dynamics in online citizen science. In: Eighth International AAAI Conference on Weblogs and Social Media. pp. 1–10.
- Mayer, F.S., McPherson Frantz, C., 2004. The connectedness to nature scale: a measure of individuals' feeling in community with nature. *J. Environ. Psychol.* 24 (4), 503–515.
- Meldrum, G., Mijatović, D., Rojas, W., Flores, J., Pinto, M., Mamani, G., Condori, E., Hilaquita, D., Gruber, H., Padulosi, S., 2018. Climate change and crop diversity: farmers' perceptions and adaptation on the bolivian altiplano. *Environ. Dev. Sustain.* 20 (2), 703–730.
- Nisbet, E.K., Zelenski, J.M., Murphy, S.A., 2009. The nature relatedness scale: linking individuals' connection with nature to environmental concern and behavior. *Environ. Behav.* 41 (5), 715–740.
- Office of Science and Technology Policy, 2019. Implementation of Federal Prize and Citizen Science Authority: Fiscal Years 2017-2018.
- Peter, M., Diekötter, T., Kremer, K., 2019. Participant outcomes of biodiversity citizen science projects: a systematic literature review. *Sustainability* 11 (10), 2780.
- Peter, M., Diekötter, T., Höfler, T., Kremer, K., 2021. Biodiversity citizen science: outcomes for the participating citizens. *People Nat.* 3 (2), 294–311.
- Phillips, T., Porticella, N., Constan, M., Bonney, R., 2018. A framework for articulating and measuring individual learning outcomes from participation in citizen science. *Citizen Sci. Theory Pract.* 3 (2).
- R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Version 4.1.3.
- Roger, E., Turak, E., Tegart, P., 2019. Adopting citizen science as a tool to enhance monitoring for an environment agency. *Citizen Sci. Theory Pract.* 4 (1).
- Saldaña, J., 2021. *The Coding Manual for Qualitative Researchers*. Sage Publishing, Thousand Oaks.
- Schultz, P., 2002. Inclusion with nature: the psychology of human-nature relations. In: Schmuck, Peter, Schultz, Wesley P. (Eds.), *Psychology of Sustainable Development*. Springer, Berlin, pp. 61–78.
- Serret, H., Deguines, N., Jang, Y., Lois, G., Julliard, R., 2019. Data quality and participant engagement in citizen science: comparing two approaches for monitoring pollinators in France and South Korea. *Citizen Sci. Theory Pract.* 4 (1), 22.
- Spipoll - Vigie-Nature, 2010 - 2018. Base de données du Suivi photographique des insectes pollinisateurs / Photomonitoring of pollinating insects database for France. Office pour les Insectes et leur Environnement - Muséum national d'Histoire naturelle, Paris, France.
- Takakura, H., Fujioka, T., Ignatyeva, V., Tanaka, T., Vinokurova, N., Grigorev, S., Boyakova, S., 2021. Differences in local perceptions about climate and environmental changes among residents in a small Community in Eastern Siberia. *Polar Sci.* 27, 100556.
- Toomey, A.H., Domroese, M.C., 2013. Can citizen science lead to positive conservation attitudes and behaviors? *Hum. Ecol. Rev.* 20, 50–62.
- Torres, A.-C., Bedessem, B., Deguines, N., Fontaine, C., 2022. Online data sharing with virtual social interactions favor scientific and educational successes in a biodiversity citizen science project. *J. Responsible Innov.* <https://doi.org/10.1080/23299460.2021.2019970>.
- Trumbull, D.J., Bonney, R., Bascom, D., Cabral, A., 2000. Thinking scientifically during participation in a citizen-science project. *Sci. Educ.* 84 (2), 265–275.
- Trumbull, D.J., Bonney, R., Grudens-Schuck, N., 2005. Developing materials to promote inquiry: lessons learned. *Sci. Educ.* 89 (6), 879–900.
- Turrini, T., Dörler, D., Richter, A., Heigl, F., Bonn, A., 2018. The threefold potential of environmental citizen science - generating knowledge, creating learning opportunities and enabling civic participation. *Biol. Conserv.* 225 (September), 176–186.