

The project design influences the quality of contributions in an online Citizen Science project

Julia Rittenschober^{a,*}, Sonia Kleindorfer^{a,b,c} and Didone Frigerio^{a,b}

a Konrad Lorenz Research Center, Core Facility for Behavior and Cognition, University of Vienna, Fischerbau 13, 4645 Gruenau im Almtal, Austria

b Department of Behavioral and Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

c College of Science and Engineering, Flinders University, Adelaide 50001, Australia

E-mail: julia.rittenschober@univie.ac.at, sonia.kleindorfer@univie.ac.at, didone.frigerio@univie.ac.at

The quality of scientific data is essential for ensuring reliability and reproducibility and can be affected by several factors. In this study we show how some aspects of project design may impact the quality of the contributions of the volunteers in an online Citizen Science project. The online platform is Zooniverse and the project is NestCams, which is about bird breeding behaviour.

*Austrian Citizen Science Conference 2023, ACSC2023
19 - 21 April 2023
Linz, Austria*

*Speaker

1. Introduction

Digital platforms enable citizen scientists to participate in research worldwide. At Zooniverse, citizen scientists can be involved in data evaluation and analysis. The project team can generate its own workflow including questions and tasks for the uploaded audio, video or image files, as well as information pages and a talk board for discussion [1].

One project NestCams, from Austria, was on Zooniverse from 2019 to 2021. Cameras were mounted near the nests of greylag geese (*Anser anser*) and northern bald ibises (*Geronticus eremita*) and citizen scientists were asked to code video sequences with duration around 20 seconds per sequence. The aim of the main project was the investigation of behavioural patterns contributing to reproductive success in the two study species. We ask for support from volunteers to analyse the behavioural videos. The quantity of data generated shows the utility of the online platform for the project: more than 400 000 classifications were made by 10 000 volunteers [2].

An essential aspect in Citizen Science projects is the reliability of the contributions [3], which can depend on many factors. Some of these factors are related to individual characteristics such as fatigue and attention, but also preferences and attitudes [4]. Some factors can be controlled by the project team, especially factors related to the project design [3,5] and some aspects of motivation, i.e. price games could have an effect [4].

The aim of this study is to investigate how aspects of project design may impact the quality of the contributions in the project NestCams. In line with the main topic “vereinen” (bring together) addressed by the 8. Austrian Citizen Science Conference in Linz, we address hereafter a well-known topic (i.e. data quality) in order to share our experience.

The following variables were considered as possibly influences on the quality of the contributions: complexity of the task (i.e. the video), complexity of the protocol (i.e. the number of questions & answer options), language (i.e. German & English) and additional motivational tools (games with prizes). The inter-coder reliability between the volunteers and the core project team was measured and an increase in reliability was expected.

In the NestCams project, two avian species were monitored. In greylag geese, only the female incubates the eggs and the nest camera captures one nest in view whereby the task of observing the behaviour at the nest is considered to be simple. In the northern bald ibis, both parents can be present at the nest and may also interact with nesting neighbours, making the nest observations and task accomplishment more complex. The data collection protocol, that is, the number of question and answer options, was simplified once over the time of the project, in spring 2020. For the greylag goose, as an example, the protocol initially included 3 questions with a total of 13 answer choices and was then simplified to 2 questions with 8 answer choices. Additionally, during the project period, German was introduced as an alternative language to English with the aim of fostering participation among local, German speaking people.

2. Methods

For the present project, we compared the videos coded by the volunteers with those coded by persons of the core project team.

The core project team consisted of three people who coded around 1,500 videos. The group of the volunteers consisted of 2,620 people who coded the same 1,500 videos several times, in total 12,598 codings. Those videos were selected randomly by Zooniverse.

For the analysis over the course of this study, we used the question about whether the bird in view ‘sits or stands’ at the nest with given answer options. This question and its answer options remained unchanged over the project period, and was a single choice asked for both animal species.

The answers to each video from each participant were exported from Zooniverse into an MS excel spreadsheet. For the coding of each Citizen Scientist entry, the data agreement with the project team member was scored as 0 (no agreement) or 1 (agreement). Based on these values, we calculated the percentage of agreement. Cohen’s Kappa and z-test was used to determine the statistical significance of the difference in agreement [6].

Regarding the complexity of the task and the protocol, all videos that fit into a specific category were chosen, regardless of whether a volunteer for example coded videos from greylag goose and northern bald ibis or not. For the analysis of language, only those volunteers who evaluated in both languages were chosen, since the native language is not known. For prize games, all contributions in the respective period were used.

3. Results

The % agreement was significantly higher in simple tasks (greylag geese) as compared to complex tasks (northern bald ibis) ($z = 41.01$, $p < 0.001$, 89 % vs. 55 %, Cohen’s $\kappa = 0.61$ vs. 0.52) (Figure 1a).

The % agreement was weakly significantly higher in the simplified protocol as compared to the original more complex protocol ($z = -2.16$, $p = 0.03$, 88% vs. 90%, Cohen’s $\kappa = 0.59$ vs. 0.61) (Figure 1b).

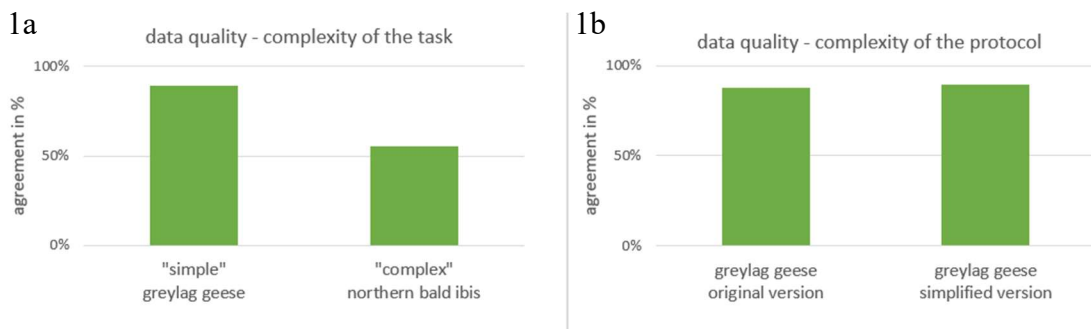


Figure 1a: The mean % agreement between volunteers and project team depending on the complexity of the task (i.e. video)

Figure 1b: The mean % agreement between volunteers and project team depending on the complexity of the protocol (i.e. number questions & answer options)

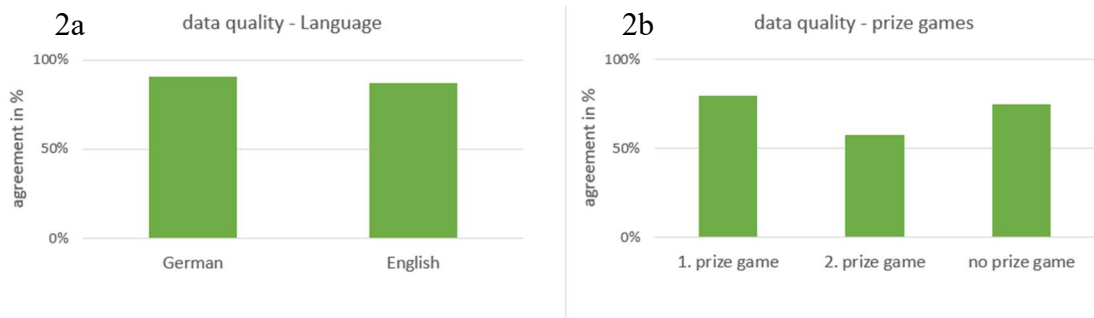


Figure 2a: The mean % agreement between volunteers and project team depending on the language.

Figure 2b: The mean % agreement between volunteers and project team depending prize games.

The % agreement was not significantly higher with German as a chosen Language as compared to English (90 %, $\kappa = 0.73$ vs. 87%, $\kappa = 0.72$) (Figure 2a). The agreement was not consistently different with prize games as compared to without prize games. (without 75%, $\kappa = 0.55$, first prize game 79%, $\kappa = 0.57$; second 58 %, $\kappa = 0.35$) (Figure 2b).

4. Discussion

The interobserver reliability between the project team and volunteers ranged between 55% and 90% agreement. Similar data collected across a shorter project period showed similar patterns of results when focussing on the greylag goose videos [7]. The contributions obtained can be considered reliable, as each single video being coded was viewed by multiple volunteers before being retrieved from the platform.

Both the complexity of the task and protocol had a moderate influence on the quality of the contributions. Therefore, a tip for future projects is to keep the task and the protocol as simple as possible.

Language and prize games had hardly any influence on agreement accuracy. However, one possibility is that the language had an influence on the volunteer's choice of the project or on the quantity of the data.

Also for prize games, the quantity increase was not measured in this study. If this is a research design issue, one should consider that the quality of the data might decrease, as in the second prize game in this study.

Other possible factors influencing the quality of contributions, such as accuracy, knowledge, and experience, may also be indirectly influenced. For example, through good communication, training, and opportunity to ask questions. Over the course of this project, the data quality of the individual users was surveyed at regular intervals and volunteers who performed very well (over 95% agreement) were invited to join the team as experts and were also listed on the project page. In this way, 47,400 videos were evaluated by 22 experts.

In summary, the project design can have a measurable impact on the quality of the data. Minimizing complexity in question and answer options can increase data quality, but there are also external factors and individual differences that need to be further considered when designing data collection by Citizen Scientists.

5. Acknowledgement

We would like to thank the more than 10,000 volunteers with the NestCams project, especially our expert users; listed by username at the following link: <https://www.zooniverse.org/projects/spotteron/nestcams/about/team>. A big thanks goes to the FFG Talente Praktikant pupils from 2018 – 2020 for preparing the videos. We gratefully acknowledge Grant Miller and the team at Zooniverse for their valuable support, Burghard Gniewosz from the Paris Lodron University of Salzburg for his supervision. We thank the Verein der Förderer der Konrad Lorenz Forschungsstelle and the Cumberland Stiftung for ongoing and long-term support.

References

- [1] <https://help.zooniverse.org/getting-started/>, last retrieved 25.08.2023
- [2] <https://www.zooniverse.org/projects/spotteron/nestcams/stats>, last retrieved 25.08.2023
- [3] B. Balázs, P. Mooney, E. Nováková, L. Bastin, & J. Jokar Arsanjani, *Data Quality in Citizen Science*, in *The Science of Citizen Science*. Springer, Cham, 2021 [doi.org/10.1007/978-3-030-58278-4_8]
- [4] G. Hofinger, *Fehler und Unfälle*, in *Psychologie sicheren Handelns in Risikobranchen*, Springer Berlin Heidelberg, 2012
- [5] J. Encarnação, M. A. Teodósio & P. Morais, *Citizen Science and Biological Invasions: A Review*, *Frontiers in Environmental Science*, 8, 303 (2021)
- [6] <https://www.socscistatistics.com/tests/ztest/default.aspx>, last retrieved 29.08.2023
- [7] J. Rittenschober, V. Puehringer-Sturmayer, & D. Frigerio, *The relevance of a digital platform for the citizen science project NestCams*, PoS(ACSC2020) (2021)