PROCEEDINGS OF SCIENCE

Identifying learning dimensions in citizen science projects

Marius Oesterheld*[a], Vincent Schmid-Loertzer,[a] Miriam Calvera-Isabal,[b] Ishari Amarasinghe,[b] Patricia Santos[b] and Yaela N Golumbic[c,d]

[a]*Wissenschaft im Dialog, Charlottenstraße 80, 10117 Berlin, Germany*

[b]*TIDE Research Group, Universitat Pompeu Fabra, Tànger, 122-140, 08018 Barcelona, Spain*

[c]*The MOFET Institute, 13 Shoshana Persitz St., Tel Aviv 61484, Israel*

[d]*The Steinhardt Museum of Natural History, 12 Klausner Street, Tel Aviv-Yafo, Israel*

E-mails: marius.oesterheld@w-i-d.de, miriam.calvera@upf.edu

Citizen science (CS) is attracting increasing interest and attention from multiple sectors of society. Educational impacts of participation, such as the development of scientific skills or increased awareness about biodiversity and conservation, are one of the most widely discussed aspects of CS. Whereas most existing studies investigate perceived or observed learning gains of citizen scientists, this paper takes an alternative perspective by examining learning-related aspects in textual self-representations of CS projects—namely project descriptions posted online. Project descriptions were chosen as objects of analysis both because they can easily be accessed and collected through automated web crawling, and because, as key elements of a CS project's online presence, they play an important role in recruiting volunteers. We have thus conducted a qualitative content analysis of 94 project descriptions with the goal of examining what they tell us about learning dimensions associated with participation in these CS projects. Building on the model of individual learning outcomes developed by Phillips et al. in 2018 [1] as a theoretical framework, our analysis shows that some learning dimensions (such as data collection or using technology), are very prominently discussed in the project descriptions we studied, while others (e.g. experimenting, study design, community action) are clearly underrepresented. In other words, the project descriptions analyzed only partially reflect the educational potential of participation in CS. In the discussion section of this paper, we suggest possible explanations and ways in which this issue could be addressed on the level of both project design and project communication.

Engaging Citizen Science Conference 2022 (CitSci2022)
25-26 April 2022
Aarhus University, Denmark

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(CitSci2022)070

1. Introduction

While the main objective of citizen science (CS) projects is generally to answer a scientific question, CS projects also have many important educational benefits for participants [2][3]. In fact, surveys show that the wish to acquire new skills and knowledge is one of the main reasons for people to join CS projects [4][5]. Since learning opportunities are an important motivational factor for participation in CS, and project descriptions posted on web platforms play a key role in attracting volunteers, we decided to examine what these texts can tell us about the educational potential of CS projects. Aside from their relevance to volunteer recruitment, a major practical advantage of studying project descriptions is that they can easily be accessed and collected on a large scale through automated web crawling.

The research question we wanted to answer through this study was "which dimensions of learning are the most prominent in CS project descriptions?".

Fig. 1 - Classification of individual learning outcomes proposed by Phillips et al. (2018)

As a theoretical framework for our classification of learning dimensions we chose the "framework for articulating and measuring individual learning outcomes from participation in citizen science" [1] that was developed in 2018 by Tina Phillips et al. to capture the informal educational potential of CS. Their model suggests six main categories of learning outcomes in CS projects: Interest; Self-Efficacy; Motivation; Content, Process and Nature of Science Knowledge; Skills of Science Inquiry; and Behavior and Stewardship (cf. Fig. 1).

This paper describes our process of qualitative content analysis (cf. Fig. 3), outlines our results and discusses the implications for CS practitioners and theorists. First, we introduce the project setting and explain points of entry for our analysis (this section). Second, we describe the process and methodology for our analysis. In a third section, we summarize and discuss our results, before ending with a brief conclusion and outlook.

2. Methods

The work presented here is part of CS Track (https://cstrack.eu), a Horizon 2020 project whose objective is to broaden our knowledge of CS by exploring its formats, demographics, impact and potential. The results will be best practice examples and policy recommendations approached from different perspectives such as learning, sustainability or citizens' motivation. One of the cornerstones of the research undertaken by CS Track is a database currently containing

PoS(CitSci2022)070

information on 4947 CS projects that was automatically extracted from more than 56 online platforms using a web crawler. The criteria applied to select websites from which to extract CS projects information were: 1) the websites must contain European CS projects or projects that citizens can participate in online, 2) project information must be available in multiple languages, and 3) data extraction must be allowed. For the study at hand we randomly selected 94 English-language project descriptions from this CS Track database and conducted a qualitative content analysis to identify different dimensions of learning. The categorization proposed by Phillips et al. (2018) [1] was chosen as a theoretical and methodological starting point for this study because it has become widely used, referenced and adapted—as evidenced by 113 citations in Google Scholar (as of 4 April 2022). Moreover, Phillips et al.'s model is broader and more multifaceted than many other recent approaches.

Building on their categorization of learning outcomes, we conducted a structuring qualitative content analysis as described by Philipp Mayring [6][7] on the 94 project descriptions selected. During this process we adapted and adjusted the coding scheme to our source corpus by using two approaches combining inductive and deductive steps: First, we applied Phillips and colleagues' classification scheme to our body of material. Second, where accumulations of relevant phrases suggested potential for additional structuring, we created new categories to accommodate the material that did not fit into the original system. In order to adequately reflect the entire range of information contained in our text corpus, we thus decided to include one additional learning dimension —Attitude Change—and two aspects related to the deliberate design of learning opportunities for participants: i) Training and Didactic Materials (provided by the project), and ii) Access to Project Results. At the same time, we chose to exclude the category of Motivation proposed by Phillips et al. [1] because we were only able to identify one text snippet that unambiguously mentions motivation as a result of participation (cf. Fig. 2).

Fig. 2 - Modified classification of individual learning outcomes created for this study

We then manually assigned phrases, sentences and short paragraphs to the resulting eight main categories (six learning dimensions and two categories relating to deliberate design of learning opportunities) and 21 subcategories. In order to reduce bias, coding of all project descriptions was performed independently by two members of the research team. Intercoder reliability was examined and the rate of agreement was found to be over 90% for six of the eight main categories (i.e. Skills of Science Inquiry, Self-Efficacy, Interest, Attitude Change, Training and Didactic Materials, Access to Project Results)—and between 70 and 80% for the remaining two categories (i.e. Content, Process, and Nature of Science Knowledge and Behavior and Stewardship). As a final step, distinctive and frequently occurring keywords were extracted from these text snippets, which can be used in future studies to train Natural Language

PoS(CitSci2022)070

Processing (NLP) algorithms. This allows us to explore the option of automatically coding all project descriptions stored in the CS Track database.

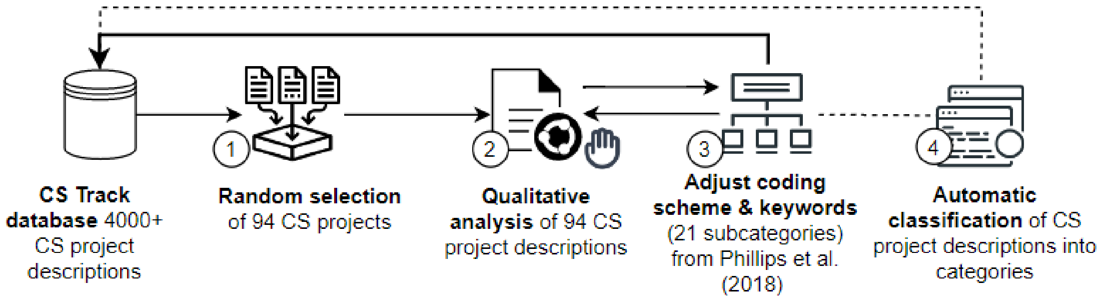

Fig. 3 - Process followed

3. Findings and discussion

As described, our study aimed to identify which dimensions of learning are the most prominent in CS project descriptions by conducting a qualitative content analysis of a random sample of 94 such texts.

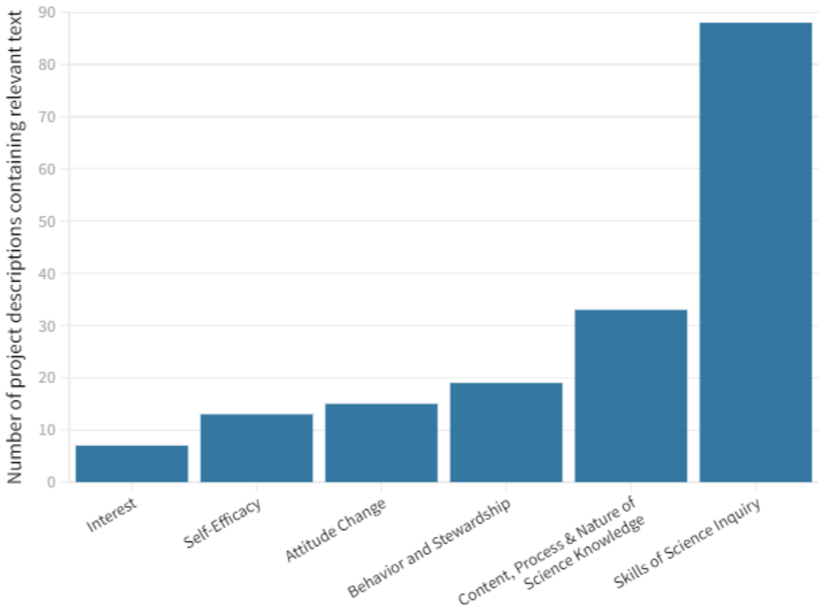

Fig. 4 - Distribution of learning dimensions across projects (main categories only)

Our results show that most project descriptions focus strongly on science-related learning dimensions (such as data collection and analysis skills or knowledge gain) while disregarding other personal or interpersonal benefits, such as self-efficacy, attitude or behavioral changes, political engagement, etc. (cf. Fig 4).

This focus on science-related learning seems to be a common bias in CS, as several publications have pointed out [8][9]—a bias that may run counter to volunteers' actual motivations and expectations. We should therefore attempt to broaden our understanding of learning to reflect all the benefits CS can have on both the individual and the societal level. This broader understanding needs to find expression in quality standards for CS, in projects' self-presentation, evaluation practices, and in deliberate design of learning opportunities for volunteers. For instance, project coordinators could ask themselves which non-scientific activities volunteers could become involved in (e.g., outreach and PR, internal communication, etc.).

PoS(CitSci2022)070

Another interesting finding is the fact that references to Content, Process, and Nature of Science Knowledge play a minor role compared to statements relating to Skills of Science Inquiry, which clearly dominate most project descriptions. That being said, Content Knowledge features more prominently than the other two knowledge types, which is in line with the results of a recent review study [10]. Within the category of scientific skills, 'Data Collection and/or Submission' and 'Using Technology' far outstrip all other subcategories with 'Data Analysis and/or Interpretation' a distant third (cf. Fig. 5). This implies that, at least judging from their self-descriptions, most of the projects we analyzed do not involve participants in anything related to study design, synthesis, communication of research results, etc. In other words, to borrow Muki Haklay's typology of levels of participation in citizen science projects [11], around 88% of the projects represented in our sample seem to fall into categories 1 and 2—Crowdsourcing and Distributed Intelligence. Or, to use the terminology coined by Bonney et al. in 2009 [12], these projects are contributory, rather than collaborative or co-created.

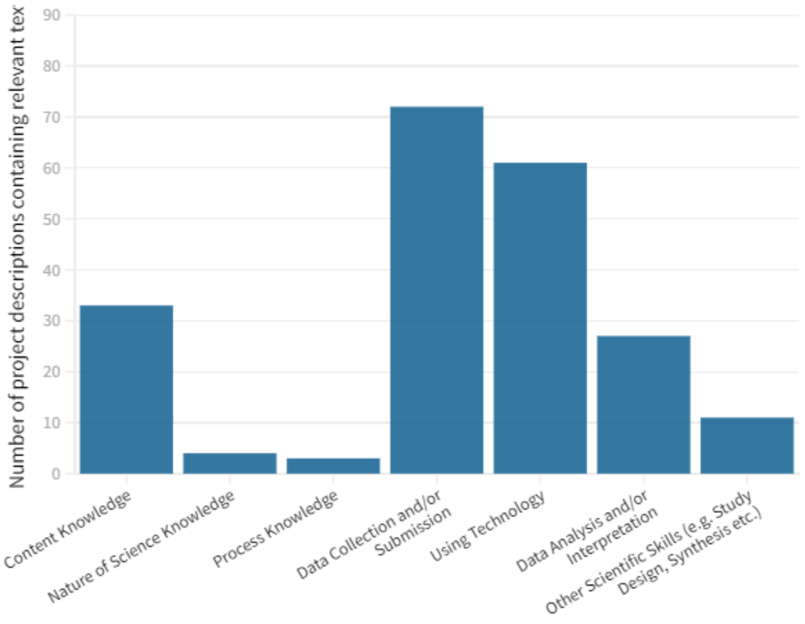

Fig. 5 - Distribution of learning dimensions across projects (selected subcategories)

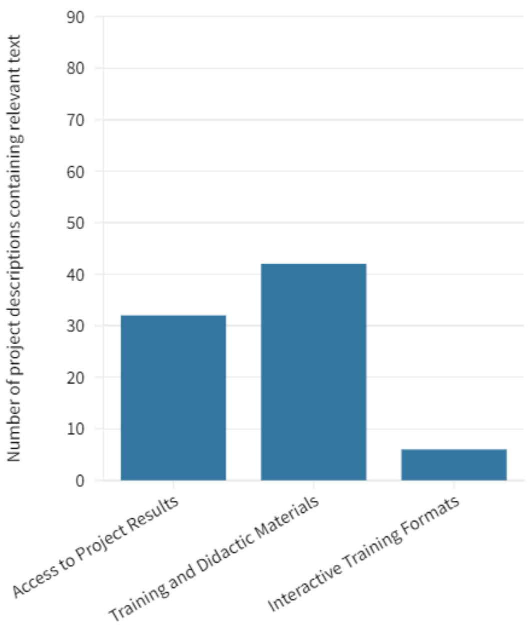

Fig. 6 - Other aspects related to learning opportunities

Of 42 project descriptions which contain information on training and didactic materials offered to participants, only 6 (14,3%) mention interactive training formats (e.g. workshops, webinars, Q&A sessions etc., cf. Fig. 6). This is unfortunate, since recent publications argue that communication and social interaction within the project is key to both learning outcomes and long-term participation [13][14][15][16]. One step in the right direction might be to establish interactive volunteer training as a standard criterion in grant applications, project evaluations, etc.

Overall, we learned in the course of this study that the quality of project descriptions varies considerably. Some are extremely short

PoS(CitSci2022)070

and contain hardly any information on the concrete tasks to be completed by the citizen scientists, others go into great detail regarding the project's scientific background while paying very little attention to societal impact. Ultimately, many project descriptions do not provide a convincing answer to the questions "what will volunteers get out of participating in this project?" or "how will they benefit?". It may therefore be advisable to define quality standards for CS project descriptions and produce guidelines for project coordinators.

4. Conclusion and outlook

Our study revealed a very uneven representation of learning dimensions within CS project descriptions. This result suggests that project initiators and coordinators either do not devote enough attention and resources to creating the broadest possible range of learning opportunities for their volunteers or do not communicate the educational potential of their project clearly enough in their project descriptions. We therefore believe it would be helpful to design guidelines or templates for project descriptions, encourage project initiators to broaden their understanding of learning, and support them in creating interactive training formats for participants. At the same time, it is important to keep in mind that CS projects at the intersection of research, science communication, and non-formal education often have too little funding to fulfil all three tasks with equal success.

It should also be noted that project descriptions do not always provide a comprehensive, detailed picture of what is happening in the respective projects. Moreover, this study was not aimed at capturing how citizen scientists experience the project and what their perceived learning gains are. In order to find answers to these questions, CS Track has conducted an online survey with more than 1000 participants, 610 of whom described themselves as "citizen scientists" (as opposed to project coordinators or professional scientists). The resulting data is now being triangulated with the analysis presented in this paper.

Furthermore, the relatively modest sample size of this study does not support ambitious generalizations. In order to extrapolate the results obtained in this data sample to the whole CS Track database, we are currently exploring the option of applying automated data classification methods.

Acknowledgements

The research presented here was conducted within the framework of CS Track, a project funded by the European Union’s Horizon 2020 research and innovation programme (Science with and for Society) under grant agreement No. 872522. It profited immensely from the kind support of Tina Phillips and her colleagues, who, in addition to providing the general theoretical basis for this study, agreed to share their team’s research data and parts of their coded dataset with us.

PoS(CitSci2022)070

References

[1] T. Phillips, N. Porticella, M. Constas, & R. Bonney, A Framework for Articulating and Measuring Individual Learning Outcomes from Participation in Citizen Science, *Citizen Science: Theory and Practice* **3**, 2, 3 (2018) 1-19. [http://doi.org/10.5334/cstp.126].

[2] R. Bonney, T. B. Phillips, H. L. Ballard, & J. W. Enck, Can citizen science enhance public understanding of science? *Public Understanding of Science* **25**, 1 (2016) 2-16. [http://doi.org/10.1177/0963662515607406].

[3] T. Turrini, D. Dörler, A. Richter, F. Heigl, & A. Bonn, The threefold potential of environmental citizen science—Generating knowledge, creating learning opportunities and enabling civic participation, *Biological Conservation* **225** (2018) 176–186. [http://doi.org/10.1016/j.biocon.2018.03.024].

[4] M.A. Vasiliades, A. Ch. Hadjichambis, D. Paraskeva-Hadjichambi, A. Adamou, & Y. Georgiou, A Systematic Literature Review on the Participation Aspects of Environmental and Nature-Based Citizen Science Initiatives, *Sustainability* **13**, 13 (2021) 7457. [http://doi.org/10.3390/su13137457].

[5] W. Ganzevoort, & R. J. van den Born, Counting bees: Learning outcomes from participating in the Dutch national bee survey, *Sustainability* **13**, 9 (2021) 4703. [https://doi.org/10.3390/su13094703].

[6] P. Mayring, *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim/Basel: Beltz (2010).

[7] P. Mayring, *Qualitative content analysis: theoretical foundation, basic procedures and software solution.* Klagenfurt (2014). [http://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173].

[8] S. Carson, J. Rock, & J. Smith, Sediments and Seashores—A Case Study of Local Citizen Science Contributing to Student Learning and Environmental Citizenship, *Frontiers in Education* **6** (2021). [http://doi.org/10.3389/feduc.2021.674883].

[9] A. J. Roche, L. N. Rickard, K. Huguenard, & P. Spicer, Who's Tapped Out and What's on Tap? Tapping Into Engagement Within a Place-Based Citizen Science Effort, *Society & Natural Resources* **35**, 4 (2022) 1–17. [http://doi.org/10.1080/08941920.2022.2056668].

[10] M. Peter, T. Diekötter, & K. Kremer, (2019) Participant Outcomes of Biodiversity Citizen Science Projects: A Systematic Literature Review, *Sustainability* **11**, 10 (2019) 2780. [http://doi.org/10.3390/su11102780].

[11] M. Haklay, Citizen Science and Volunteered Geographic Information: Overview and Typology of Pariticipation, in: Sui, D., Elwood, S. & Goodchild, M. (eds.) *Crowdsourcing Geographic Knowledge*: 105–22. Dordrecht: Springer Netherlands (2013). [http://doi.org/10.1007/978-94-4587-2_7].

[12] R. Bonney, H. Ballard, R. Jordan, E. McCallie, T. Phillips, J. Shirk, & C. Wilderman, *Public participation in scientific research: defining the field and assessing its potential for informal science education. A CAISE Inquiry Group Report*. Washington, D.C.: Center for Advancement of Informal Science Education (CAISE) (2009).

[13] A. Torres, B. Bedessem, N. Deguines & C. Fontaine, Online data sharing with virtual social interactions favor scientific and educational successes in a biodiversity citizen science project. *Journal of Responsible Innovation* **9**, 2 (2022) 1–19. [http://doi.org/10.1080/23299460.2021.2019970].

[14] N. Deguines, M. de Flores, G. Loïs, R. Julliard & C. Fontaine, Fostering Close Encounters of the Entomological Kind, *Frontiers in Ecology and the Environment* **16**, 4 (2018) 202–203. [http://doi.org/10.1002/fee.1795].

PoS(CitSci2022)070

[15] C. Jennett, L. Kloetzer, D. Schneider, I. Iacovides, A. Cox, M. Gold, B. Fuchs et al., Motivations, Learning and Creativity in Online Citizen Science, *Journal of Science Communication* **15**, 3: A05 (2016). [http://doi.org/10.22323/2.15030205].

[16] M. Luczak-Rösch, R. Tinati, E. Simperl, M. Van Kleek, N. Shadbolt & R. Simpson, Why Won't Aliens Talk to US? Content and Community Dynamics in Online Citizen Science, *Proceedings of the International AAAI Conference on Web and Social Media* **8**, 1 (2014) 315-324 . [http://eprints.soton.ac.uk/363523].

PoS(CitSci2022)070