

Chapter 25

Evaluation in Citizen Science: The Art of Tracing a Moving Target



Teresa Schaefer, Barbara Kieslinger, Miriam Brandt,
and Vanessa van den Bogaert

Abstract Evaluation is a core management instrument and part of many scientific projects. Evaluation can be approached from several different angles, with distinct objectives in mind. In any project, we can evaluate the project process and the scientific outcomes, but with citizen science this does not go far enough. We need to additionally evaluate the effects of projects on the participants themselves and on society at large. While citizen science itself is still in evolution, we should aim to capture and understand the multiple traces it leaves in its direct and broader environment. Considering that projects often have limited resources for evaluation, we need to bundle existing knowledge and experiences on how to best assess citizen science initiatives and continually learn from this assessment. What should we concentrate on when we evaluate citizen science projects and programmes? What are current practices and what are we lacking? Are we really targeting the most relevant aspects of citizen science with our current evaluation approaches?

Keywords Impact assessment · Scientific literacy · Logic model · Latent class analysis · Experience sampling method

T. Schaefer (✉) · B. Kieslinger
ZSI – Centre for Social Innovation, Vienna, Austria
e-mail: schaefer@zsi.at

M. Brandt
Leibniz-Institut für Zoo- und Wildtierforschung (IZW), Berlin, Germany

V. van den Bogaert
Museum Naturkunde Berlin, Leibniz-Institut für Evolutions- und Biodiversitätsforschung (MfN), Berlin, Germany

Institute of Educational Science, Department of Learning and Instruction, Ruhr-University Bochum, Bochum, Germany

Introduction

Evaluation comprises a systematic assessment of the operation and/or the outcomes of an activity or programme, against a set of explicit or implicit standards and criteria (Weiss 1998). Generally, we distinguish between *formative* and *summative* evaluation, where the former is considered process-based evaluation, while the latter is more outcome-oriented. Whereas outcome-based evaluation is concerned with assessing the overall goals of the activities or programmes and the benefits to the participants, process-based evaluation identifies the activities' or programmes' strengths and weaknesses. For some academics, evaluation refers foremost to assuring quality during the scientific process; for others the term is closely related to impact assessment, providing evidence for change triggered by the intervention.

A widely accepted model for defining project success is the *logic model of evaluation* (Örtengren 2004). Although coming mainly from developmental programme design and evaluation, the logic model has been widely adopted and used for evaluating scientific programmes and technology deployment programmes. However, it can also be applied to systematic analysis, implementation, monitoring, and evaluation of development and intervention projects of various kinds; and it has already been applied to citizen science projects.

The logic model provides a structured approach for project design and evaluation as it systematically relates project inputs, activities, outputs, outcomes, and impacts (Fig. 25.1). A variety of definitions and interpretations of these terms are in use. For the purposes of this chapter, we refer to inputs as the resources available to a project and the activities as what is done with those resources. Activities, in turn, deliver products or services – the outputs (e.g. data points collected, workshops conducted). Outcomes are the effects of the outputs on the target group. Impacts are long-term changes brought about on a societal level; they constitute the progress made towards high-level goals.

In the evaluation of citizen science projects, inputs, activities, and outputs are usually easy to measure with quantitative indicators that show the success, or not, of project management. In contrast, recording outcomes requires dedicated effort, and, even then, it may be difficult to causally attribute measured changes as an effect of the project; this is the focus of our chapter. When the intended outcome is a change in people's lives, it is therefore important to include the perceptions and experiences of the intended beneficiaries. For example, if a citizen scientist changes their behaviour and converts to a more sustainable lifestyle, this could be due to

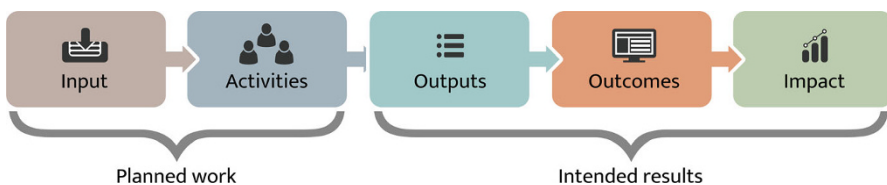


Fig. 25.1 The logic model of evaluation

knowledge and attitudes acquired through participation in a project, but it could also be because their new partner is an environmental activist. Since impact describes a much broader effect than outcome, measuring impacts is even more of a challenge. Broader impacts include primary and secondary long-term effects and aspects of sustainability, both of which have a long-term horizon. Due to its methodological complexity, *impact assessment* requires large amounts of resources, and, even if change is detected on a societal level, causal attribution of this change to one specific citizen science project may often be impossible. This applies to both directions of causal attribution: a project may cause multiple effects, and an observed effect (such as a societal change) usually has not one but many different causes. Due to these difficulties in causal attribution, outcome and impact assessments often include qualitative assessments and case studies (Goertz and Mahoney 2012).

In this chapter we discuss approaches towards outcome and impact-oriented evaluation in citizen science, showcased by concrete examples that depict the variety of practice, and reflect on current challenges as well as new developments in the field.

Historical Development of Evaluation

Evaluation in research projects and programmes has historically been associated with the research output itself, e.g. the validity of the collected data and the resulting scientific evidence. In some scientific disciplines, such as the natural sciences, this is still the prevailing approach, mostly validated via a disciplinary peer review system. However, research policy has also started to value the economic and social importance of research, defined by its mission orientation, alongside scientific quality. The use of scientific indicators in research evaluation can be traced back to the 1960s and 1970s (Leydesdorff 2005). Today's research policy tends to refer to output assessment, measuring not only research quality but also its broader impact or use. Interest in research impact started in the early 1990s, in the UK, which is often considered a leader in research evaluation (Williams and Grant 2018).

This shift towards societal impact assessment of research falls in line with the general historical development of redefining the relationship between science and society, which is observable on many levels, including the increasing institutionalisation of public engagement, the development of practices of technology assessments, and the wide support for the concept of responsible research and innovation (RRI) (Wickson and Carew 2014). Citizen science falls within the principles of RRI and, at the same time, has strong resonance with the characteristics of transdisciplinary research. When dealing with evaluation of citizen science, we have to consider its socioecological relevance, its multi-stakeholder engagement, and its societal embeddedness (also core elements of RRI evaluation).

Contemporary Evaluation of Citizen Science

Evaluation in citizen science today refers to the assessment of the value of its different outcomes and of its processes. It should be understood as a learning process that supports self-reflection and adaptive management, while also helping to understand which effects citizen science initiatives have on science, involved citizens, and socioecological systems. In the following paragraphs, we will reflect on the applied indicators for evaluation, as well as the methods used for evaluating citizen science initiatives. We will also show how different project goals and contexts influence the applied evaluation strategy via a number of case studies.

What Is Currently Evaluated in Citizen Science Projects?

A comprehensive collection of indicators for the evaluation of citizen science initiatives can be found in the *citizen science evaluation framework* (Kieslinger et al. 2018, see Fig. 25.2). This framework suggests indicators for three dimensions of participatory scientific processes: (1) scientific aspects, (2) participants, and (3) socioecological/economic systems.

For each of these dimensions, the framework suggests process-based and outcome-based evaluation: *process and feasibility* collects formative input for an adaptive project design and management; *outcome and impact* brings evidence of a project’s benefits to its participants and their surrounding contexts and shows how much an intervention’s impact contributes to the project’s expected and possibly unintended goals.

The authors of this framework suggest that both types of evaluation, process-based and outcome-based, are crucial for evaluating citizen science projects. Taking a critical look at the project design and continuous progress contributes to the successful implementation of citizen science missions. Impact assessment is

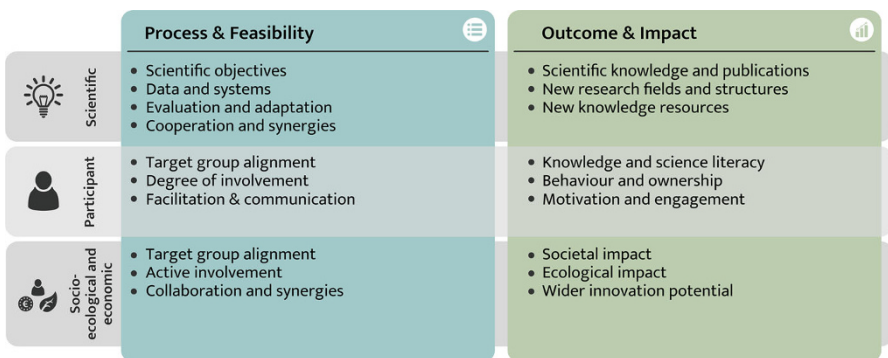


Fig. 25.2 Citizen science evaluation framework, by Kieslinger et al. 2018

increasingly being requested by science policymakers and their funding agencies and enriches our understanding of the value of citizen science.

Within the three dimensions, the citizen science evaluation framework suggests a prioritisation of indicators, adapted to the project context and specific objectives. Projects are not expected to cover all aspects of the framework equally. For example, a co-constructed project, which collaborates with citizens from the onset of the scientific process, might have a clearly defined objective of solving a pressing societal issue. Thus, the scientific outcome in terms of academic publications might be less of a concern. On the other hand, a contributory project, which has been designed with the intention of increasing science knowledge and literacy amongst the participants, should focus its evaluation on the measurable effects of the engagement on the participants.

In the following sections, we will show that evaluation can address all three dimensions suggested by Kieslinger et al. (2018).

Scientific Dimension

Scientific outputs are by far the most important outcome of those citizen science projects that understand citizen science as a research method for new knowledge. These projects also generate the highest number of scientific publications (Kasperowski et al. 2017). In quantitative terms, the largest scientific outputs can be found in the fields of ornithology, astronomy, meteorology, and microbiology, especially in citizen science projects that have developed digital platforms for volunteer contributions, such as Galaxy Zoo, Foldit, etc. (Kullenberg and Kasperowski 2016). On the other side of the spectrum, there are a number of citizen science projects, which do not yet have any output in the form of scientific publications, but it remains unclear how many of these there are (Kullenberg and Kasperowski 2016; Follett and Strezov 2015).

Academic publishing is, however, only one possible metric amongst others. The lack of peer-reviewed publications, especially in the humanities, might simply reflect the fact that many citizen science projects have objectives other than scientific publications. There are projects that prefer to publish their outcomes in societal publications, such as newspaper articles, television, or social media, reaching out to a wider audience and influencing local policies. Others have the transfer of knowledge and the raising of awareness on specific topics as their main goal, which will be examined more closely in the following two dimensions.

What we can also observe is a rising interest in citizen science as a study object and an increasing number of articles that specifically focus on methodological issues in citizen science (Follett and Strezov 2015). Other outcomes in the scientific dimension are more trustful relationships between members of society and the scientific community (e.g. Suomela 2014) and an enhanced capacity for the joint analysis of scientific findings (Bonn et al. 2016).

Participants

A large proportion of citizen science projects still do not evaluate the outcomes for individual participants (Phillips et al. 2018). Amongst those who investigate their projects' impact on individual citizens, the most common outcome documented so far is the one of learning new content knowledge.

Gaining knowledge on scientific subjects was by far the most reported type of learning in citizen science projects (Stepenuck and Green 2015). This ranges from research in climate change (Groulx et al. 2017) to public data collection projects in the field of biodiversity (Bonney et al. 2016) to a number of cyberscience projects, e.g. in the Zooniverse, where even in the absence of a direct educational motivation in the project design, participants learned about specific scientific topics (Masters et al. 2016).

Alongside these proven effects of increased content knowledge in citizen science engagement are much broader learning outcomes as well (Phillips et al. 2018). Citizen science has contributed to learning about the processes of scientific inquiry and to gaining a deeper understanding of scientific outcomes (Bonney et al. 2016), although some studies have questioned this learning outcome (Crall et al. 2012). Recently, citizen science is also being discussed in respect to its potential impact on science capital (Edwards et al. 2018).

Learning in this context is often informal or incidental; picking things up, rather than planning to learn or perceiving an activity as learning. Repeating tasks trains volunteers in specific project skills, so they gain confidence in their contribution and start to take more responsibility, for example, in supporting other learners. The community as a place to develop and exchange is a key aspect in this process, where individuals acquire skills to accomplish project tasks, scientific literacy, and on-topic extra learning through external resources (Jennett et al. 2016).

Overall, simple and visible learning outcomes that are easy to assess (i.e. content learning) are reviewed most frequently in the literature, whereas the more complex and multifaceted aspects of individual and collective learning are rarely evaluated in a systematic way (Bela et al. 2016). We also find a call for intentional learning designs, such as inquiry-based learning, in order for real learning effects to take place (Trautmann et al. 2012).

In contrast, the assessments of transformative effects of learning, such as changes in behaviour, awareness, and stewardship, are often based on assumptions and are rarely evaluated in a transparent way by projects (Bela et al. 2016; Phillips et al. 2018). Relatively few studies refer to outcomes such as a sense of empowerment, a feeling of contributing to science, or insight into one's values and interests (Groulx et al. 2017). In one example on gas drilling, Zerbe and Wilderman (2010) show that many citizens seem to care more about their residential environment than before and are also more responsive as they learn how to measure the contamination around them. Participants' involvement in citizen science proved to influence their ecological perceptions and sense of place, as it increased their understanding of the connections existing between science, place, ecosystem, and the impacts of one's

actions on the environment (Ballard et al. 2017). Also, changing attitudes towards more environmentally sustainable resource management could be observed amongst environmental citizen scientists. In one of the case studies below we exemplify how a citizen science project on air quality contributed to stewardship and citizen activism (Schaefer et al. 2020).

A good overview of learning outcomes from participation in citizen science has recently been published by Phillips et al. (2018), including a framework for evaluation that includes six types of learning outcomes, ranging from content knowledge to self-efficacy and behavioural change.

Socioecological and Economic Systems

Next to the scientific and the participants' perspective, the citizen science evaluation framework (Kieslinger et al. 2018) suggests evaluation should look at the wider social, ecological, economic, and political contexts in which projects are embedded. Considering the socioecological and economic systems is especially relevant for citizen science projects that are initiated by local communities, originating outside of academia. They involve scientists and experts in order to provide evidence in support of campaigns and political decisions regarding issues like pollution, health hazards, and species conservation. These projects do not necessarily strive for purely scientific outcomes but rather aim for transformative change and an impact on the dominate socioecological systems.

Only a few scholars address the evaluation of impacts on socioecological and economic systems, and the need for a more strategic assessment of complex science–society relations in the context of citizen science has been expressed (Bonney et al. 2016). What we currently find are studies that show how the benefits on the level of individual participants help to cascade the outcomes to whole regions and communities. Johnson et al. (2014) report that individuals diffuse their acquired skills and knowledge to peers through social networks. Individual citizen scientists feel more confident expressing their ideas to natural resource managers and figures of authority (Cornwell and Campbell 2012); thus citizen science increases their political participation.

When engaging with an ecological system and its associated social institutions (e.g. policies, management practices) through citizen science, individuals may collectively gain knowledge that increases the capacity of the contextually embedded socioecological community to reorganise and adapt to changes. In this way, learning through citizen science projects is not only an outcome to be measured on an individual level but may also be an influence or driving force for meeting project goals for a whole community or region. Phillips et al. (2018) conclude that their focus on individuals' learning can contribute to civic action and policy forming and bring supportive evidence for the democratisation of science via citizen science engagement.

A good illustration of the socioecological impact of citizen science can be found in the community-based air quality monitoring projects that use low-cost measuring

devices. These projects demonstrate a higher sense of community as an outcome, stimulate discussions with policymakers, and influence political decisions in the involved regions (Van Brussel and Huyse 2019). There is also evidence from evaluation that citizen engagement in air quality monitoring may trigger the development of measures to avoid exposure to air pollutants or to an active political involvement in seeking solutions to the problem (Schaefer et al. 2020). Another impressive example comes from Japan, where Safecast, an international, volunteer-based organisation devoted to monitoring and openly sharing information on environmental radiation and other pollutants, emerged as a response to the lack of publicly available information about radiation levels after the Fukushima Daiichi Nuclear Power Plant disaster in 2011. The organisation provided tools and community resources to help people understand the complexities of radiation measurement and to make their own informed decisions. In addition to the measurement and easily accessible provision of radiation data, the group identified relevant information sources, summarised their contents, characterised any differences of opinion and interpretation that existed, and guided people to relevant resources. This spontaneous citizen science engagement helped to close ‘crucial gaps, ultimately the timely provision of data that citizens need to make informed decisions about their livelihoods and well-being is the government’s responsibility’ (Brown et al. 2016, p. 98). In their publication, the authors state that ‘this vigorous emergence suggests that a shift in social expectations and in the balance of information is already happening, from one which favors government and large institutions, to a more egalitarian and democratic relationship driven by citizen access to objective, independent information of high quality which has been generated by the citizens themselves’ (Brown et al. 2016, p. 98).

When looking at evaluation to indicate change in social practices, we should also not forget the work being done in related areas, such as *community-based participatory research*. Especially in the social sciences, participatory action research paved the way for our contemporary understanding of participation in citizen science, even though these approaches are still often neglected (Mayer et al. 2020). Evaluation frameworks have been established for participatory research that suggests questions, indicators, and measures that provide evidence for the effectiveness of certain programmes (e.g. Nash 2015). Similar to what we know from citizen science evaluation, these concepts look at aspects such as personal knowledge development, personal research skill development, organisational/group access to and use of information, and community and organisational development.

Tools and Methods in Evaluating Citizen Science

The tools and methods used for evaluation in citizen science mostly tend to follow standard social science practice, ranging from questionnaires, interviews, focus groups, participant observations, and documented self-reflections from the involved scientists and volunteers. In their overview of citizen science projects in biodiversity,

for example, Peter et al. (2019) report on a great diversity of study designs and methods for evaluation, with many projects relying on *self-reported data*.

Surveys are amongst the most frequent instruments to be applied for self-reported data, aiming mainly at collecting evidence for learning outcomes for the participants. Citizen science practitioners can nowadays turn to a number of shared resources online that help to collect insights into participants' motivations, satisfaction, benefits, self-efficacy, etc. (Phillips et al. 2018).

Interviews are another instrument frequently used for evaluation. These range from structured or semi-structured sets of questions to very open and exploratory formats. Scholars have published their interview guidelines to gather insights into their participants' motivations, engagement activities, and benefits, amongst others (Schaefer et al. 2020). But we also find narratives and forms of *storytelling* approaches as part of the evaluation spectrum. For example, Constant and Roberts (2017) combine narrative interviews with instruments like photo essays, research diaries, and storyboards to reveal the context-based, tacit, and intangible factors involved in personal outcomes.

Other evaluation approaches are built into the interaction process or are simply applied on the data available without an a priori evaluation design. An example of the former is the *embedded assessment* approach, where a series of games or quizzes are part of the citizen science activity and help to collect insights on participants' increased skills and knowledge in playful ways without people being aware that their knowledge is tested (Becker-Klein et al. 2016). The nonintrusive, non-design specific, approach can be exemplified by Luczak-Rösch et al. (2014) who analysed the comments shared by and amongst their online citizen scientists and measured how far citizen scientists adopted technical terms in their language as a sign of new knowledge gains.

For self-reflection and self-assessment of projects, Kieslinger et al. (2018) defined a set of key questions covering the three dimensions of the citizen science evaluation framework that have been implemented in an online questionnaire, which is part of the resources available on the European platform EU-Citizen.Science. The self-assessment tool aims to support the detection of strengths and weaknesses for an adaptive management of citizen science initiatives.

Case Studies That Show the Diversity of Citizen Science Evaluation Approaches

The diversity and continuous emergence of new practices in citizen science requires special caution when trying to draw comparisons across different projects outcomes. Nor can we speak of one exemplary approach or proxy to be followed by all when dealing with such a diverse population. The type of scientific work, societal challenge, and geographic scale of participation strongly shapes the strategies that a project uses to meet its goals. Likewise, evaluation and impact assessment depend

strongly on the project goals, as well as on contextual conditions that support or impede evaluation activities (such as the availability of respondents, the resources of the project, knowledge of evaluation techniques, etc.). The following section presents different case studies that illustrate the large leeway for citizen science projects to design their own evaluation objectives and activities.

CAPTOR: Applying Classical Evaluation Instruments for Impact Assessment at Individual and Socio-economic Levels

The CAPTOR project was funded by the European Commission's H2020 programme during the years 2016–2018. It combined citizen science activities with grassroots activism to create awareness for the ozone pollution problem in three regions in Austria, Italy, and Spain. In total, 46 low-cost sensors were distributed to volunteering households and public spaces to measure the ozone pollution during the summer. The defined core project objectives were individual learning outcomes amongst the participants involved in the measurement of ozone, behavioural change, and civic activism. However, the project not only wanted to affect participants at the individual level; it also aimed to benefit whole regions and drive political change to improve air quality. This was a highly complex endeavour compared to other pollutants, because tropospheric ozone is formed in *urban* areas through chemical reactions from precursor gases but emitted mainly in *rural* environments.

With these goals in mind and given that there were only a limited number of volunteers involved, the main evaluation instrument was guided interviews – backed up by pre- and post-questionnaires about the participants' knowledge and observations in the field. In total, 53 guided interviews were conducted at the end of each measurement period. This qualitative method helped to gain deep insights into the motivations of participants and their activities as volunteering hosts of ozone measurement devices (such as promoting the project to friends and neighbours). It also revealed in how far the hosts' involvement in the project impacted them as individuals and their neighbourhoods, in terms of knowledge, changed behaviour, and more political involvement in the area of air quality protection. The comparison of the detailed insights across the three different test bed regions allowed the project leaders to learn how far different contexts influenced outcomes at both individual and community levels, and details can be found in Schaefer et al. (2020).

The analysis of data showed that at an individual level the benefits of involvement in CAPTOR were mostly an increased knowledge about and awareness of ozone pollution amongst all private participants. They became promoters for the topic in their private network, and some even took more responsibility and raised the topic with environmental organisations and local municipalities, showing some engagement in political action taking. In experimental maker workshops, where people were invited to build their own measuring devices, evaluation data confirmed that building devices provides users with a feeling of empowerment and independence. In addition, the Spanish test bed revealed considerable regional impacts, beyond

individual benefits. After 3 years of citizen science activities, ozone values are now communicated on local television, and one municipality made the CAPTOR data and data from a reference station visible in a public place as well as on the municipality website to increase transparency and awareness. In order to fight the origins of ozone, alliances with other environmental organisations were established and pressure put on politicians in Barcelona. As a result of the collective efforts, a judgement was rendered by the Spanish court, committing Spanish regions to actively fight air pollution if their air quality data exceeds EU limit values.

Reflecting on the evaluation approach revealed that participants volunteered willingly in the interviews, talking openly to researchers. While this is certainly a good method for getting very rich feedback, it requires time and effort and cannot be applied to a larger group of participants. Also, one of the main constraints for impact assessment is the limited time and resources for longer-term data collection to measure the persistence of the perceived change over time. Thus, additional impact at regional policy level might become visible only at a later stage and not be captured by the project evaluation.

Plastic Pirates: An Experimental Study Design to Assess the Impact of Citizen Science Activities on School Children

The citizen science campaign Plastic Pirates was started in 2016 as part of the research focus Plastics in the Environment, which is funded by the German Federal Ministry of Education and Research (BMBF). The scientific analysis of the data collected by the Plastic Pirates is carried out in the ocean: laboratory of the Kieler Forschungswerkstatt. The goal of this nationwide programme is to develop and establish scientific procedures, methods, instruments, and concepts for investigating plastics in the environment. Plastic Pirates invites schools and youth organisations across Germany to investigate litter contamination at a riverside of their choice. The collected data is uploaded to a database and subsequently analysed by the experts at Kieler Forschungswerkstatt. At the moment of writing, little research has been carried out on the prevalence, distribution, and range of plastic waste in German waterways; therefore these young people are making an important contribution to researching the spread of macro- and microplastics.

Evaluation in the Plastic Pirates project was specifically concerned with the influence that a conscious participation in citizen science activities may have on the interest and motivation of students. The hypothesis was that students, who know they are part of a citizen science approach, are more interested and motivated than students who participate in the project in a classic educational approach. An experimental design setup was used to compare the interest and motivation of students in different conditions. Subjects nested in classes were randomly assigned to one of two groups. One group received the intervention (information about Plastic Pirates and participation in this citizen science approach), while the control group received no information about the citizen science approach in this project. The evaluation process studied what happened to subjects in each group via questionnaires, in

order to link differences in the outcomes to the intervention. This enabled the evaluation of the impact of the citizen science approach, and it was shown that the citizen science approach had a positive impact on students' interests (van den Bogaert et al. 2018).

In order to define the most effective study design for the specific question and hypothesis, this project approached the evaluation process from a widely accepted hierarchy of evidence, which claims that the most reliable evidence comes from systematic reviews, followed by evidence from randomised controlled trials, cohort studies, and case control studies. Since empirical evidence to answer the specific question was still poor, the evaluation team used an experimental field study design, introduced an intervention, and studied the effects. As common for experimental studies, the selection was randomised, meaning the subjects (in this case subjects nested in classes) were grouped by chance. Before starting interventions in the field, experimental studies need a lot of preparation time. During the preparation, different challenges have to be faced, for example, assigning school classes to different conditions when the teachers believe that students in a control group might miss essential content. Also, whole-day interventions require a challenging coordination of arrangements with teachers and school leaders, because lessons have to be cancelled. But the prospect for schools to work on current scientific topics is an attractive one that is unique to citizen science projects.

WTimpact: Assessing Citizen Science as a Tool for Knowledge Transfer

The interdisciplinary research project WTimpact, funded by the BMBF from 2017 to 2020, aimed to assess the effects of citizen science on the participants. This research project comprised three citizen science studies. The Leibniz Institute for Troposphere Research conducted a study in which volunteers measured air quality in Leipzig. In two studies run by the Leibniz Institute for Zoo and Wildlife Research, participants recorded the biodiversity of terrestrial mammals and bats in Berlin. In all three studies, citizen scientists were provided with devices for collecting data and afterwards uploaded their data to an Internet platform. On this platform, they could also find information on the specific content and the scientific process, analyse their own data as well as the aggregated data set, and discuss their experiences, questions, and results in the forum. Education researchers and psychologists from the IPN – Leibniz Institute for Science and Mathematics Education and the Leibniz Institut für Wissensmedien evaluated the outcomes at the individual level of participants with regard to content knowledge, scientific reasoning, ownership, and attitudes towards science and citizen science.

Participants were asked to fill in detailed questionnaires including tests of their content knowledge and scientific reasoning skills before and after the project. In addition, all activities of the participants on the Internet platform were tracked in order to correlate individual learning outcomes derived from the questionnaires with participation in scientific activities on the Internet platform.

This setup proved valuable for evaluation purposes. One very clear result was that offering background materials and tools for data analysis does not necessarily lead to citizen scientists actually using these opportunities for participation in scientific activities. Thus, if evaluation results are interpreted based on the scientific activities theoretically available to participants, rather than the ones they actually engage in, this may lead to misleading results.

The questionnaire was refined in different rounds of the project, taking into account the participants' feedback. One of the challenges was to balance the aim of adhering to scientific standards in social and educational sciences (e.g. ensuring internal consistency by a certain degree of redundancy) with the amount of time and effort participants could be expected to invest in answering the questionnaires. Restricting the length of the questionnaires, in turn, meant that not all potentially relevant outcomes could be covered. Regarding the knowledge and scientific reasoning tests in the questionnaires, another challenge was that some participants expressed that they felt like they were taking an IQ test, leading them to wonder about the 'true purpose' of the study. This improved when the scientists explained to the citizen scientists in more detail what the questionnaire entailed and what they were aiming to achieve with it (albeit without providing the research hypotheses to avoid prompting participants to provide socially desirable answers).

Initial results from the evaluation showed that participants gained content knowledge in the project, and this was influenced by their scientific reasoning skills and their motivation. Also, their attitudes towards citizen science and towards science in general improved, which, amongst other things, depended on their attitudes towards the topic and their participation in scientific activities on the Internet platform.

Future Trends in the Evaluation of Citizen Science

We see future trends in the evaluation of citizen science on three levels. First, new approaches will put the evaluation of citizen and community benefits in perspective. Second, new methods will allow us to enrich the way evaluation is done. Finally, new topic areas, such as human health and food, will influence the demands on the evaluation of citizen science initiatives.

New Approaches

New approaches to evaluation will focus strongly on the dimensions of individual and socioecological benefits, by involving all actors more intensively in defining, collecting, and analysing evaluation data.

In a recent paper, Mayer et al. (2020) propose a participatory approach to evaluation, which they label *co-evaluation*. It is defined as a process that involves all relevant actors in a project in an iterative evaluation practice and combines

methods of participatory action research for evaluation purposes. It is inspired by community-based participatory research as well as science and technology studies' perspective on the evaluation of public participation exercises in research. Project goals and objectives and understanding of success, challenges, and unintended aspects are collectively discussed and documented at the beginning of a project and regularly revisited during the research design and execution, ideally even beyond the project's end. Assessment and intended impacts hence become transparent entities in the project design and important elements of the research tools inventory. With this participatory approach towards evaluation, the authors argue that citizen and community benefits, as well as the wider sociopolitical and ecological impact, can be equally assessed, alongside scientific goals, and form an integral part of the evaluation scheme.

The presented approach does not focus on individual learning outcomes but has a more social focus. During the co-evaluation process, which is conducted as a team effort that includes relevant stakeholder representatives, the assessment procedures may vary greatly in their manifestation, from surveys to storytelling to improvisational theatre, depending on the context. Another important aspect of this approach relates to the dimension of *open science*. In the process of co-evaluation, informed consent procedures and open data strategies are determined collectively by the participants.

New Methods

New evaluation methods will help to collect evaluation data and deepen the understanding of citizen science outcomes. In the past few years, a rising number of articles reflecting on the usage of specific methods for evaluation have appeared. One may assume that this increased interest in applied methods will also result in an uptake of new methods and instruments. One example is the *experience sampling method* (ESM) (Larson and Csikszentmihalyi 2014). This approach might help to better answer questions like: What is the typical citizen scientist like, and how much do citizen scientists differ from each other? What motivates citizen scientists? and What do they learn, and how do they change their behaviour? So far, it is common in evaluation research to ask participants to reflect back over weeks and months and provide a summary account of their experiences. Therefore, new approaches like ESM, which have not yet been applied in citizen science, might provide a valid instrument for systematic self-reporting, allowing the creation of an archival file of daily experiences. In ESM, upon receipt of random signals, participants respond to questions about their objective situation and their subjective state at that moment, such as their cognitive, emotional, and motivational state (Larson and Csikszentmihalyi 2014). Such data can be used to generate summary accounts without the biases introduced by retrospection over relatively long periods and allow observing changes in participants over time, as well as individual differences in such change.

Another method that can be considered in citizen science evaluation requires a shift in perspective, towards a person-centred approach. *Latent class analysis* (LCA) is a method frequently applied in social science data analysis (Collins and Lanza 2010) and is used to trace the heterogeneity in a group to a number of underlying homogeneous subgroups, at specific measurement points, allowing for their longitudinal extensions. LCA also allows empirically representing not directly measurable aspects, like social class, lifestyle, and recreational behaviour, via directly measurable variables in the form of typologies.

If applied in a citizen science context, this may allow a better understanding of the subpopulations engaged in scientific tasks and their specific characteristics that might change over time.

New Topics of Research

And finally, citizen science is finding its way into new fields of research, bringing new challenges from an evaluation perspective. One growing area of opportunity for citizen science methods is in the fields of health and biomedical research. If we take health research as an example, we come across online communities such as PatientsLikeMe, where patients share their health data for research on various conditions, generate hypotheses based on common experiences and conduct their own experiments. Such citizen science activities are certainly prime opportunities with regard to health literacy, empowerment, and active participation in public health governance (Den Broeder et al. 2018), which deserve proper project evaluation in order to show those impacts.

However, this type of citizen science raises complex ethical issues that may be of less relevance for other disciplines. Ethical aspects have to be given special attention, especially when assessing the impacts of such citizen-driven initiatives, and we need to look very carefully at potential negative and unwanted effects, like the spreading of misinformation. The peer review process of scientific publishing combined with the slow pace and high cost of clinical studies has limited broader participation in health and biomedical research so far but has the virtue of preventing the spread of misinformation (Wiggins and Wilbanks 2019) – an aspect which should be carefully considered in future evaluation activities by involving a wide range of stakeholders in evaluation and applying a diversified set of evaluation instruments that allow critically examining the citizen science outcomes and impacts from various angles.

Lessons Learned and Recommendations in Evaluating Citizen Science

For many citizen science projects, outcome evaluation, beyond the purely scientific results, is not a priority. Having to deal with the active involvement of citizens and the continuous bidirectional communication with all target groups, while driving the research process and answering the research questions, constitutes a considerably higher workload compared to traditional research practice. Many initiatives have limited resources and lack specific expertise in evaluation. However, as pressure from funding agencies, universities, ministries, etc. increases, there is a rising need to provide evidence of the outcomes and impacts of citizen science projects beyond the science itself.

Thus, evaluation of citizen science is in a continuously developing state, not unlike the field of citizen science itself. In recent years, the number of scientific articles, discussions, and demands for evaluating citizen science initiatives has been growing steadily. Examples are now emerging of citizen science projects that provide evidence for concrete outcomes and impacts, and some also share their lessons on the applied tools and methods for evaluation.

The majority of scientist-led citizen science initiatives measure their outcomes in the scientific dimension, using broadly accepted and standardised indicators, such as the number of scientific publications or presentations at scientific conferences. A smaller number of these initiatives investigate their impact on the individual citizens involved in the research. We find evidence in the literature of how the involvement of volunteer citizens in the research process in different thematic fields impacts participants' content knowledge, skills, attitudes, and behaviour concerning the topic of research. Evaluation methods that focus on the investigation of changes on individual participants have been developed and made available to the citizen science community for reuse, enabling comparisons across an increasing number of projects. Looking beyond the effects on individuals, the evaluation of outcomes that affect whole regions, communities, and socioecological systems is an even more complex task. But, here we also find a number of interesting studies that show the benefits of citizen science approaches in this regard.

Still, there is a call for the further development, sharing, and uptake of standardised, easy-to-use, and proven evaluation instruments that go beyond impacts on individuals. Such instruments could benefit both project owners and the citizen science community overall and would allow for a deeper understanding of different contexts that influence the changes in individuals, communities, and regions.

Alongside the calls for more standardisation, it is also necessary to keep discussions about evaluation open and self-reflective, not only to continually improve, but also to stay flexible and adaptable to the continuous evolution of citizen science itself. There are, for instance, standardised metrics for scientific outcomes of citizen science; but in grassroots initiatives, which prioritise the impact on socioecological systems and are not led by academics, scientific outcomes are less likely to be published in research journals. These projects require additional metrics to provide evidence of their scientific impact. Also, experimenting with new evaluation

approaches and methods, as introduced above, would be highly enriching: going beyond self-reporting to involve citizens from the very beginning in a co-evaluation process, or taking an inclusive look at unintended outcomes.

To date, there is little analysis of targets which were not achieved or of unintended results. Reporting on these is crucial for the whole field to learn from others' mistakes and improve future projects. But assessing unintended results is difficult, because evaluation instruments are geared towards the intended project goals, and unintended side effects may only come to light as anecdotal evidence.

At the very least, we need better access to information on the validated evaluation practices that are already in use. Existing knowledge and experiences are currently dispersed across countries, spread in disciplinary-focused journals, and published on institutions' websites, requiring much effort to find and extract the relevant content to be reused in different project contexts. A better overview is required of outcomes that have already been documented, applied instruments that have been shown to be successful, resulting in a knowledge base that is easily accessible and can continually grow, with new insights, instruments, and processes, via an active discussion by those involved. Complementary to such a comprehensive and accessible knowledge base, we would also recommend more workshops and training dedicated to evaluation, fostering the mutual exchange of knowledge and experiences between the members of the citizen science community and other disciplines, as well as involving political decision makers in the discussion. Such a process of consolidation needs to be nurtured continually, assuring that higher-level organisations (such as government authorities, research funders, and citizen science associations) are involved in shaping a broader societal assessment of citizen science initiatives and take up the results in their own programmes and agendas.

References

- Ballard, H. L., Dixon, C. G. H., & Harris, E. M. (2017). Youth-focused citizen science: Examining the role of environmental science learning and agency for conservation. *Biological Conservation*, 208, 65–75.
- Becker-Klein, R., Peterman, K., & Stylinski, C. (2016). Embedded assessment as an essential method for understanding public engagement in citizen science. *Citizen Science: Theory and Practice*, 1(1), 8. <https://doi.org/10.5334/cstp.15>.
- Bela, G., Peltola, T., Young, J. C., Balázs, B., Arpin, I., Pataki, G., et al. (2016). Learning and the transformative potential of citizen. *Conservation Biology*, 30(5), 990–999.
- Bonn, A., Richter, A., Vohland, K., Pettibone, L., Brandt, M., Feldmann, R., et al. (2016). *Green paper citizen science strategy 2020 for Germany*. Berlin: Helmholtz-Zentrum für Umweltforschung – UFZ.
- Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Understanding of Science*, 25(1), 2–16.
- Brown, A., Franken, P., Bonner, S., Dolezal, N., & Moross, J. (2016). Safecast: Successful citizen-science for radiation measurement and communication after Fukushima. *Journal of Radiological Protection*, 36(2), S82–S101.
- Collins, L., & Lanza, S. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health Sciences*. Hoboken: Wiley. <https://doi.org/10.1002/9780470567333>.

- Constant, N., & Roberts, L. (2017). Narratives as a mode of research evaluation in citizen science: Understanding broader science communication impacts. *Journal of Science Communication, 16* (4), A03.
- Cornwell, M. L., & Campbell, L. M. (2012). Co-producing conservation and knowledge: Citizen-based sea turtle monitoring in North Carolina, USA. *Social Studies of Science, 42*(1), 101–120.
- Crall, A. W., Jordan, R., Holfelder, K., Newman, G. J., Graham, J., & Waller, D. M. (2012). The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. *Public Understanding of Science, 22*(6), 745–764.
- Den Broeder, L., Devilee, J., Van Oers, H., Schuit, A. J., & Wagemakers, A. (2018). Citizen Science for public health. *Health Promotion International, 33*(3), 505–514.
- Edwards, R., Kirn, S., Hillman, T., Kloetzer, L., Mathieson, K., McDonnell, D., & Phillips, T. (2018). Learning and developing science capital through citizen science. In S. Hecker, M. Haklay, A. Bowser, Z. Makuch, J. Vogel, & A. Bonn (Eds.), *Citizen science – Innovation in open science, society and policy* (pp. 381–390). London: UCL Press.
- Follett, R., & Strezov, V. (2015). An analysis of citizen science based research: Usage and publication patterns. *PLoS ONE, 10*(11), e0143687.
- Goertz, G., & Mahoney, J. (2012). *A tale of two cultures: Qualitative and quantitative research in the social sciences*. Princeton: Princeton University Press.
- Groulx, M., Brisbois, M. C., Lemieux, C. J., Winegardner, A., & Fishback, L. A. (2017). A role for nature-based citizen science in promoting individual and collective climate change action? A systematic review of learning outcomes. *Science Communication, 39*(1), 45–76.
- Jennett, J., Kloetzer, L., Schneider, D., Iacovides, I., Cox, A. L., Gold, M., et al. (2016). Motivations, learning and creativity in online citizen science. *JCOM, 15*(3), A05.
- Johnson, M. F., Hannah, C., Acton, L., Popovici, R., Karanth, K. K., & Weinthal, E. (2014). Network environmentalism: Citizen scientists as agents for environmental advocacy. *Global Environmental Change, 29*, 235–245.
- Kasperowski, D., Kullenberg, K., & Mäkitalo, A. (2017). *Embedding citizen science in research: Forms of engagement, scientific output and values for science, policy and society*. <https://doi.org/10.31235/osf.io/tfsgb>.
- Kieslinger, B., Schäfer, T., Heigl, F., Dörler, D., Richter, A., & Bonn, A. (2018). Evaluating citizen science: Towards an open framework. In S. Hecker, M. Haklay, A. Bowser, Z. Makuch, J. Vogel, & A. Bonn (Eds.), *Citizen science – Innovation in open science, society and policy* (pp. 81–98). London: UCL Press.
- Kullenberg, C., & Kasperowski, D. (2016). What is citizen science? – A scientometric meta-analysis. *PLOS ONE, 11*(1), e0147152.
- Larson, R., & Csikszentmihalyi, M. (2014). The experience sampling method. Flow and the foundations of positive psychology. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 21–34). Dordrecht: Springer.
- Leydesdorff, L. (2005). Evaluation of research and evolution of science indicators. *Current Science, 89*(9), 1510–1517.
- Luczak-Rösch, 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 *Proceedings of the eighth international AAAI conference on weblogs and social media* (pp. 315–324). Palo Alto: AAAI Press.
- Masters, K., Oh, E. Y., Cox, J., Simmons, B., Lintott, C., Graham, G., et al. (2016). Science learning via participation in online citizen science. *Journal of Science Communication, 15*(3), 1–33.
- Mayer, K., Kieslinger, B., & Schäfer, T. (2020). *Co-evaluation: Re-situating participatory social research in citizen science*. Manuscript submitted for publication.
- Nash, C. (2015). *Evaluating community-based participatory research*. Guelph: Community Engaged Scholarship Institute.
- Örtengren, K. (2004). *The logical framework approach: A summary of the theory behind the LFA method*. http://www.eejp.org/resources/lfa_approach.pdf. Accessed 28 Aug 2014.
- Peter, M., Diekötter, T., & Kremer, K. (2019). Participant outcomes of biodiversity citizen science projects: A systematic literature review. *Sustainability, 11*(10), 2780. <https://doi.org/10.3390/su11102780>.

- Phillips, T., Porticella, N., Conostas, M., & Bonney, R. (2018). A framework for articulating and measuring individual learning outcomes from participation in citizen science. *Citizen Science: Theory and Practice*, 3(2), 3. <https://doi.org/10.5334/cstp.126>.
- Schaefer, T., Kieslinger, B., & Fabian, M. C. (2020). Citizen-based air quality monitoring: The impact on individual citizen scientists and how to leverage the benefits to affect whole regions. *Citizen Science: Theory and Practice*, 5(1), 6. <https://doi.org/10.5334/cstp.245>.
- Stepenuck, K. F., & Green, L. T. (2015). Individual-and community-level impacts of volunteer environmental monitoring: A synthesis of peer-reviewed literature. *Ecology and Society*, 20(3), 19. <https://doi.org/10.5751/ES-07329-200319>.
- Suomela, T. E. (2014). *Citizen science: Framing the public, information exchange, and communication in crowdsourced science* [Doctoral dissertation, University of Tennessee]. https://trace.tennessee.edu/utk_graddiss/2864. Accessed Oct 2019.
- Trautmann, N. M., Shirk, J. L., Fee, J., & Krasny, M. E. (2012). Who poses the questions? Using citizen science to help K-12 teachers meet the mandate for inquiry. In J. L. Dickinson & R. Bonney (Eds.), *Citizen science: Public participation in environmental research* (pp. 179–190). Ithaca: Comstock Publishing Associates.
- Van Brussel, S., & Huyse, H. (2019). Citizen science on speed? Realising the triple objective of scientific rigour, policy influence and deep citizen engagement in a large-scale citizen science project on ambient air quality in Antwerp. *Journal of Environmental Planning and Management*, 62(3), 534–551.
- van den Bogaert, V., Ebbeskotte, T., & Wirth, J. (2018). *Citizen science project as a teaching format – Does it work?* Presented at the fourth Austrian Citizen Science Conference 1–3 February 2018, Salzburg.
- Weiss, C. H. (1998). *Evaluation. Methods for studying programs and policies* (2nd ed.). Upper Saddle River: Prentice Hall.
- Wickson, F., & Carew, A. (2014). Quality criteria and indicators for responsible research & innovation: Learning from transdisciplinarity. *Journal of Responsible Innovation*, 1(3), 254–273.
- Wiggins, A., & Wilbanks, J. (2019). The rise of citizen science in health and biomedical research. *The American Journal of Bioethics*, 19(8), 3–14.
- Williams, K., & Grant, J. (2018). A comparative review of how the policy and procedures to assess research impact evolved in Australia and the UK. *Research Evaluation*, 27(2), 93–105.
- Zerbe, F., & Wilderman, C. (2010). Monitoring impacts of new gas-drilling technologies. *The Volunteer Monitor*, 21(1), 1–20.

Teresa Schaefer is a researcher at the Centre for Social Innovation (ZSI), Austria. With a background in economic sciences, her work focuses on participation processes in digital social innovations and the assessment of their impact.

Barbara Kieslinger is a project manager and researcher at the Centre for Social Innovation (ZSI). In her projects, she experiments with co-design and co-evaluation processes, fostering participatory stakeholder involvement.

Miriam Brandt is head of Science Management at the Leibniz Institute for Zoo and Wildlife Research (IZW) in Berlin, Germany. She leads a number of projects in knowledge transfer and citizen science and is interested in the learning outcomes of citizen science.

Vanessa van den Bogaert has a master's in education as well as a BSc in biology and mathematics. She worked at the Museum für Naturkunde Berlin on the project *Bürger schaffen Wissen*. Her research focuses on evaluation, especially the effects of participation processes in citizen science projects on motivational variables. She founded a working group on the science of citizen science.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

