



Diverse and Important Ways Evaluation can Support and Advance Citizen Science

RESEARCH PAPER

CATHLYN DAVIS

VERONICA DEL BIANCO

KAREN PETERMAN

ANDREA GROVER

TINA PHILLIPS

RACHEL BECKER-KLEIN

*Author affiliations can be found in the back matter of this article

Ubiquity press

ABSTRACT

Evaluation offers many benefits for citizen science including the ability to inform design and improve project programming; to aid in understanding impacts on volunteer outcomes; to validate project successes; and to advance best-practices in the field. However, evaluation and subsequent use of its findings in citizen science remains limited. Here, we applied an existing typology to document evaluation use among 15 citizen science project leaders who were deeply involved in a collaborative evaluation process. From their evaluation efforts, these leaders gained new and deeper understanding of their volunteers and programming (conceptual use); made critical changes to their projects (programmatic use); shared their evaluation findings with others (dissemination use); and expanded their attitudes and actions with regard to evaluation (process use). Knowledge gains from evaluation prompted the project leaders in our study to change their training, revise their protocols, add resources, and even terminate an unproductive project. Through reports, presentations, and publications, the project leaders shared findings related to skill proficiency with their volunteers, other staff members, practitioners in other citizen science projects, funders, researchers, and evaluators. Our study makes connections between the evaluation-use literature and citizen science practice, and offers recommendations to address the challenge of limited application of evaluation within citizen science. As such, this paper can help project leaders understand the important and diverse ways evaluation can support individual projects and the larger field. It also raises questions on the role of collaboration in citizen science evaluation.

CORRESPONDING AUTHOR:

Cathlyn Davis

University of Maryland Center
for Environmental Science, US
cathlyn.davis@umces.edu

KEYWORDS:

citizen science; collaboration;
embedded assessment;
evaluation use; stakeholder;
typology

TO CITE THIS ARTICLE:

Davis, C, Del Bianco, V,
Peterman, K, Grover, A, Phillips,
T and Becker-Klein, R. 2022.
Diverse and Important Ways
Evaluation can Support and
Advance Citizen Science. *Citizen
Science: Theory and Practice*,
7(1): 30, pp. 1-10. DOI: <https://doi.org/10.5334/cstp.482>

INTRODUCTION

Citizen science efforts seek to advance science research while promoting science learning among volunteer participants (National Academies of Sciences, Engineering, and Medicine 2018). Both goals depend on well-designed training and resources that prepare volunteers to follow scientific protocols (i.e., standardized methods) for collecting or interpreting data (National Academies of Sciences, Engineering, and Medicine 2018). If this preparatory support is not well aligned with a project's protocol or volunteers' strengths and motivation, then volunteers may lack sufficient skill proficiency to complete the work (Brown and Williams 2019) and data may be devalued or discarded (Burgess et al. 2017). If a protocol exceeds the interests or abilities of volunteers, these participants may not complete required tasks or may even drop out of the project (West and Pateman 2016).

These examples point to the need for robust data on volunteer outcomes, which can include volunteer proficiency on targeted science inquiry skills, their science knowledge, and their self-efficacy to make scientific observations (Phillips et al. 2018). Evaluation provides a pathway to acquiring this evidence in ways that are rigorous and can meet the goals and needs of citizen science (Phillips et al. 2018; Stylinski et al. 2020). Results from evaluation can be used to inform design and improve project programming, including recruitment, training, and protocols; to aid in understanding impacts on volunteer outcomes; to validate project successes; and to advance best-practices in the field. For instance, a citizen science leader could use a performance-based evaluation to have volunteers demonstrate their proficiency on science inquiry skills targeted by the project (e.g., identify a species, navigate to a collection site, or estimate counts), which in turn could ensure acquisition of high-quality science data and alignment with science literacy outcomes (Becker-Klein, Peterman, and Stylinski 2016). Despite these benefits, evaluation in citizen science remains limited, especially for volunteer skill proficiency (Bowser et al. 2020; Burgess et al. 2017; Phillips et al. 2018; Stylinski et al. 2020). Cited challenges include lack of time, of expertise, and of supporting resources (Phillips et al. 2018; Stylinski et al. 2020).

Here we present a baseline qualitative study that introduces and illustrates the diverse and important ways that evaluation can be used to support and advance citizen science efforts. We approach this work by directly linking the citizen science field to the extensive literature on evaluation use (described in the next section) and by viewing these linkages through the lens of practitioners (here, practitioners are citizen science project leaders).

Specifically, we ask: how do project leaders use findings from the evaluation of their citizen science effort? Our study focuses on leaders from 15 citizen science projects who were deeply involved in developing and implementing evaluation in their individual projects and in using the resulting findings. A long-held assumption in the field of evaluation is that stakeholders who are more involved in the process will be more likely to use evaluation (Daigneault 2014). Additional benefits of stakeholder participation in evaluation include feelings of satisfaction, ownership, and trust and fairness, as well as perception of higher validity and credibility (Froncek and Rohmann 2019). Thus, we describe the context of practitioner-led efforts across citizen science projects, and consider the role that collaboration played in evaluation use in our discussion.

EVALUATION USE

There is extensive literature on evaluation use, and active discussion on the definition of this complex concept that includes associated theories, influencing factors, and more (e.g., Alkin and King 2016, Alkin and King 2017, King and Alkin 2019). Patton (2020) recently provided a straightforward and broad definition of evaluation use: "whatever understandings, learnings, actions, changes, attitudes and/or knowledge [that] follow from evaluation findings and/or process" (p. 588). He notes that it varies by context and purposes, and includes intended and unintended uses and users. Drawing from their historical review, Alkin and King (2016) outline four broad types of evaluation use (conceptual, instrumental, symbolic and process). Bundi, Frey, and Widmer (2021) provide a brief summary of this four-part typology. We offer this typology in Table 1 and have coupled each type with an example relevant to citizen science. The first three types of evaluation use are based on the findings of an implemented evaluation. Conceptual use occurs when these findings deepen a project leader's understanding and shift their perspectives about volunteers and the design of their citizen science efforts. Programmatic use occurs when this new understanding leads to project revisions, such as changes in training topics, online resources, and scientific protocols. Dissemination use occurs when the evaluation findings are shared with stakeholders, including volunteers, funders, staff, and other practitioners. The final type—process use—does not come from the evaluation findings but instead centers on gains associated with participating in the evaluation process itself (Patton 1997). It is characterized by shifts in citizen science project leaders' knowledge, attitudes, or behaviors about evaluation as a result of participating in the evaluation process.

EVALUATION USE TYPE	DEFINITION (FROM BUNDI ET AL. 2021, P. 2)	CITIZEN SCIENCE EXAMPLES
Conceptual use	“Indirect use of systematically generated knowledge that opens up new ways of thinking and understanding, or that generates new attitudes or changes existing ones”	From the evaluation results, a project leader gains a deeper understanding of the importance of training volunteers on a scientific protocol (standardized method) to estimate species numbers.
Programmatic use	“Direct use of systematically generated knowledge (for example, evaluations) to take action or make decisions” (also called instrumental use).	A project leader revises the focus of their training when evaluation findings reveal volunteers lack proficiency in the skills needed to collect scientific data.
Dissemination use	“Use of evaluations to support an already preconceived position in order to legitimize, justify or convince others of their position” (also called symbolic use).	A project leader presents their evaluation findings to potential funders to convince them of the impact of their project on volunteers’ understanding of the process of science.
Process use	“Use that occurs due to the process and not due to the results of an evaluation.”	A project leader who is actively involved in developing and implementing a survey of their volunteers’ knowledge develops a greater appreciation of the time and effort required to develop robust survey questions.

Table 1 Typology of evaluation use for citizen science.

Although evaluation use is one of the most researched areas in the evaluation literature, only in the last decade have researchers examined the extent of evaluation use in different contexts (e.g., [Daigneault 2014](#); [D’Ostie-Racine et al. 2016](#); [Peterman and Gathings 2019](#); [Shaw and Campbell 2014](#)). Our baseline research contributes to this small number, and as far as we know, is the first multi-project exploration of evaluation use in citizen science.

CONTEXT

As part of a larger study funded by the National Science Foundation (DRL #1713424) to promote evaluation use in citizen science, our team collaborated with leaders to develop performance-based embedded assessments that could be implemented and used in their respective projects. Embedded assessment can be a form of evaluation. It is well matched to informal learning experiences—like citizen science—because it is seamlessly integrated into the experience and allows participants to demonstrate their knowledge and skills ([Peterman et al. 2017](#)). Embedded assessment can include analyzing science data submitted by volunteers or assessing volunteers as they perform a skill required by a citizen science protocol. For this study, one embedded assessment strategy (Secondary Analysis) focused on analyses of existing scientific data collected by volunteers to search for evidence of skill proficiency ([Peterman et al. in review](#)), while a second strategy (Shared Measures) produced embedded-assessment measures around common observation skills, such as “notice relevant features,” that could be used by more than one project ([Becker-Klein et al. in review](#)).

We had project leaders in both strategy groups actively participate in the evaluation process and associated decision-making. This included developing analysis procedures for existing data or creating evaluation measures; implementing the procedures or measures; analyzing the resulting findings; and using these evaluation findings in their projects. The leaders collaborated with each other and our team as they tackled this work over two to three years. Leaders from five Secondary Analysis projects initially gathered to discuss and reflect on the volunteer science inquiry skills that would be the focus of their separate embedded assessment efforts. They then worked independently to organize and analyze their own existing datasets (i.e., science data submitted by volunteers), meeting multiple times with our team and the other leaders to share their process, hurdles, successes, and final results ([Peterman et al. in review](#)). In a similar fashion, we guided the leaders from ten Shared Measures projects in selecting science inquiry skills for the shared embedded-assessment measures (i.e., measures that would be relevant across many projects). These project leaders then collaborated closely with our team and with each other to co-develop several embedded assessments of these skills and to share their plans, successes, hurdles, and final results from these assessments ([Becker-Klein et al. in review](#)).

To ease challenges associated with both embedded assessment strategies, we focused on a single broad science inquiry skill—scientific observation. Thus, we selected leaders for both strategy groups who directed citizen science projects that had adult volunteers collect observation-based environmental data. All leaders also expressed an interest in expanding their evaluation efforts.

Beyond this, their projects represented some of the diversity in citizen science efforts focused on scientific observations as they differed in the mode of participation (computer- and field-based), geographic scale of participation (local, regional, and nationwide), longevity (founded in 1993 to 2016), number of volunteers (80 to 20,000), and types of data collected (biotic, abiotic, and astronomical data).

INSTRUMENT AND PROCEDURES

For each project, we conducted and transcribed semi-structured recorded video interviews (30 to 90 minutes), which provided some standardization of the questions while allowing for context-specific variability associated with the exploratory nature of our work (Adams 2015). All interviews were conducted with the project leaders and sometimes with other staff members. For simplicity, we refer to this group as project leaders hereafter. Interviews were conducted midway through the evaluation work (between fall 2018 and winter 2020) and after the project leaders had completed their evaluation work (between fall 2019 and spring 2021). The midpoint interviews asked project leaders about their evaluation work thus far, while the endpoint interviews had them directly reflect on the four types of use in the context of their evaluation findings and their participation in the evaluation. In the endpoint interviews, we concluded by asking if project leaders had anything to add. If they did not mention collaboration with other project leaders in their response, we specifically followed up by asking their thoughts on the value of the collaborative aspect of this work.

ANALYSIS

We analyzed the data using thematic analysis and a deductive coding scheme (Miles, Huberman, and Saldaña 2018) to identify cross-cutting themes aligned with known types of evaluation use. Specifically, two researchers on our team created an initial codebook based on Table 1, and piloted it using a portion of the dataset to more fully describe each code and to provide examples. After a review by the rest of the team, we made refinements and produced the final version (see Supplemental File 1: Evaluation Use Codebook). Given our small sample size, they used consensus coding to analyze the data (Miles et al. 2018), with the two researchers coding each document independently with the NVivo 12 software, and discussing any disagreements until consensus was reached. We also reviewed meeting notes and any artifacts (reports and graphics) provided by the project leaders to help us understand findings associated

with their evaluation use. We included any instances in which project leaders described concrete intentions to directly use the evaluation findings or process in the near future. We did not code any project leaders' comments related to changes already underway before this study began.

RESULTS

In this section, we describe how the project leaders used findings from their embedded assessment, and we have organized these descriptions by the four evaluation use types. We also share their reflections on the importance of collaboration in the evaluation process.

EVALUATION USE

All four types of evaluation use were readily apparent among the 15 projects. Specifically, 13 project leaders cited conceptual use, 11 cited programmatic use, 13 cited dissemination use, and 15 cited process use. Every project leader used evaluation in at least three ways, and seven used it in all four. Below we describe the different ways they used evaluation in their citizen science projects.

Conceptual use

As noted, conceptual use occurs when evaluation impacts understanding and perspectives about volunteers and citizen science programming. The leaders in this study described many examples of such knowledge gains. Common across most projects was confirmation from the findings that volunteers collect robust data, but context matters when it comes to skill proficiency. For example, one project leader found that group size had an impact on volunteer-submitted data. That is, they discovered that the number of volunteers in a data-collection group was positively correlated with individual volunteers' detection rate of marine species. Another project leader found individual volunteer attributes impacted skill proficiency. In this case, the volunteers who claimed high expertise in insect identification scored higher than others on the embedded assessment.

For some of the project leaders, their findings raised additional questions about citizen scientists and their skills, such as what level of accuracy is needed in volunteer-submitted data, what are effective ways to help volunteers gain skill proficiency, and how does scientific vocabulary help them improve this proficiency. For example, one project leader pondered how to build on everyday skills associated with pattern recognition to foster scientific observation skills, and stated,

"You're always making observations and comparing [the object under observation] in your head to other similar things. And that's a native skill that people have in pattern recognition and categorizing. But what I found by 'spying' on participants is that they're using their native pattern recognition skills, but they're not able to break them down and do it comparatively...that was pretty eye-opening to think through that problem a little more and how you could...capitalize on their native pattern recognition skills to hone them in a way that a scientist hones them."

Programmatic use

Programmatic use occurs when knowledge gains from evaluation are employed to make programmatic decisions and changes. This type of evaluation use may be what most practitioners think evaluation is all about. Indeed, the leaders in our study used their new knowledge to make many programmatic changes including revisions to their project orientation, training, resources, and protocols. In some cases, the changes consisted of additions. For example, in a citizen science mapping project of trees and a parasitic fungus, the project leader assessed volunteers' proficiency in distinguishing the tree of interest from other trees. They were surprised to discover that, while volunteers had no problems identifying the tree, they were unfamiliar with the morphology of the infecting fungus. The project leader applied these findings by adding the missing guidance on fungus identification to their training.

In other cases, leaders eliminated programmatic elements. For instance, a species identification project included extensive training on many different insect species that volunteers might encounter in the field. The evaluation findings revealed that this broad review was overwhelming to volunteers, and thus the project leader streamlined the instructions to focus on the most common species and how they can be distinguished from others.

Still other leaders overhauled the protocol that volunteers followed to collect scientific data. For example, a project had volunteers hike to distant field sites, and then search for evidence of visitation by the species of interest. Their assessment revealed that some sites presented significant navigation challenges for their volunteer hikers. As a consequence, they changed their protocol, dropping the most remote sites and adding field markers to make the remaining ones more apparent.

Finally, evaluation findings can even point to the need to orient limited resources to more productive volunteer efforts. For one of the projects, the evaluation results revealed volunteers did not improve their data collection skills even with increased training and contact with staff

members. With these findings, the leader decided to discontinue the project, and focus on other citizen science efforts in their program. They reported, "We learned so much about [our project] that it actually also helped us make the decision to close [it]." Together these examples highlight the diverse ways that the project leaders directly used their evaluation findings to make decisions and take action within their citizen science projects.

Dissemination use

This use centers on dissemination of the evaluation findings to others. The project leaders in our study shared their evaluation findings of participants' skill proficiency with three key audiences: their stakeholders, their volunteers, and other practitioners and researchers in citizen science.

Almost all of the project leaders (13 of 15) provided examples of dissemination to staff, funders, and other stakeholders to demonstrate progress and to justify continuation of the citizen science project. This included the importance of sharing findings with coordinators who carry out much of the citizen science work. For example, one project leader wanted to promote the value of volunteer training with their geographically distributed staff, and thus told them, "We can prove...within some confidence that... people who are trained give us a lot better data and I think... that's...important [sic] not just from a data perspective, but from a business use case perspective."

A third of the project leaders (5 of 15) also shared findings with their volunteers to encourage their continued involvement. These leaders noted that this type of dissemination is important because volunteers are interested in findings about their skill proficiency, and have a right to know. By contrast, some project leaders expressed hesitation in sharing assessment results with volunteers. One did not want to promote the perception that the volunteers were under study, while another was simply unsure what impact this sharing would have on volunteers.

Finally, several project leaders (6 of 15) shared their evaluation findings with the larger citizen science field through conference presentations and workshops, reports, and manuscripts for peer-reviewed publications. These efforts focused on citizen science practices (e.g., how they are encouraging volunteers to collect high-quality data) and on embedded assessment practices (e.g., how they are using an online module to assess volunteers' skills).

Process use

Finally, process use occurs when there are knowledge, attitudes, or behavioral changes as a result of being involved in the evaluation process. The leaders cited multiple examples of how their active participation changed their

perspectives and actions with regard to evaluation. First, they developed a deeper appreciation for evaluation and assessment efforts, including the value of robust evaluative data versus anecdotally based perceptions or assumptions. As one project leader stated, “You’ve got to document it to be able to actually ask those questions and answer them with the right information in hand. Not just [go by] our impressions.”

Second, the project leaders in this study demonstrated a broader view of evaluation that was not limited to pre/post surveys and self-reported data, but included embedded assessments and performance-based data (i.e., in which volunteers demonstrate their skills). As one project leader highlighted, “We’re gaining maturity in our understanding of what to do and how to do [assessment], and what works and what doesn’t.” Furthermore, project leaders gained an understanding of what is required to implement a rigorous evaluation, including articulating the targeted science inquiry skills that will be evaluated, developing robust measures to collect volunteer data, and interpreting results. For example, one project leader mentioned gaining a “deeper understanding for just how much thought process has to go into deciding what to study, or how to try to put together good questions, or what it is we want to figure out.” Echoing this challenge of crafting evaluation questions, another leader stated, “Asking the question in a really clear way is really important. And I know we all know that. But you know, this was really brought home to me with this project in a very visceral way [sic] that I hadn’t really considered before.”

Third, project leaders pointed to the benefits of an effective system for organizing and managing volunteer data. This includes the scientific data collected by volunteers, as well as data on volunteer attributes (e.g., current skill proficiency, trainings attended, and extent of project participation). Ideally, all of this would be organized within a single curated volunteer-data management system. Many projects likely lack such a system, as captured by this leader who stated, “Holy mackerel, we are missing out on the opportunity to collect key benchmark data points along the way to evaluate this program effectively!” Finally, and perhaps most telling, leaders from 10 of the 15 projects stated that they will continue to use the evaluation instruments and processes from this study, and in some cases, plan to extend them to other citizen science projects in their programs.

IMPORTANCE OF COLLABORATION

In their response to the question regarding final thoughts, 13 of the 15 project leaders spoke highly of the collaborative component of this evaluation work—8 did so without being prompted by the interviewer. While they did not

specifically state how collaboration helped them use their evaluation results, they did share overall benefits. That is, they commented on the benefit of focusing and reflecting for an extended period of time on a mutual challenge with colleagues. They reflected on how the collaboration provided a unique opportunity to interact closely with other leaders working on related but different citizen science projects and to think about the challenge of evaluation from different perspectives. As one project leader noted,

“[It’s] commonplace for people to meet outside of their sector and [sic] share. But that doesn’t often happen outside of conferences...Being able to talk about our programs and think about how do we assess science learning...through embedded assessment has just been incredible.”

Likewise, another stated,

“We get accustomed to the very specific way that our project works...or the way that we are thinking about our data...It’s very helpful to get different perspectives and to think through things with others who are doing similar but not exactly the same things.”

A third leader echoed this, “It’s only when you get a diversity of people that have a diversity of ideas and approaches, I think, that you can come up with some interesting solutions.” Several leaders reported they did not think they could have achieved the same progress and results working on their own or even closely with an evaluator.

Only 2 of the 15 project leaders stated collaboration was not a critical element of their evaluation work. Both had existing in-house expertise and support to tackle evaluation questions and processes. As one of these leaders reported, “There is this robust team of social scientists...that we could draw on for advice and guidance and support. That is not always easily available to a citizen science relatively-resource-limited team.” Their comments may indicate the importance of establishing collaborations earlier rather than later in joint evaluation experiences.

DISCUSSION

In this study, we sought to demonstrate the value of evaluation within the citizen science field. Little has been published on how evaluation processes and findings can support and advance our field, which in turn may limit how practitioners view it. Here, we characterized evaluation use by applying a typology from the evaluation literature to the

citizen science field (see [Table 1](#)). To ensure this research is relevant to the broader field, we examined evaluation use through a practitioner lens by documenting the work of leaders from 15 citizen science projects who were deeply involved in evaluation development and implementation.

In our study, we found that project leaders used evaluation of their citizen science efforts in different and important ways. They gained new and deeper understanding of their volunteers and programming (conceptual use); made critical changes to their projects (programmatic use); shared their evaluation findings to persuade stakeholders, inform the field, and motivate participants (dissemination use); and expanded their attitudes and actions with regard to evaluation (process use). Many of these evaluation uses centered on guiding development or improvements of a project (also called formative evaluation). Specifically, knowledge gains from evaluation prompted project leaders to change their training, revise their protocols, add resources, and even terminate an unproductive project. By contrast, other evaluation uses sought to determine the efficacy or impact of citizen science projects (also called summative evaluation). Through reports, presentations, and publications, the project leaders shared their findings related to skill proficiency with their volunteers, other staff members, practitioners in other programs, funders, researchers, and evaluators.

We believe evaluation use by these project leaders may increase with time. That is, they may need time to consider additional ways that their new understanding and attitudes about volunteers and programming (conceptual use) and evaluation (process use) can be employed to support, improve, and promote their projects. This extended work could include reflecting on new questions revealed by the evaluation, which could provide further insights for individual efforts and potentially the field overall. For example, in the future, one of the project leaders might come up with a way to build on volunteers' everyday skills associated with pattern recognition to help them improve their scientific observation skills. Increased evaluation use over time is supported by [Shaw and Campbell \(2014\)](#), who found that process use rose in the months after their project ended. Our findings related to process use are particularly encouraging, as a deeper understanding of evaluation purpose, methods, and results may help address its limited application across the citizen science field.

We believe it is unlikely that we would have seen this level of evaluation use without the facilitation and collaboration that was central to this study. We purposely provided this facilitation in the context of use. We foregrounded evaluation use by stating from the beginning that we wanted the evaluation results to be useful for the project leaders, and we had them regularly reflect

on use at our group meetings and during our interviews. For example, in a midway meeting with the Secondary Analysis group, we had the project leaders and their data analysts share what they learned about their participants and their skills after re-analyzing existing data collected by their volunteers, along with how they are using the results thus far. Overall, this framing ensured that application of evaluation remained central as the leaders made decisions about what volunteers' science inquiry skills to assess and how to assess them.

Additionally, the leaders in our study actively participated in the evaluation of their projects. The project leaders in the Shared Measures group worked very closely to co-develop and reflect on two different embedded assessments that could be used by multiple citizen science projects. The project leaders in the Secondary Analysis group worked more independently on re-analysis of their existing datasets, but they did meet multiple times to ask questions, compare progress and results, and reflect on impacts and applications. Across both groups, most of our project leaders highlighted how these interactions helped them as they articulated the science inquiry skills that were the focus of their embedded assessments, implemented the embedded assessments of these skills, analyzed and interpreted their findings, and considered ways to use these findings.

We speculate that this deep engagement with the process and resulting data likely had a positive impact on their subsequent use of the findings. Collaborating with stakeholders—those who have a stake in the evaluation or its results such as project leaders, implementing staff, funders, and beneficiaries—is frequently discussed in the evaluation literature ([Brandon and Fukunaga 2014](#)) and is becoming more prevalent in evaluation practice ([Cousins 2020](#)). There is evidence that it promotes different types of evaluation use and particularly process use ([Cousins and Chouinard 2012](#)). We are unaware of other work on this topic within the field of citizen science, but this claim has been substantiated within the context of other informal learning collaboration. For example, [Peterman and Gathings \(2019\)](#) found that the EvalFest collaborative evaluation model was effective at promoting the use of evaluation within science festivals. Similarly, a study of the Nanoscale Informal Science Education Network (called NISE Net) found that their collaborative team-based inquiry approach resulted in museum staff who valued and used evaluation more regularly ([Bequette et al. 2019](#)). [Grack Nelson et al. \(2019\)](#) described the value of building capacity within and across NISE Net, including a diversity in evaluation approaches, a shared appreciation for evaluation, and the use of evaluation to make data-informed decisions.

The individuals who make up the evaluation team—including type of stakeholders (e.g., directors, staff, funders, beneficences)—can influence the use of the resulting findings (Brandon and Fukunaga 2014). The project leaders in our study were well positioned to use their evaluation because all were intimately involved in their citizen science endeavors, and they could influence associated decisions. That is not to say that they did not face challenges associated with evaluation use, such as limited ability to reprogram online data collection systems and insufficient funds to make needed training improvements. The relationship between stakeholders and evaluators is also critical for promoting evaluation use (Cousins and Chouinard 2012), and depends on the evaluator's technical skills and ability to interact effectively with project staff and other stakeholders (King and Stevahn 2012).

CONCLUSION

In discussing evaluation use, Patton (2020) encouraged “help[ing] people in the situation pay attention to their use options, if that seems appropriate and useful” (p. 588). Helping citizen science leaders more fully understand the important and diverse ways that evaluation can support individual projects and the larger field could increase the prevalence of evaluation in citizen science (Schaefer et al. 2021). We offer four recommendations to advance this work. First, we recommend project leaders use the typology described in this research to understand the different ways that evaluation findings can meet their project goals. Second, we encourage project leaders to be actively involved in evaluation processes and to work collaboratively with evaluators. This aligns with Stevahn and King (2016) who offered steps that informal learning evaluators can take to foster evaluation use, including building strong personal relationships with stakeholders that center on using evaluation results, providing structured interactions to help stakeholders interpret and apply results, and allowing the time needed to do this kind of deep thinking and collaboration. Third, we recommend project leaders also collaborate with each other on evaluation efforts to capitalize on shared resources and expertise around evaluation capacity and use. As shown by Grack Nelson et al. (2019) for the NISE NET community, such efforts might be particularly well suited for citizen science projects that are already part of existing networks (e.g., those associated with water quality monitoring or using shared resource such as CitSci.org). Finally, project leaders' emphasis on the importance of collaboration was an unanticipated outcome of our study, and it needs to be explored more systematically. Thus, we recommend future studies that

more deeply explore collaborative evaluation within citizen science, such as the extent of stakeholder involvement and strategies to support collaborative evaluation within the opportunities and bounds of citizen science.

DATA ACCESSIBILITY STATEMENT

The codebook with examples from the dataset are provided in the Appendix. To maintain anonymity and confidentiality of study participants, data from this study cannot be made publicly accessible.

SUPPLEMENTARY FILE

The supplementary file for this article can be found as follows:

- **Supplemental File 1.** Evaluation Use Codebook. DOI: <https://doi.org/10.5334/cstp.482.s1>

ETHICS AND CONSENT

This research was approved by the University of Maryland Institutional Review Board (IRB #1072528).

ACKNOWLEDGEMENTS

We thank the citizen science practitioners who contributed to this work and brought many important insights: Alison Cawood, Ashley Cole-Wick, Brad Schrom, Christine Goforth, Christy Pattengill-Semmons, Corey Jackson, Emma Grieg, Erica Garrouette, Gillian Cannataro, Hillary K. Burgess, Jen Meilinger, Julia Parrish, Julie Vastine, Juniper Simonis, Kevin Bonine, Laura Trouille, Lisi Lohre, LoriAnne Barnett Warren, Megan Mueller, Michael McKelvey, Natalie McNeill, Paul Millhouser, Robert Costello, Sara Fitzsimmons, Scott B. Coughlin, Shelly Grow, and Timothy Jones. We appreciate the support from additional team members (Jenna Linhart, Marshall Illiff, Amy Grack Nelson, and Evelyn Christian Ronning). We also thank our advisors for their helpful suggestions throughout this research (Anne Bowser, Mac Cannady, Joe Heimlich, and Martin Storksdieck).

FUNDING INFORMATION

This material is based upon work supported by the National Science Foundation under Grant No. DRL-171342.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

All authors were involved in the design and data collection of this study. CD led manuscript writing with VDB, KP, AG, TBP, and RBK reviewing and providing extensive edits.

AUTHOR AFFILIATIONS

Cathlyn Davis  orcid.org/0000-0002-0968-4336
University of Maryland Center for Environmental Science, US

Veronica Del Bianco  orcid.org/0000-0001-9088-9488
University of Maryland Center for Environmental Science, US

Karen Peterman  orcid.org/0000-0003-4388-9412
Karen Peterman Consulting, Co, US

Andrea Grover  orcid.org/0000-0003-4082-4138
University of Nebraska Omaha, US

Tina Phillips  orcid.org/0000-0002-5010-6052
Cornell University, US

Rachel Becker-Klein  orcid.org/0000-0002-1456-3491
Two Roads Consulting, US

REFERENCES

Adams, WC. 2015. Conducting semi-structured interviews. *Handbook of practical program evaluation*, 4: 492–505. DOI: <https://doi.org/10.1002/9781119171386.ch19>

Alkin, MC and King, JA. 2016. The historical development of evaluation use. *American Journal of Evaluation*, 37(4): 568–579. DOI: <https://doi.org/10.1177/1098214016665164>

Alkin, MC and King, JA. 2017. Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, 38(3): 434–450. DOI: <https://doi.org/10.1177/1098214017717015>

Becker-Klein, R, Davis, C, Phillips, T, DelBianco, V, Grack Nelson, A and Christian Ronning, E. (in review). Using a shared embedded assessment tool to understand participant skills: Processes and lessons learned. *Citizen Science: Theory and Practice*.

Becker-Klein, R, Peterman, K and Stylinski, C. 2016. Embedded assessment as an essential method for understanding public engagement in citizen science. *Citizen Science: Theory and Practice*, 1(1). DOI: <https://doi.org/10.5334/cstp.15>

Bequette, M, Cardiel, CL, Cohn, S, Kollmann, EK and Lawrence, F. 2019. Evaluation capacity building for informal STEM education: Working for success across the field. *New Directions for Evaluation*, 161: 107–123. DOI: <https://doi.org/10.1002/ev.20351>

Bowser, A, Cooper, C, De Sherbinin, A, Wiggins, A, Brenton, P, Chuang, T-R, Faustman, E, Haklay, M and Meloche, M. 2020. Still in need of norms: the state of the data in citizen science, *Citizen Science: Theory and Practice*, 5(1). DOI: <https://doi.org/10.5334/cstp.303>

Brandon, PR and Fukunaga, LL. 2014. The state of the empirical research literature on stakeholder involvement in program evaluation. *American Journal of Evaluation*, 35(1): 26–44. DOI: <https://doi.org/10.1177/1098214013503699>

Brown, ED and Williams, BK. 2019. The potential for citizen science to produce reliable and useful information in ecology. *Conservation Biology*, 33(3): 561–569. DOI: <https://doi.org/10.1111/cobi.13223>

Bundi, P, Frey, K and Widmer, T. 2021. Does evaluation quality enhance evaluation use? *Evidence & Policy: A Journal of Research, Debate and Practice*, 17(4): 661–687. DOI: <https://doi.org/10.1332/174426421X16141794148067>

Burgess, HK, DeBey, LB, Froehlich, HE, Schmidt, N, Theobald, EJ, Ettinger, AK, HilleRisLambers, J, Tewksbury, J and Parrish, JK. 2017. The science of citizen science: Exploring barriers to use as a primary research tool. *Biological Conservation*, 208: 113–120. DOI: <https://doi.org/10.1016/j.biocon.2016.05.014>

Cousins, JB. 2020. *Collaborative approaches to evaluation: Principles in use* (Vol. 3). Sage Publications. DOI: <https://doi.org/10.4135/9781544344669>

Cousins, JB and Chouinard, JA. 2012. *Participatory evaluation up close: An integration of research based knowledge*. North Carolina: Information Age Publishing, Inc.

D’Ostie-Racine, L, Dagenais, C and Ridde, V. 2016. A qualitative case study of evaluation use in the context of a collaborative program evaluation strategy in Burkina Faso. *Health Res Policy Sys*, 14(37). DOI: <https://doi.org/10.1186/s12961-016-0109-0>

Daigneault, PM. 2014. Taking stock of four decades of quantitative research on stakeholder participation and evaluation use: A systematic map. *Evaluation and program planning*, 45: 171–181. DOI: <https://doi.org/10.1016/j.evalprogplan.2014.04.003>

Froncek, B and Rohmann, A. 2019. “You get the great feeling that you’re being heard but in the end you realize that things will be done differently and in others’ favor”: An experimental investigation of negative effects of participation in evaluation. *American Journal of Evaluation*, 40(1): 19–34. DOI: <https://doi.org/10.1177/1098214018813447>

Grack Nelson, A, King, JA, Lawrence, F, Reich, C, Bequette, M, Pattison, S, Kollmann, EK, et al. 2019. Using a complex adaptive systems perspective to illuminate the concept of evaluation capacity building in a network. *American Journal of Evaluation*, 40(2): 214–230. DOI: <https://doi.org/10.1177/1098214018773877>

King, JA and **Alkin, MC**. 2019. The centrality of use: Theories of evaluation use and influence and thoughts on the first 50 years of use research. *American Journal of Evaluation*, 40(3): 431–458. DOI: <https://doi.org/10.1177/1098214018796328>

King, JA and **Stevahn, L**. 2012. Interactive evaluation practice: Mastering the interpersonal dynamics of program evaluation. Los Angeles, CA: Sage Publications. DOI: <https://doi.org/10.4135/9781452269979>

Miles, MB, **Huberman, AM** and **Saldaña, J**. 2018. *Qualitative data analysis: A methods sourcebook*. Los Angeles, CA: Sage Publications.

National Academies of Sciences, Engineering, and Medicine. 2018. *Learning Through Citizen Science: Enhancing Opportunities by Design*. Washington, DC: The National Academies Press. DOI: <https://doi.org/10.17226/25183>

Patton, MQ. 1997. Toward distinguishing empowerment evaluation and placing it in a larger context. *Evaluation Practice*, 18(2): 147–163. DOI: [https://doi.org/10.1016/S0886-1633\(97\)90019-2](https://doi.org/10.1016/S0886-1633(97)90019-2)

Patton, MQ. 2020. Evaluation use theory, practice, and future research: reflections on the Alkin and King *AJE* series. *American Journal of Evaluation*, 41(4): 581–602. DOI: <https://doi.org/10.1177/1098214020919498>

Peterman, K, **Becker-Klein, R**, **Stylinski, C** and **Grack Nelson, A**. 2017. Exploring embedded assessment to document scientific inquiry skills within citizen science. In: Herodotou, C, Sharples, M and Scanlon, E (eds.), *Citizen Inquiry: Synthesising Science and Inquiry Learning*, 63–82. London: Routledge. DOI: <https://doi.org/10.4324/9781315458618>

Peterman, K, **Del Bianco, V**, **Grover, A**, **Davis, C** and **Rosser, H**. (in review). Hiding in Plain Sight: Secondary Analysis of Data Records as a Method for Learning about Citizen Science Projects and Volunteers' Skills. *Citizen Science: Theory and Practice*.

Peterman, K and **Gathings, MJ**. 2019. Using a community-created multisite evaluation to promote evaluation use across a sector. *Evaluation and Program Planning*, 74: 54–60. DOI: <https://doi.org/10.1016/j.evalproplan.2019.02.014>

Phillips, TB, **Porticella, N**, **Constas, M** and **Bonney, R**. 2018. Articulating and measuring individual learning outcomes from citizen science: A conceptual model. *Citizen Science: Theory and Practice*, 3(2). DOI: <https://doi.org/10.5334/cstp.126>

Schaefer, T, **Kieslinger, B**, **Brandt, M** and **van den Bogaert, V**. 2021. Evaluation in citizen science: The art of tracing a moving target. In: Vohland, K, Land-Zandstra, A, Ceccaroni, L, et al. (eds.), *The Science of Citizen Science*, 495–514. Cham, Switzerland: Springer. DOI: https://doi.org/10.1007/978-3-030-58278-4_25

Shaw, J and **Campbell, R**. 2014. The “process” of process use: Methods for longitudinal assessment in a multisite evaluation. *American Journal of Evaluation*, 35(2): 250–260. DOI: <https://doi.org/10.1177/1098214013506780>

Stevahn, L and **King, JA**. 2016. Facilitating interactive evaluation practice: engaging stakeholders constructively. *Evaluation and Facilitation*, 2016(149): 67–80. DOI: <https://doi.org/10.1002/ev.20180>

Stylinski, CD, **Peterman, K**, **Phillips, T**, **Linhart, J** and **Becker-Klein, R**. 2020. Assessing science inquiry skills of citizen science volunteers: a snapshot of the field. *International Journal of Science Education, Part B*, 10(1): 77–92. DOI: <https://doi.org/10.1080/21548455.2020.1719288>

West, S and **Patemian, R**. 2016. Recruiting and retaining participants in citizen science: what can be learned from the volunteering literature? *Citizen Science: Theory and Practice*, 1(2): 15. DOI: <https://doi.org/10.5334/cstp.8>

TO CITE THIS ARTICLE:

Davis, C, Del Bianco, V, Peterman, K, Grover, A, Phillips, T and Becker-Klein, R. 2022. Diverse and Important Ways Evaluation can Support and Advance Citizen Science. *Citizen Science: Theory and Practice*, 7(1): 30, pp. 1–10. DOI: <https://doi.org/10.5334/cstp.482>

Submitted: 12 December 2021 **Accepted:** 14 June 2022 **Published:** 18 July 2022

COPYRIGHT:

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Citizen Science: Theory and Practice is a peer-reviewed open access journal published by Ubiquity Press.