

## NARST 2020 Paper - Strand 6

### *Embedded Assessment Pursuits: Identifying Important, Relevant, Accessible but Hidden Skills of Citizen Scientists*

Cathlyn Davis Stylinski<sup>1</sup>, Veronica Del Bianco<sup>1</sup>, Tina Phillips<sup>2</sup>, Karen Peterman<sup>3</sup>, Andrea Wiggins<sup>4</sup>,  
Rachel Becker-Klein<sup>5</sup>, and Jenna Linhart<sup>1</sup>

<sup>1</sup>University of Maryland Center for Environmental Science, <sup>2</sup>Cornell Lab of Ornithology, <sup>3</sup>Karen Peterman Consulting Co., <sup>4</sup>University of Nebraska Omaha, <sup>5</sup>Two Roads Consulting

Embedded assessment (EA) is particularly well-suited for evaluating citizen science volunteers' proficiency with science inquiry skills; however it remains uncommon in informal education. Using design-based research, we are examining processes to streamline EA development by leveraging pre-existing data within five citizen science projects to assess skill proficiency. Here, we focus on the critical first step of supporting citizen science project leaders in identifying appropriate skills that are important, relevant, accessible, and potentially hiding in plain sight in their existing data. Our research reveals that the project leaders can bring broad but uncertain conceptualizations of volunteers' skills relevant to their citizen science efforts. These leaders need time and support to refine expansive notions of skill-based outcomes into concrete and clearly defined specifics that can be assessed from their existing data. Our research shows that identifying appropriate skills for EA is a complex multi-step procedure that benefits from supporting oral and written tools. Understanding the processes for developing embedded assessment is valuable for education research in diverse venues.

### **Introduction**

Citizen science offers a learning experience that engages the public in scientific research, and occurs in diverse informal settings including recreational dives, bird-watching excursions, mountain hikes, or anywhere an Internet connection is available. Because of the emphasis on genuine participation in scientific research and unlike most types of informal learning, a key outcome of citizen science is associated with skills of science inquiry (Phillips et al. 2018). Citizen science participants must apply science inquiry skills to contribute to the research. These skills encompass all tasks required to pursue the work of science (NRC, 2012), and can include defining science questions, collecting samples and recording data, conducting data analysis, interpreting data, disseminating conclusions and translating results into action (Stylinski et al. 2020, Shirk et al. 2012).

Assessment of volunteers' proficiency with science inquiry skills could improve data quality, raise confidence in volunteers' efforts, and support learning goals (NASSEM 2018). Embedded assessment (EA) may be a particularly effective method for determining participant skill proficiency in citizen science and other informal education settings where direct testing is a poor fit (e.g., NRC 2009), disruptive to the learning experience, and could possibly reduce

participant interest and retention. As the name suggests, EAs can be integrated seamlessly into the learning experience with participants demonstrating their proficiency with targeted skills in unobtrusive ways (Becker-Klein et al., 2016). For example, an EA could be a game that requires a certain level of skill proficiency (e.g., species identification), and is played during a training session. Alternatively, it could include an analysis of learners' artifacts, such as the data that volunteers submit for a citizen science project (e.g., list of bird species observed at a particular location). However, despite benefits of embedded and other types of assessments, few citizen science projects directly assess volunteers' skill proficiency (Stylinski et al. 2020). At least one challenge is that EAs require a deliberate and intensive process that is beyond the capacity of project staff (Peterman et al. 2017).

How can we reduce the complexity of developing EAs and thus increase their use? Our team is researching two strategies to streamline the development process of authentic and performance-based EAs within citizen science. In one case, we are working with 10 citizen science project leaders to develop shared EA measures that can be broadly applied. In the second case, we are working with five additional leaders to re-analyze their existing datasets (science data collected by their volunteers) to examine skill gains among these volunteers. In this NARST paper, we will focus on the latter case and the first critical step to help citizen science leaders articulate science inquiry skills that are targeted by their projects and that are important, relevant, accessible, and potentially hiding in plain sight in their existing data.

## **Design**

We are using a design-based research methodology for this study, collaborating with researchers and practitioners (project leaders and their data analysts) of ongoing citizen science projects. This approach allowed for dual study goals for the research: (1) meeting projects' needs to understand their volunteers' skill gains, and (2) researching opportunities and challenges during the decision-making process for identifying those skills (Barb & Squire, 2004).

The five projects in our study vary in research topic, geographical location and participation settings (online vs. in the field):

- Coastal Observation and Seabird Research Team (COASST)
- Alliance for Aquatic Resource Monitoring (ALLARM)
- Front Range Pika Project
- Reef Environmental Education Foundation (REEF)
- Gravity Spy

All five projects target adult volunteers, and have numerous participants who submit or classify data multiple times—both are necessary to explore skill changes in existing data. The first four have volunteers gather data from field excursions and submit their data online. Gravity Spy occurs only online and has volunteers classify visualizations of gravitational wave data.

In accordance with design-based research, we conducted multiple cycles of testing and refinement as we explored and documented processes involved with analyzing existing data to assess volunteers' skill gains. We collected data with recorded and transcribed interviews, notes from in-person and virtual meetings with all project leader-analyst pairs, and artifacts collected from each pair. Data for this paper come from transcripts of baseline interviews (Reflection Session) and three artifacts (Decision Trees, Ranking Worksheet, Skill Hierarchy).

## **Analysis and Findings**

Our design-based research process revealed three key steps and associated tools that proved useful for citizen science project leaders identifying science inquiry skills that are important, relevant, and accessible to them. These steps consist of (1) an initial Reflection Session, (2) a Decision Tree to identify relevant skills that were supported with existing data and a Ranking Worksheet to prioritize selection of focal skills, and (3) a Skill Hierarchy that highlights the relationships between tasks and skills embedded within citizen science participation protocols. Descriptions and findings associated with each are described below.

### *Step 1: Reflection Session*

Our research team conducted a virtual project Reflection Session separately with each of the five citizen science project leaders to discuss the project, its protocols, and available data sources and current data cleaning and analysis procedures. These sessions helped our research team better understand project goals, resources, and science inquiry skills required of volunteers. We used this information to conceive and craft the Decision Trees and Ranking Worksheets described in the next step. For example, the REEF project leader recounted in detail how her volunteers (divers) survey fish, which helped our team understand that there are multiple starting points for participation that could affect whether and how a volunteer's species identification skills changed over time. That is, a volunteer's first dive might be a solo or social event or might have more or less guidance from other more experienced divers, which would all have an impact on identification skills.

### *Step 2: Decision Tree and Ranking Worksheet*

As noted, the interviews prompted us to create written guides that would help advance and organize the project leaders thinking about possible science inquiry skills for their reframing analysis. We created a Decision Tree graphic as a way to consistently capture conversations on targeted science inquiry skills of each of the five projects. Using branching logic, the Decision Tree begins with a list of potential relevant skills, and uses a stepwise process to eliminate items that are not actually skills and then those that are absent in the dataset. The remaining skills are potentially viable for evaluation. Our team drafted a Decision Tree for each project based on the reflection session, and project leaders reviewed their tree to validate them and make necessary adjustments.

For some projects, the Decision Tree resulted in too many possible skills to analyze. To select the focal skills for the secondary analysis among these options, we had project leaders review the skills that were identified in the Decision Tree using a Ranking Worksheet, which challenged them to think more carefully about their existing database and the importance of each identified skill to their overall project development priorities. For example, a skill like walking a straight line might be relevant and accessible in the database but not have high importance for a protocol focused on data accuracy. Together, the Decision Tree and Ranking Worksheet helped project leaders make an evidence-based and goal-oriented selection of inquiry skills that could have the greatest benefit to project operations.

We will use COASST to illustrate this process. This 19-year-old project involves over a thousand volunteers in California, Oregon, Washington, and Alaska in monitoring marine ecosystems. After extensive training, volunteers document and attempt to identify seabird

carcasses along a selected stretch of beach. The COASST project's initial list of skills in the Decision Tree consisted of 35 items, which included "make body measurements," "species identification accuracy," and "photos." They then removed items that were not volunteer skills, such as "time since training" and "number of intact birds identified." Next they removed skills not accessible in their existing datasets such as "beach search effort" and "non-birds," leaving 21 possible skills. Guided by the Ranking Worksheet and their own expertise, the COASST team selected one skill for their EA, "make body measurement," which they broken down into three subskills: "measure bill," "measure wing," and "measure feet."

After identifying their subskill(s), project leaders and their data analysts cleaned and organized their existing data to allow assessment of these skills. Broadly speaking, all five project leaders selected "accurate data collection" as the focus of their reframing analysis, although this focus supported different goals across the projects (e.g., "measure feet" for COASST and "pattern recognition" for Gravity Spy). Each project identified variables in their existing records that could serve as evidence of the accuracy of the selected skill(s) or a proxy for accuracy. For example, REEF project volunteers are instructed to enter data only for the species of fish that they can confidently identify on a dive. Consequently, the addition of new species not previously reported indicates improved species identification skills of a volunteer. Thus, the project leader and her data analyst decided to analyze accumulation of species over time in each volunteer's records as a proxy for accuracy.

### *Step 3: Skill Hierarchy*

These experiences with project leader-analyst pairs led our team to create a third tool to support understanding and connection of selected subskills to each project's data collection protocol. To clarify we developed a Skill Hierarchy graphic that illustrates connections between finer resolution subskills and their overarching skill associated with data collection accuracy. Project leaders filled in the Hierarchy, and discussed them with our research team and other project teams. Returning to the COASST example, their overarching skill associated with data collection accuracy is "identify and document seabird carcasses." Skills necessary to achieve this include "photograph carcass" and "make body measurements." The subskills for "make body measurements" are "measure bill," "measure wing," and "measure foot."

While these Skill Hierarchy graphics did not lead to changes in their selected skill, they provided greater clarity and understanding for the project leaders and our research team because they illustrated the hierarchy of skills. This helped highlight subskills that were particularly important because they support successful completion of multiple skills, or which had potential for multiple measures that could be used for verification of the assessment. The Skill Hierarchy graphic, along with the Decision Trees, has also served as a key communication tool that has helped projects present their work to each other and to other practitioners (e.g., Peterman et al. 2019).

## **Discussion**

This research is part of our larger effort to discover the processes for developing broadly applicable EA methods that can be used in citizen science and other informal science education settings. Here, our process uses existing data submitted by citizen science volunteers, which can provide "traces" of individual growth of volunteers' skills (i.e., reveal what is hiding in plain

sight) through secondary analysis to surface these traces. For this paper, we focused on the first phase of this EA process; that is, supporting project leaders in identifying important, relevant, and accessible data representing the skills of citizen science volunteers as the focus of the EA. Because this EA strategy uses existing data, it has the potential to reduce barriers to adoption of this form of assessment by other citizen science projects.

Our research of the EA development process reveals that the project leaders can bring broad but uncertain conceptualizations of science inquiry skills relevant to their citizen science efforts. They need time and support to refine these expansive notions of skill into concrete and clearly defined specifics that could be assessed with their data. Time was also necessary to identify proxies in their existing records for selected science inquiry skills and other related variables. For example, data necessary to explore volunteers' ability might be collected but not stored in a consistent way. The Front Range Pika Project estimated spending 200 hours on cleaning and organizing their data for the analysis. Time required to reflect more fully on targeted project outcomes associated with volunteer skills, to search for proxies of these skills in the data, and to collect, clean, organize, and prepare data for analysis could be reduced with clear planning at the project inception or at key development stages. Indeed, some of our project leaders identified skills that could readily be evaluated in the future, contingent on updates to their current organizational processes for capturing information about volunteers or their participation.

In summary, our research reveals that identifying appropriate science inquiry skills for EA is a complex multi-step procedure that can be supported with appropriate tools. We found that starting with an in-depth Reflective Session of project activities can help leaders consider measures of skills that might be included in their existing datasets. A Decision Tree coupled with a Ranking Worksheet can filter these skills into those that are important, relevant, and accessible within the project, while a Skill Hierarchy can illustrate links between subskills and overall project protocols. Overall, this process can be useful for other informal science endeavors seeking to articulate and assess participants' skills.

## References

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, 13(1), 1-14.
- 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. DOI:<http://doi.org/10.5334/cstp.15>
- National Academies of Sciences, Engineering, and Medicine (2018). *Learning Through Citizen Science: Enhancing Opportunities by Design*. The National Academies Press. DOI: <https://doi.org/10.17226/25183>
- National Research Council (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. The National Academies Press. DOI: <https://doi.org/10.17226/12190>
- Peterman K., Becker-Klein R., Stylinski C. & Grack Nelson A. (2017). Exploring Embedded Assessment to Document Scientific Inquiry Skills within Citizen Science. In C. Herodotou, M.

Sharples, & E. Scanlon (Eds.), *Citizen Inquiry: A Fusion of Citizen Science and Inquiry Learning* (pp. 63-82). Routledge.

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. DOI: <http://doi.org/10.5334/cstp.126>

Shirk, J.L., Ballard, H.L., Wilderman, C.C., Phillips, T., Wiggins, A., Jordan, R., ... & Krasny, M.E. (2012). Public participation in scientific research: a framework for deliberate design. *Ecology and Society*, 17(29). DOI: <http://www.jstor.org/stable/26269051>

Stylinski C., Peterman K., Phillips T., Linhart J. & Becker-Klein R. (2020) A review of assessment of volunteers' science Inquiry skills targeted in citizen science. *International Journal of Science Education Part B*, 10(1), 77-92.