

ESSAY

Embedded Assessment as an Essential Method for Understanding Public Engagement in Citizen Science

Rachel Becker-Klein*, Karen Peterman† and Cathlyn Stylinski‡

Citizen science is an important way of engaging a broad range of audiences in science inquiry by participating in research that asks novel questions and unearths new knowledge and new questioning. Though citizen science projects are quite diverse in their scientific pursuits, all projects share the common element of involving volunteers directly in some aspect of science inquiry. Thus, it is essential for citizen science projects to determine their participants' capacity to learn and successfully perform science inquiry skills, such as making scientific observations, collecting and analyzing data, and sharing findings. Such skill gains are essential to (a) ensure high quality data that can be used in meaningful scientific research, and (b) achieve broader goals such as developing a participant's identity as a contributor to science. However, we do not yet fully understand how improvement in participants' inquiry skills through citizen science advances our knowledge of public engagement with science. In this essay, we offer embedded assessment as an effective method to capture participant skill gains, and encourage citizen science leaders, evaluators, and researchers to develop authentic methods that address the complexities of measuring skill development within the context of citizen science.

Keywords: Embedded Assessment; Science Skills

Citizen science is an important approach for engaging non-scientists in research that asks novel questions and unearths new knowledge and new questioning (Bonney et al. 2009; Dickinson and Bonney 2012). Though citizen science projects are diverse in their scientific pursuits, all share the common element of involving volunteers directly in some aspect of science inquiry. Citizen science participants need to develop robust science inquiry skills to (a) ensure high quality data that can be used in meaningful scientific research and (b) achieve broader goals such as developing a participant's identity as a contributor to science. Consequently, citizen science projects must determine their participants' capacity to learn and successfully perform various inquiry skills such as making scientific observations, collecting and analyzing data, and sharing findings. This essay discusses embedded assessment (EA) as an effective method for capturing and measuring gains in participant skills.

As defined by Wilson and Sloane (2000: 184), EAs are "opportunities to assess participant progress and performance that are integrated into instructional materials

and virtually indistinguishable from day-to-day [program] activities." As such, EAs allow learners to demonstrate their science competencies through tasks that are integrated seamlessly into the learning experience itself. Here we explore development of science inquiry skills in the context of citizen science and offer examples of how EAs could be applied to understand skill gains in ways that do not interfere with the free-choice nature of learner-driven experiences.

Science Inquiry and Citizen Science

The extent to which a project includes various science inquiry skills has been used to define models of public participation in citizen science (Bonney et al. 2009; Shirk et al. 2012). Specifically, *contributory* citizen science projects primarily limit the involvement of volunteers to the collection of data (usually through observation, identification, and monitoring), while *collaborative* citizen science projects expand on data collection to include developing explanations, designing data collection methods, and analyzing data. *Co-created* projects include the inquiry skills from the other citizen science models while also providing opportunities for the public to define the research question, gather information to support the study rationale, interpret data to make conclusions, disseminate results, and pose questions for further study.

Most current citizen science projects are considered contributory in nature. However, as the citizen science

* PEER Associates, US

† Karen Peterman Consulting Co., US

‡ University of Maryland, Center for Environmental Science, US

Corresponding author: Rachel Becker-Klein
(rachel@peerassociates.net)

field grows, more collaborative and co-created projects are being developed in which the public plays more active roles in the design, implementation, analysis, interpretation, and dissemination of research. Such a shift in project design requires participants to develop a deeper understanding of the science process, while also supporting their “use of critical thinking skills in their everyday lives and their use of science in relevant contexts, such as Earth stewardship and scientifically informed decision making” (Dickinson et al. 2012).

The current array of tested methods for demonstrating the effects of citizen science on skills is limited and consists largely of self-report scales (Phillips et al. 2012). This assessment challenge extends to other informal learning activities, for which traditional (formal) assessment measures offer a poor fit (National Research Council [NRC] 2009; Fenichel and Schweingruber 2010). As such, we believe that the citizen science community is in need of more innovative and performance-based methods to collect data about skill-based outcomes. Taking steps in this direction would enhance our understanding of the ability of citizen science to achieve both educational and scientific goals and could serve as a model for assessing project outcomes in the larger field of informal science, technology, engineering, and math (STEM) learning.

Embedded Assessment and Citizen Science

Because they require that participants demonstrate their skills, EAs offer an innovative way to understand the impacts of citizen science participation. EA methods can include *performance assessments*, in which participants do something to demonstrate their knowledge and skills (e.g., scientific observation), and *authentic assessments*, where the learning tasks mirror real-life problem-solving situations (e.g., the specific data collection techniques used in a citizen science project; Rural School and Community Trust 2001; Wilson and Sloane 2000). EAs can be used alongside more traditional research and evaluation measures and also may be useful for measuring volunteers’ skill development across time. An analogy is the current method for obtaining a driver’s license in the United States, which requires both a written exam and driving test. The written exam provides information about whether someone knows the rules of the road, while the driving test demonstrates actual driving skills. Just as it would be unwise to judge a person’s driving ability based solely on one or the other of these evaluation components, it is equally unwise to measure the impacts of citizen science participation through self-reported methods alone.

EAs, especially performance-based EAs, are particularly appropriate for free-choice learning programs—such as citizen science—which require the use of skills throughout their implementation and that have educational or practical outcomes related to skill gains among volunteers. One of the challenges cited by the NRC was the fact that surveys, tests, and assessments do not feel authentic to informal learning contexts (NRC 2009). EAs, by contrast, are integrated into the learning experience itself and thus authentic in every way.

Although EAs have not been widely used to advance understanding of the impact of participatory informal science education efforts, interest in their use is growing. For example, since 2006 EA has been featured regularly at STEM and evaluation conferences including presentations at the American Evaluation Association conferences, the National Educational Computing Conference, and the Out of School Time Conference (Becker-Klein et al. 2014; Na’im et al. 2009; Peterman 2006; Peterman and Muscella 2007; Peterman et al. 2009). Most recently, the NRC created a consensus report on how to best assess the performance expectations articulated in the Next Generation Science Standards (NGSS), which describe skills as science practices (NRC 2014). The authors emphasize the importance of creating assessments that are able to capture three-dimensional science learning to adequately assess students’ mastery of performance expectations. They specifically refer to the promising development of classroom-embedded assessments as “assessment tasks that have been designed to be integral with classroom instruction” (NRC 2014: 4). Finally, performance-based assessments were mentioned in the 2011 consensus report on *Learning Science through Computer Games and Simulations* (NRC 2011), which examined the role that technology (specifically computer gaming and simulations) plays in the assessment of student learning.

Embedded Assessment Examples

EAs can take many forms and can be used in a variety of settings. The essential defining feature is that EAs document and measure participant learning *as a natural component of the program implementation* and often as participants apply or demonstrate what they are learning. EAs must be created through a deliberate and intensive process of development, including the involvement of both program staff and evaluators or education researchers. Scientific data quality assessments and validation provide additional opportunities to integrate EA into citizen science projects while also contributing critical information for science research. Several examples of EA are described below.

Performance-based embedded assessment

Co-author Peterman collaborated with Deborah Muscella of the Girls Get Connected Collaborative (GGCC) to create a series of games that were used to document students’ technology and data-collection skill gains for the National Science Foundation-ITEST project Technology at the Crossroads (DRL-0423588). In this project, students participated in a summer camp experience that included the Greater Boston Urban Forest Inventory (GBUFI), a citizen science project dedicated to identifying all the trees in the city of Boston.

The project evaluation consisted of a number of games that were implemented as part of a field day competition. One such game included a competition to see which team could inventory trees the fastest and most accurately. The GGCC team conducted an inventory of several trees that were close to the summer camp. The evaluation team then had groups of students inventory the same trees as part

of the competition. Students conducted their first inventory after one week of camp and a second inventory at the end of the two-week camp experience. They were timed as they conducted their inventories and their data were compared to those collected by the GGCC staff. Time and accuracy rates were also compared from the first to the second week of camp and to volunteer data made available by the GBUFI. The results showed statistically significant increases in data accuracy from the first to the second week of camp and a statistically significant decrease in the amount of time it took to collect the data. In addition, the accuracy of the data and the time needed to collect the data matched or improved on the accuracy and time provided by the GBUFI (Peterman and Muscella 2007). This activity was developed specifically for the purposes of project evaluation.

Data quality assessment

Another option for EA is to take advantage of data-validation procedures that a project already has in place. Several recent articles offer strategies for validating citizen science data quality (Bonter and Cooper 2012; Dickinson and Bonney 2012; Havens et al. 2012). For example, Bonter and Cooper (2012:36) state that “one promising development for addressing the need [to identify plausible but erroneous observations] is the creation of online games or quizzes.” Such games, whether online or in person, could be used not only to improve data quality but also could be a form of EA that assesses participants’ skills in scientific observation and data collection. In this way, citizen science data can serve multiple purposes—improving quality assurance/quality control, engaging volunteers in project activities, guiding project revision to deepen participants’ understanding of science, and tracking the development of participants’ observation skills over time through EA.

As an example, Bonter and Cooper (2012) describe a data validation protocol developed for Project FeederWatch. This protocol was designed to measure participants’ monitoring skills and accuracy with the intention of increasing confidence in the data collected. The project created a set of filters embedded into the FeederWatch website that resulted in a checklist of “allowable” species for the area reported. The authors note “when data violated the smart filter criteria, the submission was flagged and the participant immediately shown an error message informing them that the observation was unusual (p. 305).” In

this way, participants’ observations were checked and the data were deemed more valid. If smart filter data for individual participants were tracked across time, these data could be used to demonstrate skill gains associated with the project—a form of EA using online technology.

Embedded assessment with accompanying rubric

Co-author Becker-Klein (2011) developed a technology-based EA in collaboration with Bob Coulter from the Missouri Botanical Garden and Eric Klopfer from the Massachusetts Institute of Technology to evaluate a project funded by the National Science Foundation’s ITEST program called Community Science Investigators – CSI (DRL-0833663). The goal of this assignment was to measure student skills in working with a technology called Augmented Reality (AR). Students in an out-of-school-time program were challenged to create an AR game after learning the technology, and an accompanying rubric (see below) was developed to assess learners’ skill in the tasks assigned. Project leaders and evaluators collaborated to determine criteria for what constituted a “good AR game” that would demonstrate what participants had learned (Table 1). Becker-Klein used a recording sheet to document the quality of student work. One of the criteria related to a category called Player Experience, which stated that “students have put thought into what the player will be doing and how a new player would experience the game.” Results indicated that the challenge activity was doable and appropriate for students with at least some experience with AR technology. The task and scoring rubric were sensitive enough to detect differences in participants’ skill level with this technology. With further testing and validation, this tool likely could be used to assess participants’ skills in this area.

Importance of Common Methods for Citizen Science Assessment

Another important aspect to consider is the adoption of common measures for assessing science inquiry skills. For example, a report on the need for systematic assessment in informal learning environments published by researchers from the Program in Education, After School & Resiliency at Harvard University and McLean Hospital (Hussar et al. 2008) concluded that, while many tools to assess science interest and skills do exist, there is a need for new tools in the field that can be used across multiple programs and

	Excellent	Proficient	Developing	Needs Attention
Player Experience	The AR game is extremely intuitive, with the key steps well thought out and easy for a new player to follow. Where a player will start is abundantly clear, and it will be simple for them to choose where to go.	It is pretty clear how a new player should play this AR game. There are a couple of different places that a player might start, and then it will be simple for them to choose where to go.	Overall, it is fairly clear how a new player should play this AR game, but there are some fuzzy pieces of the game. There are several different places that a player might start the game, and it is less clear how a player will choose where to go from the starting point.	The AR game is not at all easy to play, and it is unclear to a new player what he/she is supposed to do in the game. The starting points and where to go from these points are not specified.

Table 1: Rubric for Assessing Augmented Reality (AR) Challenge Activity.

which could contribute to creating a stronger evidence base for the impact of informal science programs. The desire to create sharable measures for informal science education also is exemplified by a recent online forum of informal learning practitioners and researchers hosted by the Center for Advancement of Informal Science Education (<http://informalscience.org/community/groups/forum/CAISEvaluationinISEInitiative/viewforum/56/>).

Following this idea, common measures are in development within the field of citizen science. Examples are the tools and scales developed through a project called DEVISE, funded by the National Science Foundation (DRL-1010744), which are intended for use by environmentally focused citizen science projects (Phillips et al. 2014). The DEVISE scales include generic and customized scales intended to measure a range of participant outcomes: Interest, self-efficacy, motivation, skills, and behaviors. As such, the DEVISE scales are a significant asset to the field of citizen science specifically and the field of informal learning overall. However, while skill gains are included in the DEVISE instruments, they are limited to self-reported measures and thus do not document a participant's performance of those skills within the context of citizen science (T. Phillips, personal communication, March 2015). EAs have the potential to fill this gap.

Challenges of Embedded Assessment

Development of EAs pose some challenges that have so far prevented their widespread adoption in citizen science specifically and informal science education more generally. Much of what we know about EAs is based on research that was conducted on the use of performance assessments in formal learning environments. Research to document the benefits and challenges of performance assessments did not begin in earnest until the 1990s (Lawrenz et al. 2001; Ruiz-Primo and Shavelson 1996), and debates are ongoing about whether performance assessments can truly be a practical assessment solution (Gorin and Mislavy 2013; Gott and Duggan 2002; Roberts and Gott 2006). For the successful application of EA to citizen science, three significant challenges must be addressed: (1) the lack of a standard EA development process; (2) the tension between validity and reliability; and (3) the dearth of professional development related to EA.

Lack of a standard embedded assessment development process

A primary challenge preventing wider adoption of EAs is the lack of a framework for developing them. By their very nature, EAs must be customized to the content area of each project. They take time to develop, often require individual administration, and can be complex to score (Johnson et al. 2009; Stufflebeam 2001). To combat the development challenge, Solano-Flores and colleagues (Solano-Flores et al. 1999) created a task shell that was used to create performance assessments for science students. A similar process could potentially guide the development of EAs for science outcomes that are common to citizen science.

Methodologists recently have begun to develop a standard process for creating performance assessments

with explicitly articulated procedures that include identification of the need for an assessment and creation of a reliable and valid measure (Johnson et al. 2009). Some processes also offer standardized guidance on the creation of rubrics (Educational Testing Service 2006). We suggest that a similar set of exemplary practices, tips, and a standardized process for creating EAs would aid evaluators in creating and using such methods.

Tension between validity and reliability

Some researchers have claimed that EAs may be more valid but less reliable than traditional tests. Gipps (1995) exemplified this tension by describing a reading assessment for 7-year olds. In this case, the children read aloud from books of their own choosing and then were asked questions about the characters and their actions, an assessment that was quite high on content and construct *validity* because the materials used for the assessment task were authentic to the students' classroom experiences. However, the fact that many children chose familiar books made their task easier, which decreased the *reliability* of the assessment.

Research on the coding of students' science lab notebooks as a performance assessment may be particularly relevant to the citizen science field. Ruiz-Primo, Baxter, and Shavelson (1993) concluded that notebook scores served as both a reliable and a valid surrogate for performance assessments. This example also meets our definition of EA given that it is a natural part of the learning process that takes place in many science classrooms. Because many citizen science projects involve data collection procedures and forms that have similarities to school lab notebooks, this research has the potential to serve as a model for coding datasheets as a measure of citizen science participant data collection procedures and/or data quality. If coding schemes are created with both validity and reliability in mind, this type of strategy may be able to provide rigorous data that would contribute to our growing understanding of public engagement through citizen science.

Dearth of professional development related to embedded assessments

Few evaluators or researchers are comfortable with the practice of creating and using innovative methods if they do not have prior experience or training; unfortunately, trainings to aid in developing and using EAs are lacking. For example, a search through past conferences of the American Evaluation Association yielded only a few presentations and no demonstrations or professional development workshops about EA. Evaluators and researchers who wish to develop EAs are therefore largely on their own unless they have the fortune to find and subsequently work with others who are engaged in creating such measures. While this collaborative cross-project model is happening more often, it still is not common in the field.

Conclusion

There is a strong need within citizen science to measure science inquiry skill gains in order to understand the impact of citizen science projects on their participants. EA

provides a valuable way of measuring skill gains in context, something that is sorely needed within citizen science and other free-choice learning opportunities. EAs can complement traditional evaluation and research methods without undermining the voluntary nature of learning in informal science contexts. However, there are significant challenges to developing and administering EAs, including questions about the reliability of the method and the lack of evaluator and researcher training in this area.

Despite these hurdles, we believe that EA is a critical method for citizen science and the broader informal science education field. We are beginning to tackle many of these hurdles through an NSF Advancements in Informal Science Learning (AISL) grant. Our *Embedded Assessment for Citizen Science* (EA4CS; DRL-1422099) project will examine the inquiry skills common in citizen science projects and how those skills are or could be assessed. Our project will also explore the development process of EAs for three distinct citizen science projects. From this work, we will initiate a process for developing EAs that can be both reliable and valid, which meets the needs of the case study projects, and that has the potential to generalize to the broader citizen science community. We hope that this project will demonstrate the utility of EA as a valuable method that can be used across projects that share inquiry skills.

In the meantime we encourage citizen science leaders, evaluators, and researchers to help us push the envelope by thinking critically about the inquiry skills fostered by their citizen science projects and ensuring that those skills are measured as part of evaluation and research plans. Citizen science leaders should consider whether their projects include practices that could be used as an EA of skill development and, if so, take advantage of those systems for evaluation and research purposes. In cases where existing practices are untenable for evaluation or research, we encourage science evaluators and researchers to develop authentic methods that address the complexities of measuring skill development within the context of citizen science. Finally, given the critical role that inquiry skills play in the success of citizen science projects and research, we invite citizen science evaluators and researchers to share these experiences broadly with the citizen science community in an effort to highlight the valuable role that citizen science can play in engaging the public with science.

Acknowledgements

We thank Leslie Goodyear, Chris Lysy, Martin Storksdieck, and Tina Phillips for reviewing early versions of this manuscript. This work was supported by the National Science Foundation under Grant number DRL-1422099.

Competing Interests

The authors declare that they have no competing interests.

References

Becker-Klein, R., 2011. Using student performance assessment as an evaluation tool. A presentation at the 2011 National Science Foundation ITEST Summit, Washington, DC.

- Becker-Klein, R., Stylinski, C., Peterman, K. and Phillips, T., 2014. Using embedded assessment to measure science skills within STEM education. A think tank presented at the annual meeting of the American Evaluation Association, Denver, Colorado.
- Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J. and Wilderman, C.C., 2009. *Public participation in scientific research: Defining the field and assessing its potential for informal science education*. A CAISE Inquiry Group Report. Washington, DC: Center for Advancement of Informal Science Education (CAISE).
- Bonter, D.N. and Cooper, C.B., 2012. Data validation in citizen science: a case study from Project Feeder-Watch. *Frontiers in Ecology and the Environment*, 10(6): 305–307. DOI: <http://dx.doi.org/10.1890/110273>
- Dickinson, J.L. and Bonney, R., 2012. *Citizen science: Public participation in environmental research*. Ithaca, NY: Cornell University Press.
- Dickinson, J.L., Shirk, J., Bonter, D.N., Bonney, R., Crain, R.L., Martin, J., Phillips, T. and Purcell, K., 2012. The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, 10(6): 291–297. DOI: <http://dx.doi.org/10.1890/110236>.
- Educational Testing Service, 2006. Metarubric: Criteria for judging the quality of rubrics – a rubric for rubrics. Retrieved from: <http://www.pfsd.com/uploads/Metarubric.pdf>.
- Fenichel, M. and Schweingruber, H.A., 2010. *Surrounded by science: Learning science in informal environments*. Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Gipps, C., 1995. Reliability, validity, and manageability in large-scale performance assessment. In: Torrance, H. (Ed.) *Evaluating authentic assessment: Problems and possibilities in new approaches to assessment*. Open University, pp. 105–123.
- Gorin, J., and Mislevy, R.J., 2013. (September) Inherent measurement challenges in the next generation science standards for both formative and summative assessment. In: *Invitational research symposium on science assessment*, Washington, DC.
- Gott, R. and Duggan, S., 2002. Problems with the assessment of performance in practical science: which way now? *Cambridge Journal of Education*, 32(2): 183–201. DOI: <http://dx.doi.org/10.1080/03057640220147540>
- Havens, K., Vitt, P. and Masi, S., 2012. Citizen science on a local scale: the Plants of Concern program. *Frontiers in Ecology and the Environment*, 10(6): 321–323. DOI: <http://dx.doi.org/10.1890/110258>
- Hussar, K., Schwartz, S., Boiselle, E. and Noam, G.G., 2008. *Toward a systematic evidence-base for science in out-of-school time: The role of assessment*. Study prepared for the Noyce Foundation. Cambridge, MA.
- Johnson, R.L., Penny, J.A. and Gordon, B., 2009. *Assessing performance: Designing, scoring, and validating performance tasks*. New York, NY: The Guilford Press.
- Lawrenz, F., Huffman, D. and Welch, W., 2001. The science achievement of various subgroups on alternative

- assessment formats. *Science Education*, 85(3): 279–290. DOI: <http://dx.doi.org/10.1002/sce.1010>
- Na'im, A., Harris, D., Hilliard-Clark, J., Lowes, S., Nolte, P., Peterman, K. and Suter, L., 2009. (July) *Evaluation tips and strategies: Experiences of ITEST evaluators*. A panel presented at the 2009 Annual National Educational Computing Conference in Washington, DC.
- National Research Council, 2009. *Learning science in informal environments: People, places, and pursuits*. Committee on Learning Science in Informal Environments. Washington, DC: National Academies Press.
- National Research Council, 2011. *Learning Science Through Computer Games and Simulations*. Committee on Science Learning: Computer Games, Simulations, and Education. Washington, DC: The National Academies Press.
- National Research Council, 2014. *Developing Assessments for the Next Generation Science Standards*. Committee on Developing Assessments of Science Proficiency in K-12. Washington, DC: The National Academies Press.
- Peterman, K., 2006. (April) *Technology at the crossroads: A youth-based science and technology program with an innovative approach to evaluation*. A panel chaired and presented at the American Educational Research Association's Annual Meeting in San Francisco, California.
- Peterman, K., Kaminsky, A. and O'Sullivan, R., 2009. (October) *What counts as learning in OST and why?* A panel presentation at the OST STEM: Building Experience, Building Bridges Conference. Washington, DC
- Peterman, K. and Muscella, D., 2007. (November) *Games as embedded assessments*. A panel presentation at the American Evaluation Association Annual Meeting. Baltimore, Maryland.
- Phillips, T., Bonney, R. and Shirk, J.L., 2012. What is our impact? Toward a unified framework for evaluating outcomes of citizen science participation. In: Dickinson, J.L. and Bonney, R. (Eds.) *Citizen science: Public participation in environmental research*. Ithaca, NY: Cornell University Press, pp. 82–96. PMCID: PMC3523681.
- Phillips, T.B., Ferguson, M., Minarchek, M., Porticella, N. and Bonney, R., 2014. *User's guide for evaluating learning outcomes in citizen science*. Ithaca, NY: Cornell Lab of Ornithology.
- Roberts, R. and Gott, R., 2006. Assessment of performance in practical science and pupil attributes. *Assessment in education*, 13(1): 45–67. DOI: <http://dx.doi.org/10.1080/09695940600563652>
- Ruiz-Primo, M.A., Baxter, G.P. and Shavelson, R.J., 1993. On the stability of performance assessments. *Journal of Educational Measurement*, 30(1): 41–53. DOI: <http://dx.doi.org/10.1111/j.1745-3984.1993.tb00421.x>
- Ruiz-Primo, M.A. and Shavelson, R.J., 1996. Rhetoric and reality in science performance assessments: An update. *Journal of Research in Science Teaching*, 33(10): 1045–1063. DOI: [http://dx.doi.org/10.1002/\(SICI\)1098-2736\(199612\)33:10<1045::AID-TEA1>3.0.CO;2-S](http://dx.doi.org/10.1002/(SICI)1098-2736(199612)33:10<1045::AID-TEA1>3.0.CO;2-S); [http://dx.doi.org/10.1002/\(SICI\)1098-2736\(199612\)33:10<1045::AID-TEA1>3.3.CO;2-Q](http://dx.doi.org/10.1002/(SICI)1098-2736(199612)33:10<1045::AID-TEA1>3.3.CO;2-Q)
- Rural School and Community Trust, 2001. *Assessing student work*. Retrieved From: http://www.ruraledu.org/user_uploads/file/Assessing_Student_Work.pdf.
- Shirk, J.L., Ballard, H.L., Wilderman, C.C., Phillips, T., Wiggins, A., Jordan, R. and Bonney, R., 2012. Public participation in scientific research: a framework for deliberate design. *Ecology & Society*, 17(2): 207–227. DOI: <http://dx.doi.org/10.5751/ES-04705-170229>
- Solano-Flores, G., Javanovic, J., Shavelson, R.J. and Bachman, M., 1999. On the development and evaluation of a shell for generating science performance assessments. *International Journal of Science Education*, 21(3): 293–315. DOI: <http://dx.doi.org/10.1080/095006999290714>
- Stufflebeam, D., 2001. Evaluation models. *New directions for evaluation*, 2001(89): 7–98. DOI: <http://dx.doi.org/10.1002/ev.3>
- Wilson, M. and Sloane, K., 2000. From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13(2): 181–208. DOI: http://dx.doi.org/10.1207/S15324818AME1302_4

How to cite this article: 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): 8, pp. 1–6, DOI: <http://dx.doi.org/10.5334/cstp.15>

Submitted: 30 March 2015 **Accepted:** 16 September 2015 **Published:** 20 May 2016

Copyright: © 2016 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.

OPEN ACCESS