

Promising Practices In 4-H Science



4-H Science

May 2011

Evaluation of 4-H Science Programming

Katherine E. Heck & Martin H. Smith
University of California, Division of Agriculture and Natural Resources
4-H Youth Development Program

Why evaluate?

4-H Science programs provide opportunities for young people to learn science content, improve their science process skills, and develop positive attitudes toward science. This is accomplished within a positive youth development program structure that allows youth to form positive relationships and to grow as individuals. Evaluations can provide important information about 4-H Science programming, including whether or not, or to what degree outcomes have been achieved; areas necessary for program improvement; and to demonstrate to funders the value of a particular program.

Evaluations may include formative or summative elements, or both. Formative evaluations are primarily concerned with process and how well a program worked, and the information gained is typically used to feed back into program improvements. Summative evaluations focus on the outcomes of the program; for example, they might attempt to determine what skills the young people learned.



Formative evaluations: Evaluation of process

Many evaluations of out-of-school time programming have been published. Research has demonstrated that formative evaluations can improve the quality of programming (Brown & Kiernan, 2001). Often conducted in the early stages of program implementation, formative evaluations typically use observational data, surveys, and/or interviews to determine whether the program is being implemented in accordance with its stated goals, and whether it adheres to a positive youth development framework. Participant satisfaction is one element commonly addressed. Formative evaluations could identify challenges that may help the program to improve. Such challenges might involve structural issues (such as problems with attendance), staff development needs, issues of collaboration or other concerns (Scott-Little, Hamann, & Jurs, 2002). Additionally, formative evaluation may be used in 4-H Science programs to see how closely they adhere to the recommended 4-H Science checklist, which states that programs should use inquiry strategies and an experiential learning approach, and should be delivered by trained and caring adults who include youth as partners.



Summative evaluations: Evaluation of outcomes

A summative evaluation seeks to determine whether, or to what degree, youth who participate in a particular program achieve targeted outcomes. Summative evaluations are typically conducted on established programs, rather than in the first few months of program implementation. They seek to identify the impacts programs have on their participants.

The 4-H Science Initiative promotes the acquisition of a specific set of science process skills within a framework of positive youth development and based on the National Science Education Standards. Youth who participate in 4-H Science programming are expected to improve their “Science Abilities,” a group of science process skills. These 30 abilities, which include skills such as hypothesizing, researching a problem, collecting data, interpreting information, and developing solutions, are delineated on the 4-H Science Checklist, available at <https://sites.google.com/site/4hsetonline/liaison-documents/4-HSET-Checklist2009.pdf>. Youth in a 4-H Science-Ready program are also expected to have the opportunity to develop mastery and independence, to be able to contribute and feel generosity, and to feel a sense of belonging within the group. The acquisition of science content, the development of science process skills, and youth development outcomes are all possibilities for summative evaluation.

Evaluation methods

Evaluation of science programs can take a variety of forms, such as observation of the program, surveys of participants, interviews with program staff or volunteers, or focus groups with participants. Each method has advantages and disadvantages, and selection of the appropriate method will depend on the goals of the evaluation.

- An initial interview with the program director is important in determining the goals of the program, which is the first step in identifying the most appropriate methods for evaluation. If the program has a logic model, or can develop one, that can help clarify the goals and outcomes to be evaluated. The evaluation plan needs to be specific and appropriate to the program that is being evaluated.
- Individual interviews or focus group interviews with staff, volunteers, or participants may provide insights into process issues that could be useful in formative or summative evaluations.
- Observation of a program can provide useful information for a formative evaluation in determining how well the program delivery is working. Observers need to be prepared in advance for what to look for and how to note or code the interactions or other behaviors. One potential resource for observational data collection focusing on quality of programming is the Out-of-School Time (OST) Evaluation Instrument developed by Policy Studies Associates available from <http://www.policystudies.com/studies/?id=30>.
- Surveys of program participants are one of the most common methods for evaluation. Surveys, particularly those involving quantitative data, have the advantage of typically being quicker and simpler, and often less costly, to complete than qualitative methods. Traditional survey data collection was done with paper and pencil surveys, but new forms that may be used include online surveys or group surveys done with computer assistance.
- Authentic assessment is another method of evaluation. In authentic assessment, learners are asked to demonstrate their knowledge and skills through real-world tasks (Palm, 2008). Authentic assessments can be utilized for formative or summative purposes and can be used independently or in conjunction with more traditional assessment strategies such as surveys, observations, or interviews. With respect to 4-H Science, authentic assessment strategies such as learners' responses to open-ended questions, written data presented in a science notebook, developing, conducting, and explaining the results of an experiment, or designing and building a model bridge can be used to assess youths' understanding of specific science content and the development of Science Abilities. The activities or results are then coded independently by the evaluators.



Best practices in science program evaluation

Evaluation tools need to be piloted prior to implementation to ensure that they are able to be used and understood, and to identify and correct any problems. Ideally the items on a survey or interview should be validated and the reliability of the instrument ascertained. The reliability of an instrument is the extent to which the instrument achieves consistent results, which may include consistency among the participants or among the individuals who are coding or rating the responses. The validity is a determination of whether the instrument is measuring what it is intended to measure; whether the questions accurately and fully provide a picture of the concepts or constructs intended.

There are several considerations to take into account when identifying a particular evaluation method and an appropriate tool to use. Some of these might include:

- The age or abilities of the potential respondents. For example, paper and pencil surveys or surveys with complex questions are not appropriate for very young children.
- The number of respondents. Large numbers of respondents are often more suitable for quantitative methods, research that involves data that are typically measured numerically, with objective questions, often using multiple choice or other fixed response categories. Small numbers of participants may make qualitative methods more appropriate and feasible. Qualitative research involves a text-based method of developing themes, theories, and ideas from observations, short or essay answers, focus groups, or interviews.
- The amount of time available for completing the evaluation. Some methods are quicker to complete than others.
- Whether the evaluation should be longitudinal and follow participants over time, or whether a point in time, cross-sectional evaluation is appropriate.
- How large a burden the evaluation needs to place on participants in terms of time they will need to spend responding to a survey or participating in an interview or focus group.
- What specific outcomes or processes the evaluation will be ascertaining.

When developing new instruments or questions, it is important to try to prevent bias in the survey that may result from sampling issues or from the wording of particular questions. Bias results when the survey or sampling design causes error in measurement. Participants in evaluation surveys should ideally be either the entire program or a representative sample of all program participants. Questions should be worded to be as objective as possible, and written in a way that respondents are able to answer accurately. (For example, surveys asking about past behavior may be answered inaccurately if the respondent cannot recall behavior from a long time ago.)

Some states, such as Tennessee and Louisiana, have begun routinely collecting and summarizing evaluation data from program participants. Tennessee has developed a program evaluation network that includes a limited set of evaluation items that are completed by all youth participating in 4-H programming in the state. The specific questions that youth respond to vary depending on the projects in which they are engaged. The questions in the survey bank were drawn from previously validated tools, such as, for science programming, the Science Process Skills Inventory (available at <http://www.pearweb.org/atis/tools/18>).

Resources for science program evaluators

There are many pre-existing validated, reliable surveys that can be used focusing on science content and science process skills. One useful resource for these tools is the Assessment Tools in Informal Science (ATIS) website, maintained by Harvard University's Program in Education, Afterschool and Resiliency. This website is located at <http://www.pearweb.org/atis>. Previously validated instruments are available for a range of participant ages, from preschool through college, and focusing on a range of science outcomes, from skills to interest to content knowledge. In addition to ATIS, there are a bank of instruments on science and other youth development outcomes available at the CYFERNET website, at https://cyfernetsearch.org/ilm_common_measures. These instruments are used in the evaluation of National Institute of Food and Agriculture (NIFA)-funded Children, Youth, and Families At Risk (CYFAR) projects, and were selected from pre-existing, validated instruments on a variety of topics, including science and technology but also demographics, youth program quality, leadership, citizenship, nutrition, parenting, workforce preparation, and others.

In addition to pre-existing validated tools, several journals also exist which publish articles that provide guidance on methodological issues around program evaluation. Among these are Practical Assessment, Research & Evaluation (<http://pareonline.net/>); the American Journal of Evaluation (<http://aje.sagepub.com/>); Evaluation Review (<http://erx.sagepub.com/>); and Evaluation (<http://www.uk.sagepub.com/journals/Journal200757>).

References

- Brown, J.L., & Kiernan, N.E. (2001). Assessing the subsequent effect of a formative evaluation on a program. *Evaluation and Program Planning*, 24(2), 129-143.
- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment Research & Evaluation*, 13(4). Available online: <http://pareonline.net/getvn.asp?v=13&n=4>
- Scott-Little, C., Hamann, M.S., & Jurs, S.G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation*, 23(4), 387-419.

National 4-H Science Professional Development: Janet Golden, jgolden@fourhcouncil.edu, 301-961-2982
Graphic Design and Layout: Marc Larsen-Hallock, UC Davis, State 4-H Office