Chapter 2
Designing Evaluations for K–12 Robotics Education Programs

Kristen Stubbs
Electra Studios, formerly of iRobot Corporation, USA

Jennifer Casper
The MITRE Corporation, USA

Holly A. Yanco
University of Massachusetts Lowell, USA

ABSTRACT

While a large number of robotics programs for K-12 students have been developed and deployed in the past twenty years, the effect that these programs have on students' motivations to enter science, technology, engineering, and mathematics (STEM) careers has yet to be fully determined. In order to demonstrate the value of these programs, researchers must make a concerted effort to measure their impact. Based on prior work in the evaluation of educational robotics programs, the authors of this chapter present frequently-utilized evaluation and measurement methods as well as guidelines for selecting these methods based on factors such as a program's duration, size, and maturity. This chapter is intended for use as a reference guide for designing evaluations of K-12 educational robotics programs.

1 INTRODUCTION

For over twenty years, a wide variety of programs have introduced K-12 students to robotics. Large programs such as FIRST, started in 1989, and Botball, started in 1997, have brought robotics to tens of thousands of students. There are many reports, mostly anecdotal, that students are motivated by such experiences and by other exposures to robotics in their classrooms. However, enrollment in STEM fields (science, technology, engineering and mathematics) at the college level continues to be a concern, even after more than twenty years of organized programs using robotics to attract students. Are these robotics education programs having any impact in increasing student interest in STEM fields?

In order to determine the answer to this question, robotics researchers, designing and offering

DOI: 10.4018/978-1-4666-0182-6.ch002

programs for K-12 students, must evaluate their programs and their impact upon the students who participate in them. The importance of evaluation is known and recognized (Brophy, Klein, Portsmore, & Rogers, 2008). This chapter presents several methods for evaluating robotics programs. For example, a combination of pre- and post-tests can evaluate the impact of the program in the short term, measuring students' assessments of their attitudes about issues that the program has been designed to impact. The use of comparison groups, considered the gold standard in educational evaluation, is not always possible; for example, students who self-select to participate in after school programs cannot be compared to the students who did not choose to participate. Despite these and other challenges, it is vital to perform evaluations of K-12 robotics programs in order to improve their impact.

Short-term evaluation is almost universally undertaken. However, it is much more difficult to perform a longitudinal evaluation in the years following a program. Contact information for students is often outdated within a few years after the end of a program. Yet it is the long-term evaluation that will allow us to know if interventions that use robotics truly have an impact. We do not know if students participating in one program are then participants in another program, but this type of questioning would allow us to further determine the success of our programs.

The intent of this chapter is to serve as a reference for designing evaluations for robotics outreach programs. There are many factors that influence an evaluation: program maturity, length, size of target audience, target age group, and logistical support (such as funding and personnel). This chapter addresses these factors, discussing what has been done in prior work. It concludes by suggesting ways to evaluate programs based upon the factors listed above.

2 THE EVALUATION DESIGN PROCESS

Evaluation can be incorporated throughout the implementation and execution of robotics education programs. Formative evaluations are conducted as a program is being developed (Friedman, 2008, p. 17). The purpose of the evaluation is to provide feedback about how the robot platforms, curricula, training materials, and other aspects of the program should be modified before the program is conducted again. Summative evaluations are conducted to determine the impact of a program on its target audience (Friedman, 2008, p. 9). For robotics education programs, summative evaluation often involves measuring what students have learned and how their attitudes have changed after participating in a program.

When designing either formative or summative evaluations, it is necessary to determine both the evaluation method(s) and measurement method(s) which will be used. An evaluation method represents an overall strategy for data collection: when will data be collected and from what groups of participants. Measurement methods are the means by which data will be collected from these participants, such as conducting interviews or questionnaires.

The following sections outline a process for evaluations intended to improve a program's design, and for evaluations intended to assess a program's effectiveness, respectively.

2.1 Evaluation for Program Improvement

The development of robotics education programs tends to follow a model similar to that shown in Figure 1. Program development is an iterative process. Initial materials are developed, along with measurement methods. These materials and methods are then piloted with a small number of students or instructors. Analysis of the evaluation data indicates what changes should be made. Both