

Evaluation Can Be Fun and Meaningful at the Same Time

By Karen Peterman, Ph.D.

Imagine a group of students racing across a local park with GPS units in hand, working together to complete a scavenger hunt that consists of successive locations. As they run from one location to the next, they look for a second team, racing on its own scavenger hunt course to the same finish line. Passersby stop to ask, "What's going on? Is this a Summer Camp or something?" "Yes," a staff member replies, "and it's also our evaluation."

The above is an example of an embedded performance-based assessment, a method that GRG evaluators have begun integrating into our multi-method evaluation designs in recent years. Performance-based assessments can be defined as evaluation activities that require participants to demonstrate their skill or knowledge by performing program-related tasks. The use of this technique is considered an embedded assessment when the performance-based tasks are seamlessly integrated into a program, such that the evaluation activities and program activities are virtually indistinguishable.

There are many benefits to using this method. First, performance-based assessments do not burden participants with tasks, such as surveys, that can feel unrelated to a program's activities. Perhaps more importantly, these assessments have high construct validity (i.e., the degree to which an assessment measures an underlying concept) since they require participants to demonstrate their abilities. As such, performance-based assessments offer an accurate and authentic measure of the program activities. They also are enjoyable for participants to complete.

GRG's recent focus on embedded performance-based assessment began with our evaluation of *Technology at the Crossroads*, a program created by the Girls Get Connected Collaborative and Simmons College in Boston. Working closely with the project team, GRG designed and implemented a field day event consisting of games that required students to use the science knowledge and technology skills they acquired from the program in a competition against other teams. The GPS scavenger hunt presented at the beginning of this article is an example of one of these games.

Embedded authentic assessments can also include the analysis of existing program activities, which is the case with GRG's evaluation of an online interactive game being created by the University of California San Diego as part of their NSF-sponsored Information Technology Engineering and Environmental Education Tools (IT-E3) project. The game includes a series of missions that require students to use open-

ended inquiry skills to solve real-world problems in virtual Antarctica. It also captures every move that students make in the virtual world, and GRG is collaborating with the project team to analyze these data to demonstrate students' use of inquiry and process skills within and across missions.

GRG's evaluation of the Simmons College Libraries MassBLAST (Building Library Awareness and Staff for Tomorrow in Massachusetts) project has also taken advantage of planned program activities. This project team created a number of performance-based assessments to measure the success of their sites at providing library interns with key concepts (e.g., arranging books by call number). These assessments allow library sites to monitor teaching success, and GRG to measure the interns' library knowledge.

Existing program plans can also be expanded to serve both the project and the evaluation. For example, the Dan River Information Technology Academy (DRITA) program at the Institute for Advanced Learning and Research provides high school students with training in specific IT and workforce skills. GRG is gathering baseline data on students' interview skills (prior to receiving training in this area) by holding mock job interviews and then coding those interactions for skills that will be featured in the training. These baseline data will be compared to students' performances in a second mock interview that is part of the program's implementation, and they will also serve as information the project team can use to target training for individual students.

Regardless of how embedded performance-based assessments are created and integrated into a program, these innovative techniques require a strong client-evaluator collaboration and a proactive approach to planning both program and evaluation activities. GRG encourages program developers to think about ways to incorporate these methods into project and evaluation plans. We have found the professional experiences and relationships that result from their use to be some of the most rewarding for ourselves, our clients, and most importantly, program participants.



Karen Peterman, Ph.D., Senior Research Associate, joined GRG in October 2002, and has completed over 25 evaluations at GRG. She began conducting performance-based embedded assessments in 2005, and these methods continue to play a prominent role in her work.