Articles

Assessing Precollege Science Education Outreach Initiatives: A Funder's Perspective

Debra A. Felix,* Mark D. Hertle, Jill G. Conley, Lori B. Washington, and Peter J. Bruns

Howard Hughes Medical Institute, Precollege Science Education Program, Chevy Chase, MD 20815

Submitted October 2, 2003; Revised April 19, 2004; Accepted April 20, 2004 Monitoring Editor: Marshall Sundberg

> We describe an assessment of the collective impact of 35 grants that the Howard Hughes Medical Institute (HHMI) made to biomedical research institutions in 1999 to support precollege science education outreach programs. Data collected from funded institutions were compared with data from a control group of institutions that had advanced to the last stage of review but had not been funded. The survey instrument and the results reveal outcomes and impacts that HHMI considers relevant for these programs. The following attributes are considered: ability to secure additional, non-HHMI funding; institution buy-in as measured by gains in dedicated space and staff; enhancement of the program director's career; number and adoption of educational products developed; number of related publications and awards; percentage of programs for which teachers received course credit; increase in science content knowledge; and increase in student motivation to study science.

Keywords: primary, secondary, assessment, evaluation, impact, precollege, K-12, grants.

INTRODUCTION

Assessment is critical to program planning and implementation. It is a necessary tool for funders to use to evaluate the effectiveness of their programs and to look ahead. In many ways, the summative evaluation of individual grants is a formative exercise in planning new grant initiatives; the grantee of one initiative informs and aids the grantor in planning the next. Yet many grantees view assessment as merely completing a report card in which the amount of effort expended is the sole measure reported.

Many program evaluations simply record outputs, such as the number of participants served. Although a valid count, it is at best a minimum measure. In addition, it is an unfortunate reality that evaluation quality varies widely depending on the sophistication of a program's evaluation team, which is often insufficiently broad to advance beyond the most easily measured evaluation criteria. As an example, a program might report the *output* that it trained 125 biology teachers to teach inquiry-based science. It would be more meaningful, however, if the evaluation measured an *outcome*, such as changes in the quality of their teaching or how many

DOI: 10.1187/cbe.03-10-0017

*Corresponding author. E-mail address: felixd@hhmi.org.

of these teachers continued teaching science compared with their peers who had not participated in the program. Even more powerful, and ultimately more important, would be to characterize the *impact* of the program by seeing whether the students of a participating teacher learned more science than the students in a neighboring classroom whose teacher was not trained by the program. Unfortunately, measuring impact can take years and is often difficult to assert without a considerable number of caveats.

Because individual grantees can at best measure only the results of their own programs, it is important for HHMI to assess the cumulative effect of its initiatives both across grantees and over time. Our practices have evolved toward increasing cooperative engagement with grantees to evaluate their projects by funding grantee-led evaluation efforts and cluster evaluations such as the one described here.

In late 2002, we undertook a study of the outcomes of the 4-year science education grants HHMI made to biomedical research institutions in 1999. In that year, \$12.6 million was awarded to a cohort of 35 medical schools, biomedical research institutions, teaching hospitals, and academic health centers; individual grants ranged from \$225,000 to \$500,000. The purpose of the grants was to encourage science-rich institutions to share their knowledge and resources with teachers and students and to promote the understanding and appreciation of science to people of all ages. See

http://www.hhmi.org/grants/office/precolprog/biomed.html for a full description of this initiative and Appendix A for a list of the 35 grantees.

The programs carried out by these 35 grantees were specifically targeted to students, their parents or caregivers, and teachers from preschool through 12th grade. The projects involved one or more of the following activities: teacher professional development, outreach activities for students, science education directed at families and the general community, or the creation of science curricula and associated educational resources. Some grantees conducted multiweek summer professional development programs for teachers. Others ran student-in-the-lab programs or summer science camps. Some sent scientists to classrooms to collaborate with a teacher. Yet others produced educational products, such as a statewide elementary school science curriculum and a series of online laboratories for the entire K-12 population. Our hope was that these programs would fill the biomedical pipeline with young people who would come to love science and who would later take up careers in health or medicine. At the same time, we hoped to increase science literacy and give scientists more opportunities to interact with and influence the general public in their communities.

Through annual progress reports, financial reports, and site visits, we tracked the progress of the 35 grantees and were pleased with the quality of science outreach work being done at the individual project level. A more comprehensive look at what these institutions have been able to accomplish as a group seemed in order, however. To this end, in February 2003, during the last year of their 4-year grants, we sent a questionnaire to the directors of all 35 programs. The questions asked were largely based on what grantees had identified in their yearly progress reports as the most compelling measures of their project outcomes. Follow-up was done individually, by phone, and by e-mail to ensure questionnaire completion and occasionally to clarify answers. All but one institution completed the form. This study was not an element of the review process for a grant competition under way at the time, because we were interested in the aggregate impact of these grants, not the individual performance of applicants to a new program.

CONTROL GROUP

Although it was difficult to find an identical control group for these 35 institutions, we felt it was important to try to provide a comparator for our results. To that end, we chose as a control group the 50 institutions that were closest to, but just below, the funding cutoff based on the final stage of the 1999 grant competition. These institutions had proposed projects of sufficient quality that they progressed to the end of a rigorous review process, making them the closest available parallel to the funded institutions. Of the 50 nongrantees contacted, 19 responded with information about the outreach programs-if any-they were able to implement in the same time period (between 1999 and 2003) even without HHMI funding. One institution had a program funded by an ongoing grant from another HHMI initiative. Its data were eliminated, leaving us with 18 valid responses.

Ĥow comparable are these institutions to our grantees? All 18 institutions included as controls are also medical schools,

biomedical research institutions, teaching hospitals, or academic health centers. Five of them had been awarded grants from HHMI in the previous competition in 1994. Fourteen applied to the most recent (2003) competition within this initiative, and two were awarded grants. Furthermore, despite being denied funding from HHMI, 13 of the 18 were able to conduct some form of the program they had proposed in 1999 using other resources and grants.

Finally, as an indication of the amount of research at each institution, the 35 grantees had an average of 74 R01 and R37 grants from the National Institutes of Health (NIH) at the time of the 1999 competition, although this average falls to 58 if we eliminate three outlier institutions that significantly skew the data. The control group at the time of the competition had an average of 45 R01 and R37 NIH grants. Overall, we believe that this group of near-grantees, although perhaps slightly less developed than the group of institutions we chose to fund, is the most comparable control group possible.

The questionnaire sent to grantees was designed to elicit meaningful data reflecting the degree to which 1999 grantees realized the goals of the initiative using the grant funds they received from the HHMI Precollege Biomedical Research Institutions initiative. We chose measures, indicated as headings in the Results section, that would show impacts attributable to HHMI funding.

The questionnaire sent to the control group was identical except for this additional item:

Please check one of the following:

- __ We were not able to conduct the program we had proposed due to a lack of funding.
- ___ We were able to conduct some form of the program we had proposed, without HHMI funding.

Follow-up questioning was again done by phone and email to ensure complete forms and clarification of answers. For reasons of confidentiality and to encourage participation in this type of study, we have not listed the control group institutions in this report.

RESULTS

Ability to Secure Funding

Grantees reported the dollar amount of each additional grant received by the institution for which the HHMI grant served as leverage. Similarly, nongrantees reported the number and value of any grants they received during the same 4-year period.

	No. grants (avg.)	Grant amount (avg./grant)
Grantees (34)	5.4	\$268,746
Nongrantees (18)	2.6	\$177,525

Institutional Buy-In

A key goal of this initiative was "to encourage research institutions to engage in community-based outreach to students and teachers from preschool through high school." Because we do not require in-kind contributions from grantreceiving institutions, we consider it an extremely positive sign when an institution decides to allocate valuable resources to precollege-level outreach activities, and a success when the receipt of an HHMI grant contributes to that decision. Such buy-in also helps ensure the long-term sustainability of the outreach program. Specifically, we asked program directors whether their institutions had set aside space on site for their projects' outreach activities or had devoted salary money or additional headcount to program staff:

	Gained space	Staff support
Grantees (34)	47%	44%
Nongrantees (18)	11%	22%

As an example of this, the Fox Chase Cancer Center in Philadelphia wrote:

Prior to obtaining the 1999 HHMI grant, Fox Chase Cancer Center had no formal precollege science education program. A few scientists took high school students into their labs [approximately 3 per year] and a few scientists interacted sporadically with local schools [approximately 1 per year]. Upon receipt of the grant, Fox Chase created a position for the Program Director, provided and completely renovated office space, and covered half of the secretarial costs for the program. The number of scientists volunteering their time also increased dramatically (to 27 by the third year of the grant). Scientist mentors host student scientists in their labs for a full year and often continue longer, including full-time in the summer for eight weeks. Mentors cover all costs for their students' materials, provide all training and assistance needed to help students complete a research project, and occasionally supplement the students' summer stipends.

Impact on Program Director's Career

The value placed on science education outreach varies among institutions. At some institutions, a program is well supported and boosts its director's career. At other institutions, outreach is not a priority, and the grant can even be seen as taking an employee away from his or her usual (more valued) activities. Program directors were asked whether their careers had been enhanced as a result of their participation in precollege science outreach efforts:

	Career enhanced
Grantees (34)	68%
Nongrantees (18)	50%

As an example, the program director at Harvard Medical School was promoted from assistant dean to associate dean to dean on the administrative side and from instructor to assistant professor on the academic side while managing the institution's precollege HHMI grant.

Project Products

A significant proportion of our precollege funding is used to develop science curricula and associated educational products, such as kits that include lesson plans, reusable materials, and consumable supplies; online labs; and Web sites through which lesson plans can be accessed. Less frequently, other educational tools, such as CD-ROMs and videos, are produced. We asked program directors how many educational products had been produced by their programs and how many learners had used these products:

	Kits		Curricula		Online labs		Web sites	
	Avg. no.	Users/ kit	Avg. no.	Users/ curr.	Avg. no.	Users/ lab	Avg.	Hits/year/ site
Grantees (34)	1.7	1,534	21	750	3	1,405	0.4	3,984,335
Nongrantees (18)	1.7	401	2	148	0	0	0.4	13,308
					Othe	r Produ	icts	
			A	vg. no.		U	Jsers/p	product
Grantees (34)) (18)			25			35	56

As an example, the UCLA School of Medicine developed more than 100 online labs requiring scientific problem solving. The online labs range from elementary school science through medical school basic science and cover disciplines such as biology, chemistry, molecular biology, and earth science. In each simulation, students are given a problem to solve; a kindergarten example is, "What animal am I?" Students have the opportunity to gather evidence to confirm or reject any hypotheses they have. For example, a child can click on "Food" to find out what the animal eats, "Habitat" to learn where it lives, or "Color" (which they discover does not help them very much). At any point, a student can propose a solution to the problem. Each problem set has between 5 and 50 cases that give students ample opportunities to practice and develop expert problemsolving skills. In the first 3 years of the grant, 48,000 students had completed more than 140,000 cases. (See http:// www.immex.ucla.edu/ for more information.)

Product Adoption

We believe that when a school district, city, or state deems a product to be of sufficient quality that it is formally adopted as part of the curriculum, that product has received an objective validation indicating significant program impact. Thus, we asked those programs that had developed educational tools whether their products had been officially adopted city-, county-, or statewide:

Product adopted			
22%			
11%			

As an example, teams of scientists at the Pennsylvania State University College of Medicine's Division of Developmental Pediatrics and Learning Center for Science and Health Education teamed up with practicing teachers in Pennsylvania to write 30 science activity modules for elementary school classrooms. The modules form a comprehensive, standards-based, hands-on, inquiry-driven science curriculum that has been adopted by more than 1,000 K–6 teachers across Pennsylvania, reaching approximately 25,000 students. New school districts are requesting the program and are successfully obtaining funding to establish dedicated science laboratories in their elementary schools (similar to traditional art or music classrooms) in which to implement the modules. Outside Pennsylvania, the program has been licensed and is marketed by Cognitive Learning Systems as the LabLearnerTM Program.

Publications and Awards

To measure program success in another way, we collected data on the number and types of publications and awards received by the program director, teacher participants, or student participants since 1999. Grantees were asked to list only publications and awards that resulted from their HHMI-funded program. They also included the name of the person who wrote the published piece or who won the award and the name of the publication or award. We then determined the quality of the honor using the following scale:

- Class I: A publication or award at the highest professional level, such as a publication in a peer-reviewed, national journal or participation in the INTEL Science Talent Search
- Class II: A state or local publication, abstract, or science fair award
- Class III: An in-house or self-published work

	Publications and awards		
	Class I	Class II	Class III
Grantees (34) Nongrantees (18)	25 7	356 40	6 1

Earning Credits

As with product adoption, when a university or government entity deems a program to be of sufficient quality that teachers receive official course credit for their participation, we believe we are seeing an unbiased validation of the worth of the program activities:

	Teachers earned credits	
Grantees (32)	50%	
Nongrantees (18)	22%	

For example, teachers who complete the Summer Science or Mathematics Institute given by the Carnegie Institution receive three graduate credits from George Mason University. The total number of teachers who have received graduate credits since 1994 is more than 500.

Content Knowledge

It is important to establish whether participants in HHMIfunded programs gain science content knowledge and associated skills, such as math skills and laboratory techniques. We asked grantees for evidence that teacher or student participants had acquired science content knowledge from the program. Most programs collected data, but unfortunately many of the control group institutions provided no comparisons of participants with nonparticipants; we only counted data that included a control group of similar participants or data collected on the same participants before and after the intervention. We rejected selfreported and anecdotal evidence of impact.

For example, the University of Cincinnati College of Medicine reported the percentage of students who had passed a science proficiency exam that all children must pass to graduate from high school. (The exam is given in the ninth grade, and students can retake it as often as they like until they pass it.) The grantee program looked at the pass rate of its participants (seventh and eighth graders) after their ninthgrade year, a year or two after the children had attended a yearlong Saturday Science Academy. In contrast to the statewide pass rate of 23%, the pass rate of program participants was 100%.

In another example, the Yale University School of Medicine reported an increase of 62 points in the math Scholastic Assessment Test (SAT) scores of participants between the beginning of the program and its close. Yale also reported that the average grade point average of its participants in science courses increased from 2.5 to 2.8 by the end of the program.

Finally, another grantee reported that 87% of the students in its program said they had learned "a moderate amount," "quite a bit," or "a lot" by working on a science project during the program.

As is evident from these examples, not all grantees measured the same parameters in the same way, so it is impossible to simply consolidate the responses and conduct statistical analyses on them. Instead, we were forced to evaluate each claim individually. Did we think the institution's program had succeeded in establishing whether it had imparted science knowledge and skills to its participants? With respect to the examples given here, we answered yes to the first three claims and no to the fourth because it comprised only self-reported claims without any comparison, pre-post, or control group data. Only the first three were included in the data tables. We indicate this approach in the following tables by noting in parentheses both the number of reporting institutions with relevant programs (denominator) and the number of responses that we considered relevant (numerator).

	Teachers gained knowledge		
Grantees (11/32) Control group (1/18)	34% 6%		
	Students gained knowledge		
Grantees (17/23) Control group (1/16)	$74\% \\ 6\%$		

Student Motivation to Study Science

We asked respondents if they could show that their programs had motivated participants to continue their study of science. Although most projects collected such data, we again counted the data only if they included a control group of similar participants or if results on the same students were collected before and after the intervention to demonstrate the change. As in the previous section, we rejected self-reported or anecdotal evidence of impact. One acceptable measure, which several grantees reported, was the percentage of program participants who went on to major in the sciences in college compared with national averages:

Grantee	% Postprogram science majors
University of Cincinnati College of Medicine	83
University of Nevada School of Medicine	63
Robert C. Byrd Health Sciences Center of	59
West Virginia University	
University of Mississippi Medical Center	59
Cleveland Clinic	53
National average	32
National average for underrepresented mino-	5
rities (most program participants)	

Although 10 of 23 grantees demonstrated that their programs produced a large number of students who went on to study the sciences, none of the control institutions did.

DISCUSSION

The results suggest that, as a group, HHMI grantees achieved positive outcomes and measurable impacts on all parameters measured: ability to secure non-HHMI funding; institutional buy-in as measured by dedicated space and staff, enhancement of the program director's career, number and adoption of educational products developed, number of related publications and awards, percentage of programs for which teachers received course credit, increase in science content knowledge, and increase in student motivation to study science. In addition, the results show that institutions that did not receive HHMI funding were either unable to implement an outreach program or were, on average, able to implement only a less effective program than that of the average HHMI grantee.

Although we realize that reliable, quantitative measures of impact are difficult to obtain, we judged the data critically to determine whether they adequately and convincingly answered our questions. The results of the study show that, overall, the \$12.6 million in HHMI precollege funding made to biomedical research institutions in 1999 achieved its mission of increasing young people's and teachers' exposure to, interest in, and understanding of science. Furthermore, the grants were used to attract additional support and resources that grantees could use to enhance their success.

Using the results of this study, we have designed an instrument to help grantees collect their outcomes data. The instrument is closely aligned to the questionnaire we used in this study (see Appendix B). Our hope is that by giving grantees a relevant and clear framework for evaluation before their grants begin, they will be able to design and implement meaningful evaluation processes that will simultaneously allow us to make summative (although not comparative) analyses of our initiatives.

To further help grantees learn how to evaluate their programs, we have implemented a comprehensive peercentered process that includes evaluation training, reciprocal site visits to see how peer institutions evaluate similar programs, access to the services of a professional evaluator for guidance and technical assistance, and the dedicated time to focus on and reassess one's own program-evaluation activities.

Clearly, program assessment should not come only at the close of the program. It should be planned at the outset, implemented throughout, and inform change long before the program has run its course. Moreover, program evaluation can be a positive collaboration between grantor and grantee. It is our hope that the evaluation assistance we are now providing to our grantees will help them fine-tune their evaluation practices and, using the results, improve their programs for the benefit of all participants.

Finally, this study should provide applicants of future HHMI grant competitions with insight into what we hope the projects might achieve and what methods we endorse to assess the success of our initiatives.

ACKNOWLEDGMENTS

Funding for this study was provided by the Howard Hughes Medical Institute.

REFERENCES

Chubin, D.E. (1995). Education program evaluation at NSF: What difference does it make? In: Innovating and Evaluating Science Education: NSF Evaluation Forums, 1992–94. http://www.ehr.nsf. gov/ehrPublications.cfm (accessed 13 July 2004).

Daniels, S., McIntosh, W., and Leonard, J. (1996). Process or outcome? Foundation News & Commentary (March-April), 46–48.

Easterling D. (2000). Using outcome evaluation to guide grant making: Theory, reality, and possibilities. Nonprofit and Voluntary Sector Quarterly 29, 482–486.

Fine, A.H., Thayer, C.E., and Coghlan, A. (1998). Program Evaluation Practice in the Nonprofit Sector. A study funded by the Aspen Institute Nonprofit Sector Research Fund and the Robert Wood Johnson Foundation, Washington, DC: Innovation Network.

Fitzpatrick, J.L., Sanders, J.R., and Worthen, B.R. (2004). Program Evaluation: Alternative Approaches and Practical Guidelines, 3rd ed., Boston: Allyn and Bacon.

Patrizi, P., and McMullan, B. (1998). Evaluation in Foundations: The Unrealized Potential. Prepared for the W.K. Kellogg Foundation Evaluation Unit. http://www.wkkf.org/Programming/ ResourceOverview.aspx?CID=281&ID=773 (accessed 13 July 2004).

Sundberg, M.D. (2002). Assessing student learning. Cell Biol. Educ. 2002 1(Spring/Summer), 11–15.

United Way of America. (1996). Measuring Program Outcomes: A Practical Approach, Alexandria, VA: United Way of America.

Appendix A

GRANTEES STUDIED

Baylor College of Medicine Boston University School of Medicine Carnegie Institution of Washington Cleveland Clinic Foundation Cold Spring Harbor Laboratory Columbia University College of Physicians and Surgeons Creighton University School of Medicine Fox Chase Cancer Center Fred Hutchinson Cancer Research Center Harvard Medical School Massachusetts General Hospital Oklahoma Medical Research Foundation Pennsylvania State University Rockefeller University Rush-Presbyterian St. Luke's Medical Center University of Alabama at Birmingham

University of California-San Diego, School of Medicine

University of California–Los Angeles

University of California-San Francisco

University of Chicago Division of the Biological Sciences and Pritzker School of Medicine

University of Cincinnati College of Medicine

University of Kentucky

University of Massachusetts Medical Center

University of Medicine and Dentistry of New Jersey-New Jersey Medical School

University of Minnesota-Twin Cities

University of Mississippi School of Medicine University of Nevada School of Medicine University of South Dakota School of Medicine University of Utah School of Medicine University of Washington University of Wisconsin Medical School Wake Forest University School of Medicine Washington University School of Medicine West Virginia University, Robert C. Byrd Health Sciences Center Yale University School of Medicine

Appendix B

OUTCOMES QUESTIONNAIRE

The Questionnaire

How Do You Measure Your Program's Success?

Please answer as many of the following questions as you can, including a brief explanation of your evaluation technique and data to support your findings.

You may not be able to answer every question positively, and that is fine. Just answer as many as you can, because we are interested in gathering as much information about your program's successes as possible. As far as is practical, please restrict your comments to reflect only the work that HHMI has supported.

Also, wherever possible, please include a control group (ideally), context, benchmark, or other comparator that will help us to understand the scope of each accomplishment.

If you have not yet collected any data, please describe your evaluation plan and estimate when you will have results. Thank you again for taking the time to provide us with this critical information.

Please complete this survey by typing your answers below after each question and e-mailing the form back to us (or print it out and mail it to us) by February 25, 2003.

We do not expect you to answer every question. Please type "NA" if a question does not apply to your project. Thank you again for your time.

1. Resources Gained

Were you able to leverage your 1999 HHMI grant to attract additional resources for your program? By resources we mean additional funding, donations of space, volunteer hours, the creation of a new outreach position by your institution, etc. Please tell us what resources you have gained as specifically and quantifiably as you can. Don't forget to include a context, benchmark, or other comparator that will help us to understand the scope of each accomplishment—for example, that you were able to get NSF funding because the HHMI grant constituted your non-Federal cost-sharing commitment, or that you were given space or administrative help for your program contingent on a particular level of funding.

2. Career Enhanced

Has your own position been enhanced in any way (salary, title, place in your institution's hierarchy, tenure) as a result of your HHMI-funded work on this program?

3. Educational Products Developed and Disseminated

Has your program produced any educational products?

Please provide the following information:

The number of each type of product produced

A general description of the product(s)

The estimated number of users (for online tools please cite the number of Web "hits")

The measured impact of the product(s), as seen in pre-post usage data, or in the product's adoption by a local or state school system for integration into its official curriculum

Product	#	General description	No. users	Impact
Kits				
Curricula				
Online labs				
Web sites				
Other				

4. Teachers Gained Knowledge or Skills

If you have a teacher professional development program, have you been able to show:

A significant increase in the science content knowledge or teaching quality of the teachers you've trained? If so, please state your findings.

-OR-

That teachers' confidence in or attitudes toward science improved as a result of their participation in your program? If so, please state your findings, including a control group (ideally), context, benchmark, or other comparator that will help us to understand the scope of the change.

5. Students Gained Knowledge or Skills

If you have a teacher professional development program, have you been able to show that the science knowledge of students increased as a result of their teachers' participation in your program? Please present your results, including control data.

-OR-

If your program serves students directly, have you been able to show that the science knowledge of students increased? Please present your results, including control data.

6. Students Were Motivated to Study Science

By such measures as the number of science courses elected or choice of college major, can you show that program participation increased students' motivation to study science at an advanced level? Please enter your findings, including control data and data collection methods.

7. Students or Teachers Won Awards or Coauthored Papers

Please list any participants in your program who

Won awards at science fairs (please list details) or

Coauthored papers for peer-reviewed journals as a result of participation in your program.

Please provide comparators that will put this information into context.

8. Teachers Earned Credits

Did teachers in your program earn more graduate level and continuing education credits than a comparable group of nonparticipating teachers? Please present your findings here.

9. Students Graduated

Did students who participated in your program graduate from high school at a higher rate than their peers? Please provide this data for participating students and for a *comparable*, or matched, group of nonparticipating students.

10. Families or Community Members Served

Did your program have a measurable impact on local families or community members served? Please elaborate, including comparators that will put this achievement into context.

[Note: There were not enough responses to this question to include them in the analysis.]

11. Other

Are there other quantitative measures of your program's success that have not been captured here of which you believe our Board of Trustees should be apprised?

[Note: There were not enough responses to this question to include them in the analysis.]

Thank you again for your time.