

National Science Foundation

The Cultural Context of
Educational Evaluation:

The Role of Minority
Evaluation Professionals

**WORKSHOP
PROCEEDINGS**

JUNE 1-2, 2000

*Arlington Hilton and Towers Hotel
Arlington, Virginia*



Directorate for Education and Human Resources
Division of Research, Evaluation and Communication

ACKNOWLEDGEMENTS

These proceedings are an edited version of a two-day meeting with minority evaluation professionals. Session summaries were prepared by invited participants, several of whom were also asked to prepare papers/presentations to frame the discussion. Dr. Elmima Johnson organized the conference and served as editor of this document. We would like to acknowledge the contributions of all attendees, including members of the Directorate for Education and Human Resources who gave invited comments or served as workshop discussants. We also would like to acknowledge the assistance of Joy Frechtling and Martine Brizium, Westat, Inc., in conference planning and logistics under Contract No. REC-9412965.

The Cultural Context of Educational Evaluation:

The Role of Minority Evaluation Professionals

National Science Foundation
Directorate for Education and Human Resources
Division of Research, Evaluation and Communication

Any views, findings, conclusions, or recommendations expressed in this report are those of the participants, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

CONTENTS

	Page
INTRODUCTION	1
OPENING SESSION	2
SESSION ONE: Evaluation of Educational Achievement of Underrepresented Minorities	5
Guiding Question	5
Discussion Highlights	6
Papers/Presentations	10
SESSION TWO: Participation of Minority Professionals in Educational Evaluation	24
Guiding Questions	24
Discussion Highlights	25
Papers/Presentations	28
SESSION THREE: Training Issues: Barriers, Recruitment and Outreach Strategies, and Training Mechanisms	37
Guiding Questions	37
Discussion Highlights	38
Papers/Presentations	41
WORKSHOP RECOMMENDATIONS	56
CLOSING REMARKS	58
APPENDICES	
Appendix A: Workshop Agenda	60
Appendix B: List of Conference Participants	62
Appendix C: Biographies of Invited Participants	65

INTRODUCTION

The National Science Foundation (NSF) Directorate for Education and Human Resources (EHR) sponsored a two-day workshop on the cultural context of educational evaluation June 1-2, 2000, at the Arlington Hilton and Towers in Arlington, Virginia. Invited participants included 15 nationally recognized minority evaluation professionals as well as EHR staff. The meeting served as a platform for presentation of seven invited papers and talks. It also provided a forum for examining a host of issues associated with evaluation.

The NSF Assistant Director for EHR and the Director of the Division of Research, Evaluation and Communication delivered opening remarks and framed the meeting's purpose within NSF and the Directorate. Three workshop sessions followed, organized around two major themes:

- Academic achievement by underrepresented minorities and
- Training and participation of minority professionals in the evaluation of mathematics and science programs.

Invited papers and presentations stimulated dialogue. Invited discussants, experienced educators familiar with program accountability issues across the educational continuum, offered their views.

The workshop represents the opening of a dialogue and will serve as a reference point for the Directorate as it determines its role in building capacity within the field of educational evaluation. This report presents the formal workshop papers and presentations as well as highlights of the discussions. Appendices list the workshop agenda and participants.

OPENING SESSION

Presiding

Elmima C. Johnson

Staff Associate

Division of Research, Evaluation and Communication (REC)

Directorate for Education and Human Resources (EHR)

National Science Foundation (NSF)

Welcome

Judith S. Sunley

Interim Assistant Director

EHR/NSF

Dr. Sunley welcomed participants, noting that they were a distinguished group with expertise and experience in the evaluation of education programs. As such, she said, they are aware of the issues and needs that pertain to evaluation and capable of proposing effective strategies. The participants also were charged with considering the challenges of training NSF staff and grantees as well as other professionals.

She offered a contextual framework for the meeting and its importance to the Directorate, noting that the requirements of the Government Performance and Results Act (GPRA) have increased the need for substantive evaluation of NSF's activities. Dr. Sunley stressed the importance of having the "numbers" reflect what is happening in the field in a consistent manner in order to provide guidance to program managers and help them describe project performance relative to program objectives. She also noted that the meeting was a follow-up to the Evaluation Program's January 2000 workshop on training evaluators to work in mathematics and science. That workshop suggested the need for more SMET (science, mathematics, engineering and technology) evaluators, including those from traditionally underrepresented groups.

She suggested two major meeting tasks. First, participants were asked to consider, when discussing equity and diversity, whether it matters who formulates and asks questions and conducts the evaluation. The second task pertained to a broader issue: how to connect interim and ultimate program outcomes. Because long lead times are common, NSF has difficulty tracing causal pathways between expenditures and outcomes. Focus on interim outcomes can help address this challenge.

Dr. Sunley closed by committing support to evaluator training and related activities, including attention to the participation of minority evaluators. She agreed to consider carefully the recommendations of the workshop, particularly because of the projected need for a substantial number of new evaluators who are familiar with mathematics and science as well as evaluation techniques.

Greeting

Eric R. Hamilton

Interim Division Director
REC/EHR/NSF

The issues raised at this meeting — building the capacity of the national education evaluation enterprise and, specifically, increasing the presence and participation of culturally attuned and minority-represented evaluators — collectively reflect a strategic national concern. As former director of NSF's Comprehensive Regional Center for Minorities in Chicago in the early 1990s, I am deeply aware of the difficulties involved. We labored long and hard to devise and implement a pipeline of programs for K-12 youngsters in the city that would bring them into the urban higher education centers prepared to excel in science and engineering. We also sought to build capacity in the school system to make sure more minority youngsters had the opportunity to successfully navigate the gatekeeper courses of middle school and high school mathematics. But we were always functioning *ad hoc* as we generated the formative evaluations that would allow for continuous, robust improvement of the program pipeline. Our problems were a microcosm of the broader national need to elevate education evaluation, especially as it pertains to minority-focused programs. On behalf of NSF and the REC Division, please accept our welcome to this effort to help build deeper leadership and participation by minorities in evaluation. Their involvement is critically needed.

Remarks

Conrad Katzenmeyer

Senior Program Director
REC/EHR/NSF

Unlike days past, evaluation has become an important issue for NSF's programs and projects. Once, evaluation was considered an add-on. Program officers dropped it from budgets as they struggled to fund a few more grants. Now, all programs require the inclusion of evaluation plans in proposals. Also, many programs require a percentage of each award (often 10%) to be directed to evaluation. Under these circumstances, it is critical to have trained evaluators available. Often, we hear complaints from Principal Investigators who say they cannot locate evaluators to work with them. Clearly, there is a shortage of evaluators — particularly minority evaluators — with mathematics and science backgrounds and experience.

Although our budget is limited, the Evaluation Program has tried for several years to have an impact on this problem. In 1993, we supported the publication of a *User-Friendly Handbook for Project Evaluation*, developed by Westat, Inc. It provides our Principal Investigators with the basics of evaluations. In 1997, we published a companion document, *The User-Friendly Handbook for Mixed Methods Evaluation*. To apply these materials, we also sponsored a series of short workshops, primarily for Principal Investigators and their staff.

In addition, for doctoral-level evaluation training, we support the work of the American Educational Research Association. It coordinates four programs for training evaluators with strong mathematics and science backgrounds. We also support the Evaluation Center at Western Michigan University to provide short, intensive evaluation training. By supporting month-long summer institutes for faculty and advanced graduate students, we also sought to provide evaluation that goes beyond awareness, but also is less than full-scale doctoral training. Finally, we initiated an Online Evaluation Resources Library developed by SRI, International. It provides materials for evaluators to use in their work on NSF projects.

In January 2000, we held a workshop for 40 invited specialists in evaluation training. Its purpose was to discuss and assess what was known about the impact of our efforts and how to proceed with them. Our planning continues. However, one issue that clearly emerged was the need to prepare minority evaluators. That finding led to this conference. Another issue that has arisen is how we can work most effectively with professional evaluation associations, particularly the American Evaluation Association. Work in that area is still in the planning stage. However, EHR's Evaluation Program intends to pursue a vigorous training effort and it welcomes input about how best to proceed.

SESSION ONE: Evaluation of Educational Achievement of Underrepresented Minorities

Session Chair

Beatriz Chu Clewell

Principal Research Associate and Director
Evaluation Studies and Equity Research Program
Education Policy Center
The Urban Institute
And former Executive Director
Commission on the Advancement of
Women and Minorities in Science,
Engineering and Technology (CAWMSET)
National Science Foundation

Presenters

Gerunda B. Hughes

Assistant Professor
Howard University/School of Education
Center for Research on the Education
of Students Placed at Risk (CRESPAR)

Carlos Rodríguez

Principal Research Scientist
American Institutes for Research

Discussant

Jane Butler Kahle

Division Director
Elementary, Secondary and Informal Education
EHR/NSF

Guiding Question

- Several Federal agencies — for example, NSF, NASA, DOE and ED — support science and mathematics education reform. Evaluation of their efforts includes paying attention to student academic achievement. What issues surround the evaluation of science and mathematics achievement, especially with respect to underrepresented populations? The discussion should highlight the cultural context of this area of evaluation in light of relevant literature.

Discussion Highlights

Gerunda B. Hughes

Session chair, Dr. Beatriz Clewell, opened the meeting by highlighting the importance of collecting, analyzing, properly interpreting, and reporting achievement data. The results can help educators and policy makers set an agenda for educational reform and improvement, she said. The benefits have been seen on the international level through the Third International Mathematics and Science Study (TIMSS).

The relatively poor showing of U.S. students when compared with some of their peers from other nations drew negative public reaction. The study revealed a gap in mathematics and science achievement the size of which few knew existed. The TIMSS results prompted many recent mathematics and science reform efforts in U.S. classrooms. In evaluating these reforms, asserts Dr. Clewell, there is a need to measure not only achievement, but also the factors that may influence it. This information fosters better understanding of student performance and facilitates the design of program and interventions modeled after "best practices." Such practices are most useful when they reflect the goals and objectives of American education, especially the goals to provide equal and equitable learning opportunities.

In light of the gap between white and underrepresented minority students in the area of mathematics and science achievement, we face an important question: Can we be any less diligent in implementing programs, policies, and practices to eliminate the national gap than in addressing the international gap? Dr. Clewell asserted the need to be even more diligent nationally because achievement by "American students" is inextricably tied to performance of all sub-populations, including "underrepresented minorities," who comprise an increasing proportion of the total student enrollment. NSF is a leader in support of projects and programs that improve mathematics and science achievement. But are NSF-funded education reform projects implementing effective evaluation designs that will capture data that is necessary to make sound decisions about project effectiveness and success? Regarding the assessment of achievement by underrepresented minorities, what issues, cultural and otherwise, must be considered?

Following Dr. Clewell, Dr. Carlos Rodríguez addressed three topics: contextual considerations, misconceptions and myths, and guiding principles. America is in transition in many ways, he said. Although minorities are quickly becoming the majority, critical masses of minority students remain scientifically and technologically illiterate. Dr. Rodríguez also noted that many people do not meet minimum levels of proficiency in literacy and quantitative capability. Most are Latinos, blacks, and poor whites. If these problems are ignored, many minorities will not receive the material rewards that follow competent performance, nor will they be able to participate fully in a democracy. According to Dr. Rodríguez, these circumstances pose serious implications for peace on the homefront and challenge the United States' position as a global leader in technology.

Dr. Rodríguez suggested that program evaluators bear much of the responsibility for bringing about effective reform. "What we do today in program evaluation is important because we are setting the pace, envisioning the future, changing the present," he asserted. "Part of what is setting the pace today is the call for high expectations (including high standards) and disciplined effort." Dr. Rodríguez added that, too often,

these noble goals are confused with high stakes testing, which is often found where students have not had an adequate opportunity to learn. "Are standards, frameworks and high stakes testing simply yet another excuse [for] telling minority students how deficient they are?", Dr. Rodríguez asked. If this is the tenor and context of current program evaluations, he continued, we have not progressed very far.

Dr. Rodríguez also warned that if our program evaluations produce still more "blaming the victim" litanies about minority students' bleak performance in mathematics and science without also examining their caretakers' willingness to improve, we do nothing but subscribe to myopic views. He said that the result is captured in the Spanish saying, "*El que adelante no va, atras se queda*" ("Who doesn't go forward, stays behind.") Dr. Rodríguez cited recent reports by the National Assessment of Education Progress that Hispanic eighth graders are more likely than non-Hispanic whites to take no science courses, while Hispanic and black students are more likely than any other group of students to take remedial mathematics and English classes. He said we continue to invest in schools and teachers like amateur gamblers, placing small bets but expecting huge winnings. It is no wonder, he said, that we have yet to realize either the ideals embodied in *A Nation at Risk* (1983) or the goals of *Goals 2000: Educate America Act* (1990).

Dr. Rodríguez identified the following misconceptions and myths, related to the dynamics of race, culture and language, which affect student learning and the organization of learning and opportunities to learn:

- Past inequities and inequalities have been addressed and no longer require attention.
- Merit can be defined by test scores.
- Fairness is best achieved through race-neutral policy.
- The goals of excellence and equity are irreconcilable.
- Test scores alone tell the whole story.

Dr. Rodríguez further noted that many variables can affect a student's test score, including:

- The quality of the student's education,
- The student's skill, ability or knowledge about a particular topic,
- Preparation for the test, or even
- What the student may have eaten for breakfast on the day of the test.

Dr. Rodríguez suggested that program evaluators and teachers need to spend time together to reflect on the diagnostic potential of tests and assessments to improve science and mathematics achievement among minority students. He also shared the five guiding principles of evaluation as adopted by the American Evaluation Association in 1994:

- Systemic inquiry,
- Competence,
- Integrity and honesty,
- Respect for people, and
- Responsibility for the general and public welfare.

“There are insidious notions in popularized ideas about the very ability of non-mainstream students to learn that appear ready to sabotage and derail any progress put forth by program evaluations, policy, research and practice,” Dr. Rodríguez warned. He said that political and ideological processes may essentially be determining the “improvement” of science and mathematics achievement among underrepresented minority students, and he added that program evaluators should not ignore or deny them.

Dr. Gerunda Hughes began her presentation by noting how often evaluation designs are very narrow in scope, focusing on the achievement of students as operationally defined by a test score. According to her, this often leaves many questions unanswered, especially when determining the effectiveness or success of a project that has been designed to improve science or mathematics achievement. Dr. Hughes noted that improved student achievement is a direct or indirect goal of many educational reform efforts but that academic achievement does not exist in a vacuum; it is correlated with factors that may or may not be within the control of the reform effort. To the extent that these factors are controllable, Dr. Hughes said, they may be planned for in the project design. If they are not controllable, they can at least be measured and statistically controlled in the analysis of achievement data. According to Dr. Hughes, school factors include: the quality of instruction; the opportunity to learn the material being assessed; and teacher characteristics such as efficacy, knowledge of content, teaching skills, attitudes and beliefs about the children’s capacity to learn and years of experience. Personal factors include cultural orientation and socioeconomic status. The assessment of all of these factors (and others) can increase understanding of the academic achievement of minority and non-minority students, Dr. Hughes said.

Like Dr. Rodríguez, Dr. Hughes noted that the goal is neither to belittle the test score nor to ignore standards. Minority students must be held to the same high standards as other students, if they are going to be prepared to compete in the global marketplace, she said. However, she suggested that the test score does not tell the whole story. She said program evaluators must be sensitive to the widely divergent educational experiences, backgrounds and cultures of students and explore the ways in which those factors interact with the cultures of teaching, learning and assessment. With such sensitivity, Dr. Hughes said, evaluators and other measurement professionals are better able to identify and use traditional or innovative approaches to testing and assessment — methods that yield valuable information about what individuals know and can do.

In addition, Dr. Hughes said, it is possible that changes in the focus of curricular goals have a major impact on the kind of data that program evaluators collect and how it is collected. As an example, she noted that the NSF Working Group on Assessment in Calculus acknowledged that the mathematics community needs to change fundamental ways of assessment at all grade levels because of an increased understanding of what it means to think mathematically. The group suggested the use of paper-and-pencil tests as well as performance tasks, open-ended items, investigations and projects, observations, interviews, portfolios and self-assessment. Although there is some overlap in their purposes, each alternative assessment method taps into a slightly different aspect of learning. Dr. Hughes said that this has implications for the professional development of teachers in assessment, testing and measurement, especially as they work with program evaluators to provide useful information that can inform instruction and explain outcomes more fully. In closing, Dr. Hughes challenged all evaluators, and minority evaluators in particular, to go beyond the obvious and pay attention to factors that correlate with student achievement.

Dr. Jane Butler Kahle, the session discussant, stressed the need to look for achievement trends, rather than to try to establish causality. She also reinforced the importance of disaggregated data, citing her recent study that showed different factors affected the science achievement of 8th grade African American children. She encouraged using multi-variant analyses, such as Hierarchical Linear Modeling, to analyze the complex factors surrounding minority achievement, as well as combining qualitative data with quantitative data. She stressed that until an adequate number of minority evaluators could be educated, efforts to educate majority evaluators in understanding minority issues, cultures, and contexts needed to expand.

Asked to identify the important points of session discussions, participants suggested the following guidelines:

- Utilize
 - quantitative and qualitative approaches in the evaluation design
 - multiple measures for project evaluation
 - multiple data sources
 - short-term, intermediate and long-term objectives
- Define “success” in multiple ways (e.g., not just in terms of test scores and standardized measures, but also in terms of increased attendance and decreased mobility rates; positive student-teacher interactions; increased parental involvement in the math or science education of the child; increased self-esteem among students; increased persistence rates; and improved attitudes about schooling).
- Disaggregate data according to race, gender and ethnicity (also along contextual lines, if appropriate).
- Include evaluators who are sensitive to issues of diversity (especially minority evaluators) and who will frame the right questions.

Papers/Presentations

Evaluation of Educational Achievement of Underrepresented Minorities: Assessing Correlates of Student Academic Achievement

Gerunda B. Hughes

The evaluation of the educational achievement of children involves more than analyzing and reporting test results. It carries the responsibility of providing important information to stakeholders that facilitates both internal and external decision-making about a project. This additional charge is reflected in the definition of evaluation as stated by the Joint Committee on Standards for Educational Evaluation (1981). In its report, the committee defined evaluation as "the systematic investigation of the worth or merit of an object." Similarly, Webster's International New Dictionary defines evaluation as "the examination of the worth, quality, significance, amount, degree or condition of [an object]." The target of most educational evaluation is a project, program or product that has as one of its primary or secondary goals, improved student achievement. The evaluation of achievement usually comes near the end of the project, for example, and is part of the summative evaluation. Inferences and conclusions about the success or failure of a project are drawn from data collected during the project. Sometimes these evaluations involve the use of controlled or matched groups and sometimes comparisons of student achievements are made on the basis of sex, race or ethnicity. In the latter case, what is often reported in the research literature and what characterizes many evaluation reports about minority students' educational achievement is limited to "the amount and degree of achievement." This is best illustrated by the first sentence of the first chapter in the book, *The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips. It states: "African Americans currently score lower than European Americans on vocabulary, reading, and mathematics tests, as well as on tests that claim to measure scholastic aptitude and intelligence (p. 1)." This statement may be true. But is it the whole story? Are there factors that may explain differences in the academic achievement of African Americans and European Americans? Can project staff plan effectively for the influence of such factors on project outcomes? If so, how? One way is by systematically assessing those factors known to correlate with student achievement. The additional information can provide greater insight into why certain outcomes, such as gaps in educational achievements, exist.

Goal of Educational Reform

Improved student achievement is, directly or indirectly, a major goal of educational reform efforts. Yet, the academic achievement of children does not exist within a vacuum. It is influenced by and correlated with a variety of school and personal background factors. These operate in ways to facilitate or inhibit the academic achievement of children in different contexts. School factors include how children are taught, how they are assessed, teacher expectations and opportunity to learn. Personal background factors include cultural orientation and socioeconomic status. Assessment of these factors in evaluation studies can lead to a more comprehensive understanding of the academic achievement of minority and non-minority students.

Researchers in the field of testing, measurement and assessment have noted that the systematic assessment of African Americans, Native Americans, Hispanics, women and persons of low socioeconomic status is as appropriate, as desirable and as necessary as it is for any other group (Davis, 1948; Johnson, 1979; Gordon, 1996). Additionally, there is no argument against the logic that individuals within these groups must develop the same body of skills and expertise that standards require. What is argued, however, is that a single test score does not reflect all of reality, that it should not be used as the sole basis for making inferences about individuals or groups. Rather, it is indeed necessary to look beyond the test score to the widely divergent educational experiences, backgrounds and cultures of the test taker and explore how these interact with the cultures of teaching, learning and assessment (Johnson, 1979; Ladson-Billings, 1995). By doing so, evaluators and other assessment professionals are in a better position to identify more effective approaches to testing and assessment that yield valuable information about what individuals know and can do. These approaches to testing and assessment very often take into account the multidimensional variables associated with educating people of color —especially how they are taught and how they learn.

Culture and Cognition

Issues about culture have always played an important part in schools. Though these issues were not always directly addressed, children were, nonetheless, judged and evaluated in terms of how much of the mainstream culture they espoused. In essence there is a positive correlation between the amount of mainstream acculturation one has acquired and achievement on traditional mainstream tests of cognition. In fact, not so long ago children who came from impoverished environments and who generally scored lower on tests of achievement and cognition were described as “culturally deprived.”

Fortunately, being “cultured” is no longer associated with having high test scores. In its most basic form, culture entails the way a particular group of individuals translates reality. Many of these translations are evident in how children respond to test items. Boykin (1994) and Giroux and McLaren (1986) suggest that culture embodies a set of practices and ideologies from which different groups draw to make sense of the world. It embodies belief systems and ways of knowing and valuing. Sometimes the culture of schooling and the assessment of what goes on in schools matches that which the children bring to the classroom, and sometimes it does not.

Delpit (1988) argues that a culture of power exists in American classrooms. That culture of power reflects the practices of those in power and consequently reflects mainstream culture. Some children come to school with the rules of the culture already understood. They come to school inclined to be receptive, if not endowed with the culture because of their previous experiences with it. Other children, however, do not have or have not been sufficiently exposed to the “rules of the culture” and hence, are directly or indirectly penalized for not knowing them. In other words, some children possess the requisite “cultural capital” while others do not. The latter are among the “culturally deprived.”

Boykin (1994) notes that the school culture and, I might add, the present mainstream “assessment culture,” reward individual possession of specific intellectual and social attributes. By emphasizing competition among individuals and devising reward systems to honor such accomplishments, schools and the testing professions reinforce what they value. But what if tests and assessments were designed to be aligned with the cultural integrity of the children and at the same time maintain the content validity they were designed to possess? It may be possible to demonstrate that the much talked about gap between black and white children is not as large as one might think — or at least there may be a way to explain the variance in the context of culture.

Snow and Lohman (1989) note that “cognitive study of the targeted aptitudes, achievements and content domains for which educational measures are to be built might suggest alternative measurement strategies and refinements for existing instruments.” With this perspective, they help provide a rationale for seeking alternative ways of assessing what different subgroups of the populations can do on tests and assessments. Furthermore, they note that much of cognitive research on the nature and development of ability suggests that learner experience, which includes out-of-school experience and the cultural context of the home, is an important determiner of what attributes are measured by any test and how many different attributes are measured.

Teachers and Teaching

The role that teachers play in the academic achievement of minority children cannot be overemphasized or understated. Teachers’ personal and cultural attributes as well as their attitudes, beliefs and behaviors are important. They influence self-concept and attitudes of students as well. Irvine (1990) notes that students identify teachers as significant others in their lives, and how a child feels about himself or herself may be — to a large extent — based on how the child perceives the teacher feels about him or her. Many children who believe their teacher does not like them, in turn, do not like themselves or school and eventually fail academically. This effect is exaggerated for low-income and minority students because they are more teacher-dependent and are more likely to hold the teacher in high esteem.

Teacher expectations of students’ performance are also related to students’ academic achievement and are mediated by factors such as the characteristics of the teacher and the students. In content areas, and especially in mathematics and science, achievement among minority students can be greatly inhibited by teachers’ low expectations simply because of the small numbers of minorities who are currently in the those fields.

Using Multiple Means of Assessment in Mathematics

In the Assessment Standards for School Mathematics, the National Council of Teachers of Mathematics (NCTM) (1995) encourages the use of multiple forms of assessment for both short-term and long-term instructional planning. Multiple forms of assessment provide evidence about student learning that may be difficult to capture by administering singular formats. Observations and questioning, for example, offer opportunities for understanding the influences of students’ unique prior experience. Furthermore, making valid inferences about students’ learning requires familiarity with every student’s response in a variety of modes, such as talking, writing, graphing or illustrating in a variety of contexts. The NCTM also notes that “culture considerations

are also important; however, care should be taken not to make assumptions based on cultural stereotypes, because each student has unique responses to experiences in and out of school" (p. 52). The recognition by the mathematics community of alternative ways to assess what students know and can do, is indeed a step toward accommodating cultural differences, where appropriate. Thus, in addition to the use of paper-and-pencil assessment tasks, the NCTM recommends the use of open-ended items, student-constructed tests, performance tasks, investigations and projects, interviews, portfolios and self-assessment in mathematics instruction and assessment. Finally, the expanded use of alternative assessments, including the use of performance assessments in the classroom, must be accompanied by appropriate training and professional development for classroom teachers and university faculty (Johnson et al. 1998).

Role of Evaluators in Mathematics and Science Projects

Clearly, the charge for minority evaluators is to go beyond the obvious. While there may be gaps in the academic achievement of minority and non-minority students, it is not enough to leave such gaps unexplained. By paying attention to the factors that are correlated with minority student achievement, and designing and implementing evaluation models that measure these factors, evaluators can inform the instructional and assessment practices that aim to improve student achievement in general, and minority student achievement in particular.

References

- Boykin, A. W. (1994). Afrocultural expressions and its implications for schooling. In E. R. Hollins, J. E. King, & W. C. Hayman (Eds.), Teaching diverse populations: Formulating a knowledge base (pp. 225-273). Albany, NY: State University of New York.
- Davis, A. (1948). Social class influences upon learning. Boston: Harvard University Press.
- Delpit, L. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. Harvard Educational Review, 58, 280-298.
- Giroux H. & McLaren, P. (1986). Teacher education and the politics of engagement: The case for democratic schooling. Harvard Educational Review, 56, 213-238.
- Gordon, E. W. (1996). Toward an equitable system of educational assessment. Journal of Negro Education, 64(3), 360-372.
- Irvine, J. (1990). Black students and school failure: Policies, practices, and prescriptions. New York: Praeger.
- Jenck, C. & Phillips M. (1998). The Black-White test score gap. Washington, D.C.: Brookings Institution.
- Johnson, S. (1979). The measurement mystique: Issues in selection for professional schools and employment. Washington, D.C.: Howard University, Institute for the Study of Educational Policy.
- Johnson, S., Thompson, S., Wallace, M., Hughes, G., Manswell-Butty, J. (1998). How teachers and university faculty perceive the need for and importance of professional development in performed-based assessment. Journal of Negro Education, 67(3), 197-210.
- Joint Committee on Standards for Educational Evaluation. (1981). Standards for Evaluation of Educational Programs, Projects, and Materials. New York: NY: McGraw-Hill.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. American Educational Research Journal, 35, 465-491.
- National Council of Teachers of Mathematics (1995). Assessment Standards for School Mathematics. Reston, VA: Author.
- Snow, R. E. & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), Educational measurement (pp. 263-331). Phoenix, AZ: The Oryx Press and American Council on Education.

Presentation

Assessing Underrepresented Science and Mathematics Students: Issues and Myths

Carlos Rodríguez

Good Morning. Thank you for the opportunity to be here. I have prepared my remarks to focus on three areas related to the question we are to address today. What are the issues surrounding the evaluation of science and mathematics achievement, especially the academic assessment of underrepresented populations? I understand that the relevancy of this question to you and the purpose of this meeting are in the context of science and mathematics program evaluation and the training of science and mathematics program evaluators. First, I will offer you some contextual considerations; secondly I will dispel some common myths about assessments and minority students; and thirdly, I will articulate some principles that should guide evaluation practices that I hope you take with you as tools in the important work you do.

Contextual Issues

"Many in our society and its educational institutions seem to have lost sight of the basic purposes of schooling, and of the high expectations and disciplined effort needed to attain them." High standards, high stakes testing and even program evaluations rarely talk about high expectations and disciplined effort.

What I have just presented to you is paraphrased from the introduction to *A Nation at Risk* written 17 years ago in 1983. Almost two decades have passed and these words are as true today as they were then. If minorities are quickly becoming the majority, does this mean that we should expect even greater failure rates, even less success? We answer this question with a resounding "NO."

Success, access and opportunity, however, do not happen accidentally for most people. They happen deliberately, and unfortunately, slowly. What we do today in program evaluation is that much more important because we are setting the pace, envisioning the future, changing the present. There is a saying in Spanish: "*El que adelante no va, atraz se queda*" ("Who doesn't go forward, stays behind.")

Many people in the United States today do not possess the minimum levels of know-how in literacy and numeracy, and training essential to this emerging millennium— and they are mostly Latinos, blacks, and poor whites. In fact, with Latinos this is true for almost half of the total U.S.-born population, and slightly less so for African Americans. Many then, of these groups, our brothers and sisters, will effectively be left out, not simply from the material rewards that accompany competent performance, but also from the chance to participate fully in our democratic society. These levels of educational achievement are inadequate when 80% of all new jobs this century will require at least some post-secondary education or training (Carnavale, 1999).

Without enumerating all the data, you should know as well as I do that, with critical masses of our minority students, we are raising a new generation of Americans of color that is scientifically and technologically illiterate. There is a real world of techno haves and have-nots. With program evaluations, we cannot be content to simply continue to represent the gap.

Program evaluators have the opportunity to close the gap by identifying the missing pieces in program interventions — the missing links, if you will. Program evaluators have the opportunity to build the bridges among policy, practice and research.

Schools may be emphasizing such basics as reading and computation at the expense of other essential skills such as comprehension, communication, analysis, problem solving and drawing conclusions. Do you know what the standards-based curriculum really contains in your state, school district, schools and classrooms? Do you know how well your testing programs are aligned with these standards and the enacted curriculum (what's really going on behind closed doors)? How well do you know if the standards are really changing how the poor and minorities are taught?

If you are not engaged in answering these kinds of questions in your evaluation efforts, you have a lot of work to do. Or have standards and frameworks and high stakes testing simply become yet another excuse for telling minority students how deficient they are? The training of program evaluators in science and mathematics must include deep understanding of contextual issues, and curriculum, and pedagogy, and assessments — both the promises and the limitations of each of these things. No mean task. But remember, comprehensive problems require comprehensive solutions. Comprehensive solutions can be built from comprehensive program evaluation approaches. Piecemeal solutions and Band-Aid approaches yield piecemeal and Band-Aid solutions.

If, for example, as we evaluate the loss of our black and Latino students from science and mathematics programs funded by NSF, NASA, the JPL, and the like, we stay at the level that reiterates the deleterious effects of resource-poor educational opportunities, without condemning the insistence of the state or district to perpetuate same, we relegate our greatest resource, our students, to shame. If our program evaluations produce yet more blame-the-victim litanies about minority students' bleak performance in mathematics and science, without also examining the willingness of their interveners to correct same, we do nothing more than subscribe to myopic views.

Let me turn to mathematics and science, then, for a few more minutes. We know that the vast majority of elementary school teachers are not prepared to teach mathematics and science. In fact, we can observe that the elementary curriculum is primarily organized around the language arts. At the middle school level, mathematics and science courses introduce students to inductive and deductive reasoning and provide them with experience in the practice of science. Minority students who do not have full access to these courses are unlikely to do well in the high school mathematics and science courses that are prerequisites for college entry. Hispanic eighth graders are more likely than non-Hispanic whites to be taking no science courses, and Hispanic and black students are more likely than any other group of students to be taking remedial mathematics or English courses. Most Hispanic high school students lag substantially (four years) behind non-Hispanic whites in science and mathematics proficiency, and black students are not far behind. Hispanic and black students are often denied access, that is, "tracked out" of regular science and mathematics courses. Minority students also are generally "tracked into" non-college preparatory courses. Today, black and Hispanic students are receiving, on average, more vocational course credits than academic credits, and are less likely to take algebra, geometry, and science courses — the minimum requirements for college admission. Did you know that only six states require algebra 1 of all students for high school graduation and none require

geometry? We need to get serious about high standards and make school boards and state departments of education accountable to provide the level of resources required for high standards of learning. We continue to invest in schools and teachers like gambling — we place small bets and expect huge winnings. Yet, we seem not to have any difficulty in finding enough money to barrage students with tests. Can you guess which groups of students are getting the better results with educational reforms based on high standards and high stakes testing?

So we see a rather easy-to-identify trajectory for the poor performance of most minority students in mathematics and science through high school. Even in high school, we know we have a cadre of influential and entrenched mathematics and science teachers who do not believe that the ability to learn mathematics and science is ubiquitous, but is really reserved for a select few. We organize learning in school as if intelligence is not normally distributed.

Dispelling Misconceptions and Myths

Program evaluators of science and mathematics often need to integrate some measurement data of academic performance, i.e., assessment and test scores.

So let's get to the misconceptions and myths about testing and assessment. This information should be axiomatic to the competent program evaluator of minority student academic achievement in science and mathematics.

Misconceptions (Change et al., 1999) that relate to the dynamics of race, culture and language affect student learning and the organization of learning and opportunities to learn:

Misconception One

Past inequalities in access and opportunities that racial and ethnic groups have suffered have been sufficiently addressed and no longer require attention. If this were true, most of this meeting would be unnecessary and the millions of dollars NSF invests in URM (Underrepresented Minority) programs a fraud. The fact is that low-income and minority children have significantly poorer access to quality schooling experiences, are concentrated in resource-poor schools and, due to persistent tracking and ability grouping modes, are usually found in the lowest groups.

Misconception Two

Merit can be defined by test scores. There appears to be a fairly common public notion that there are ways of measuring merit that are fairly precise and scientific, which may be true psychometrically. However, while tests may be shown to be statistically sound, policies based on such narrow definitions of merit inevitably exclude some students. The factors that determine merit and capacity for success — a mixture of ability, performance, talent and motivation — are not measured by standardized tests. The misuse of test scores beyond which they have been validated has had a systematic adverse effect on minority students.

Misconception Three

Fairness is best achieved through race neutral policy. This misconception contends that all individuals, regardless of race or ethnicity, should be judged on the same established criteria of competence, which are considered objective. Using the same standards, however, to judge individuals from majority and minority groups is unfair because differences in opportunities to learn prevent groups from having equal opportunity.

Let me turn to some common myths (Coleman, 1999) that also should be axiomatic to program evaluators when making the case about minority student academic achievement in science and mathematics:

Myth One: The Goals of Excellence and Equity Are Irreconcilable

Excellence and equity are not irreconcilable. They are neither intrinsically nor necessarily competing theories. Only when excellence is defined as flowing from some idealized notions of level playing fields, objectified neutral knowledge and meritocracy are excellence and equity irreconcilable. The educational foundations that guide policies promoting excellence can be — and should be — fully aligned with the promotion of equal opportunity for all students.

Myth Two: Test Scores, Alone, Tell the Whole Story

The value that test results can provide when making educational decisions about students does not mean that test scores should, as a matter of good educational practice, trump the need for thoughtful educational decision-making. A test's value as an educational tool is dependent upon its design, the context in which the test is administered and the ultimate uses of the test. Even when a test is used for purposes consistent with its design, a test is one tool among many. Just as tests are not perfect barometers of learning, conclusions based on those test results are not always error free. Many variables can affect a student's test performance, including: the quality of the student's education; the student's skill, ability or knowledge about a particular topic; preparation for the test; or what the student ate for breakfast on the day the test was administered. Does this mean that we should do away with tests? No. What it does suggest is precisely what test measurement standards affirm: the importance of considering multiple and educationally appropriate measures when making life-defining decisions about students. In 1985, the American Psychological Association (APA) Standards for Educational and Psychological Testing reminded us, for instance: "In elementary and secondary education, a decision that will have a major impact on a test taker should not automatically be made on the basis of a single test score" (APA Standard 8.12).

Conclusion to Misconceptions and Myths

Ultimately, good educational practices highlight the importance of considering objective measures such as tests in appropriate ways when making decisions about students. Not all assessments and tests are created equal, and tests should be used in ways that are valid for the particular purpose for which they are used. We must guard against cookie-cutter approaches to learning and assessment. Human beings are much

too complex to be reduced to learning and assessment approaches that assume that everyone, for example, learns how to read and solve problems and perform calculations in the same way.

Let me also sum up the challenges to school reform, including the reform of assessment: As states, districts, and schools use elements of standards-based pedagogical and curricular reform as well as standards-based assessment reform to enhance education capacity, they will face several continuing challenges.

The most critical challenge is to place learning at the center of all reform, assessment and program evaluation efforts — not just improved learning for students, but also for the system as a whole and for those who work in it. If the improvement of the learning experience for URMs is not a primary goal of the assessment practices and program evaluations in science and mathematics, they are not worth doing. If this improvement is not the tangible, realized goal of assessment and evaluation data, what end does it serve? For if the test givers — i.e., funders, teachers, administrators — are not themselves learning how to improve learning through assessment and program evaluation, and if the system does not continually learn from practice, then there appears to be little hope of significantly improving opportunities for all our youth to achieve to the new standards (CPRE 1995). This is what should be the grist of effective program evaluation — what the evaluator can say about the observed linkages between input and outcome with due consideration for the contextual considerations, while capturing all the nuances of unanticipated outcomes.

In my opinion, program evaluations need to contribute to the development of coherent and strategic approaches to capacity-building that take into account the needs and goals of the individual learner, school, district and state, not just for the immediate initiative, but for the long term. Resources are obviously a critical aspect of organizational capacity. A key target in addressing resource needs will be expanding available time to school personnel — time for teachers to collaborate in planning and assessing their instruction; time for teachers and administrators to participate in learning opportunities outside the school; and time for reforms to mature without falling prey to policymakers' readiness to halt reform if student test scores do not rise immediately. Allowing schools and districts to reconfigure schedules to provide time for teacher collaboration and learning is possibly the most cost-effective means of providing at least some of the additional time required. Finally, teachers and program evaluators need time together to reflect on the diagnostic capacity of assessments and the program evaluation itself to improve science and mathematics achievement among minority students.

Principles to Guide Evaluation Practices

In 1994, the American Evaluation Association adopted the following principles. The order of these principles does not imply priority among them; priority will vary by situation and evaluator role.

Systematic Inquiry: Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.

Competence: Evaluators provide competent performance to stakeholders.

Integrity/Honesty: Evaluators ensure the honesty and integrity of the entire evaluation process.

Respect for People: Evaluators respect the security, dignity and self-worth of the respondents, program participants, clients and other stakeholders with whom they interact.

Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.

Let me warn you, however, about certain seductions. Evaluators and educators must be very careful not to assume or be seduced into thinking that if program evaluations, policy and research and practice are perfectly linked, school improvement and advancement of minority students in science, mathematics or any subject will result as logical occurrences. There are insidious notions in popularized ideas about the very ability of non-mainstream students to learn that appear ready to sabotage and de-rail any progress that program evaluations, policy, research and practice put forth.

All are susceptible to manipulation. This is the "Who benefits?" question. Allington and Woodside-Jiron in the November 1999 *Educational Researcher* provide compelling and disturbing evidence, for example, for major policy manipulation at the state level that is currently on-going and is targeted at implementing a more "code-oriented" or phonics emphasis curriculum framework for reading instruction in the states, indeed nationally. In this analysis, they uncover the masquerades of research in the form of "expert opinion" to promote a one-size-fits-all approach to developmental reading instruction. They cite Benviste's work, *The Politics of Expertise*. In this work, Benviste indicates that through the political use of expertise, policy advocates consolidate a monopolistic position by promoting the appearance of an external professional consensus on a policy issue, often achieved by using highly selective research teams whose advice may not be easily dismissed.

Another example may be given drawing on *Issues in Education Research* (1999) edited by Ellen Condliffe Lagemann and Lee Shulman. Theodore Mitchell and Analee Haro in their chapter "Poles Apart: Reconciling the Dichotomies in Education Research" acknowledge that scholarly efforts "to put research knowledge into practice, working collaboratively and in mediated fashion, have been extraordinarily powerful, engaging teachers, researchers, and parents in discussions of educational aims and means and developing a sense of common purpose around the task of educating children." They warn, however, that "effective practices" often become cookie-cutter formulas for success, rather than tools for continued improvement. They posit this as a consumption problem — of a national appetite for solutions to the frustrating and complex problems facing education. A hunger that includes voters, parents, some school professionals, university presidents and, increasingly, funders of education inquiry and practice. Thus are fed the reductionist approaches that seek to take carefully crafted interventions and wholesale them. This is very appealing and seemingly cost-effective, and seductive.

Certainly it can be hypothesized that program evaluations, policy, practice and research can be aligned, and examples given. But we must be diligent at unmasking and respecting the discreet elements within each domain that may really be at play. The political underpinnings of program evaluation, policy, research and practice are real, yet not often made explicit or articulated. Political underpinnings refer to values, belief systems, power arrangements and divisions of labor.

Margaret Barrego Brainard reminds us that the word assessment is derived from a Latin verb *assidere*, which mean to sit beside. Evaluate comes from the Latin *valorem*, which means to place a value on something or to ascertain the value of something. Neither assess nor evaluate means to sit on top of or to hold back or to judge. They suggest that in order to reveal what a student really knows, it is necessary to be close to them, perhaps even moving alongside them on a path of learning. It means that we are challenged to see all our students, the teachers who teach them and those who organize their learning environment, the faculty that research them, and, yes, especially our minority students — we are challenged to see them in new ways, especially in ways that stop telling us that they are unteachable; in ways that convince us that each of us can contribute to increasing the degree to which our schools and our evaluations of the programs to enhance student learning can serve the full range of diversity that our students represent.

Where is the voice today of program evaluation surrounding the evaluation of science/mathematics achievement, especially the academic assessment of underrepresented populations, outside this conference? Outside this conference, it is in the voice of high standards and high stakes testing in the public discourse. Which, (at the risk of sounding repetitive) if they do not provide diagnostic information upon which to improve student's learning, then they are but another face of blaming the victim. Blaming the victims because they attribute the causes of failure to deficiencies in the social, cultural and linguistic experiences of the students' themselves, and tangentially, if at all, to the organization of learning and the resources committed to it. In short, political and ideological processes, more often than not, may be determining the improvement of science and mathematics achievement among underrepresented minority students, and should not be ignored or denied by program evaluators.

Recommendations

The following questions and answers I recommend as suggestions for the thematic content of this conference:

- How do we determine if adequate integration exists among program evaluators, administrators, teachers and faculty and students on issues affecting SCIENCE AND MATHEMATICS learners?
 - Early and often.
 - Through collaborative inquiry between education evaluators and practitioners. By engaging teachers as researchers and evaluators in action research/program evaluation projects, for example.
 - By taking the long view. By establishing short-term accomplishable goals and placing them in a long-term continuum.
 - By remaining aware that programmatic and policy interventions almost always have political implications. At the site or grantee level, program evaluation may often require political strategies and political thinking.

- How is integration manifested?
 - Through deliberate discourse building.
 - Getting evaluators, researchers, practitioners and policymakers in the same room and talking through the process and building consensus on the desired outcomes.
 - When people know they are doing it.
- What barriers exist to integration among program evaluation, practice, research and policy?
 - Contradictory political values.
 - The lack of willingness to engage in comprehensive approaches.
 - Ignorance.
- How can diverse stakeholders advocate for minority students in SCIENCE AND MATHEMATICS in ways that will advance improved achievement?
 - Early and often.
 - Change the discourse.
 - Influence the alignment of assessments to what is actually taught. Require the use of multiple measures of students' abilities to inform instructional decisions.
 - Dispel the myth of a single American "mainstream." As my colleague at ETS, Tony Carnavale, so succinctly states: The notion of a single American culture is inconsistent with American history and ignores the realities of the development of individual and group identities in the modern world.
 - Integrate into program evaluations the rich natural resources of linguistic and cultural diversity in our communities.
 - Maintain comprehensive, holistic, ways of seeing the effects of program interventions on minority student achievement in science and mathematics.

Summary

The promotion of program evaluation of science and mathematics performance among underrepresented minorities to advance academic achievement provides one of the few tools we currently have to promote inclusionary values that foster learning success in schools and education writ large. Inclusionary values honor, respect and give dignity to the innate wonder and beauty and promise of each and every child, of each and every student. There is nothing more worth doing.

References

- American Psychological Association (1985). *Standards for Educational and Psychological Testing*. Washington, DC: APA
- Carnavale, Anthony P. (1999). *Education=Success: Empowering Hispanic Youth and Adults*. Washington, DC: ETC/HACU.
- Change, Mitchell; Witt-Sands, Daria; Jones, James; and Hakuta, Kenji (Eds.) (1999). *The Dynamics of Race in Higher Education*. Stanford, CA: AERA, Center for Comparative Studies in Race and Ethnicity.
- Coleman, Arthur E. (1999). Public briefing.
- CPRE (1995). *Policy Brief: Building Capacity for Education Reform*.

SESSION TWO: Participation of Minority Professionals in Educational Evaluation

Session Chair

Stafford Hood

Associate Professor

Division of Psychology in Education

Arizona State University

Presenters

Rodney K. Hopson

Assistant Professor

School of Education

and Center for Interpretive and Qualitative Research

Duquesne University

Discussant

Norma Dávila

Co-Principal Investigator

Puerto Rico Statewide Systemic Initiative

University of Puerto Rico, Rio Piedres

Guiding Questions

- What is the motivation for increasing the number of minority evaluators with advanced training and experience in the field of educational evaluation? (Minority evaluators are defined as persons from those groups underrepresented in the fields of evaluation.) What is the importance of including minority evaluators in the evaluation of science and mathematics education programs?
- What mechanisms are available to identify the current population of minority evaluators, in particular those with expertise and experience in science and mathematics education? (Information sources include a survey of professional organizations, university programs, etc.)

Discussion Highlights

Norma Dávila

This session included a brief presentation of two background papers by their authors. It is followed by a group discussion, a summary of the discussion and identification of issues to be addressed in the future by the discussant.

The paper by Rodney Hopson presented the problem of the lack of available mechanisms to identify the current population of minority evaluators. It took the position that: "sustainable substantive interventions and mechanisms are needed to increase the number of minority evaluators with advanced training and experience in the field of educational evaluation, particularly those with expertise in science and mathematics education" (p. 1). The paper's theme was "that any proposals garnered to address the paucity of minority participation in educational evaluation were necessary but insufficient unless they both provided structural and adequate nurturance to a pipeline of scholars and created conditions whereby minority evaluators sought to advance knowledge production in the field for the purpose of emancipation and social justice" (p. 1).

Hopson explained that the role of professional organizations in addressing this problem was just getting started. He stated that organizations such as AEA [American Evaluation Association] were beginning to build diversity among the evaluation community including the development of mentoring programs within the organization. He further suggested that to be effective efforts to increase the numbers and participation of minority evaluators should include the use of multiple strategies.

Two major questions framed the paper by Stafford Hood: 1) "What ought to be the motivation for increasing the number of minority evaluators with advanced training and experience in the field of educational evaluation?" and 2) "What might be the benefits of including minority evaluators in the evaluation of science and mathematics education programs?" Hood believed that the motivation to increase the number of minority evaluators stemmed from practical considerations, since urban districts have increasing numbers of minority students. Yet very few minorities served as external or internal evaluators of educational programs in these districts.

Hood further stated that graduate programs have not done enough to address this problem, yet there were many contributions to be made by minority evaluators. Among the potential contributions were: better understanding of the programs serving minority students, value of the programs for those intended to be served, refinements to improve the benefits of the programs, interpretation of the meaning based on "who's doing the looking?" and observation and translation of the non-verbal communication that occurs in interactions. The author also raised the issue that, according to the Program Evaluation Standards, it is necessary to include stakeholders in the evaluation, particularly minorities. However, in the case of minority evaluators, racial similarity should not be the only requirement considered beyond competency. The paper ended with an invitation to look at covariates of success for achievement in mathematics and science by children of color.

The group discussion began with agreement that more minority evaluators are needed. Dr. Judith Sunley, Interim Assistant Director of EHR, clearly recognized this need in her welcoming remarks. The group expressed the belief that the identification and description of minority evaluators are important steps to justify the creation of

programs to increase their numbers. As part of this discussion, it was suggested that NSF conduct a survey to determine the number of minority evaluators and their work settings. There also was a proposal to develop a framework and plan for a database of the current pool of minority evaluators in school districts, state education agencies, colleges and universities, as well as in professional organizations, as part of the identification and description process.

In terms of strategies to recruit minority evaluators, some group members suggested that we think across the educational pipeline to identify and train minority evaluators. We also need to make a list of the “cogs” of the wheel such as universities, government agencies and professional organizations that can help to identify and to produce minority evaluators. Another suggestion was that we develop recruitment strategies at the undergraduate level to tap future evaluators before they enter graduate school. Two additional suggestions were to recruit an advisory team of education research and evaluation faculty from minority-serving institutions to plan a pipeline program and to create an interagency working group to develop a paper on research and evaluation needs.

Several strategies to train minority evaluators were introduced. The need to emphasize the importance of building cultural sensitivity at multiple levels and to consider cultural context in the evaluation of programs was stressed throughout the discussion. It was proposed that colleges and universities (including predominantly minority serving institutions), government agencies and professional organizations develop training programs for prospective minority evaluators; their proponents could look at training centers for models for these programs. Another strategy was to identify the best methods to recruit and train minority evaluators. Because it is so difficult to find evaluators who can be solely responsible for mathematics and science projects, the group debated whether the training of minority evaluators should be based on content or on strategies and suggested that individuals should be trained across disciplines and across levels (i.e., social sciences, undergraduate, master’s level). Further, an intern program for minority evaluators was presented as a training alternative, and the need for incentives for students and institutions to support the training of minority evaluators was highlighted.

In addition to identifying and training potential minority evaluators, design strategies were suggested to keep these evaluators working in the field. The group proposed four major strategies to address this need: 1) have senior faculty at universities serve as mentors and role models; 2) identify mechanisms to nurture a critical mass of minority evaluators within universities; 3) identify potential employment opportunities for minority evaluators and make this information available to eligible candidates; and 4) utilize evaluation teams that are culturally diverse and include subject matter content experts.

The group addressed the importance of supporting minority evaluators from multiple types of agencies. It was stated that agencies should fund programs that will prepare evaluators — especially minority evaluators — effectively. It was further suggested that agencies need to build on current successful efforts, particularly where a critical mass of individuals is currently involved in the training of evaluators, particularly minority evaluators. It was suggested that NSF could become the leader in increasing training opportunities for minority evaluators among federal agencies.

The discussion and subsequent written comments by the participants raised the following questions for future consideration:

- How do we connect the conversation about student achievement with the issue of capacity building for minority evaluators?
- How do we increase the number of minority evaluators who can think systemically?
- How do we “bring minority evaluators into the loop?”
- Is there a difference in the evaluation process in different settings — schools (K-12), colleges and universities, research labs, business and industry — or would one model suffice for all?
- How deeply should research labs and science entities be involved in training evaluators for projects with an emphasis on science and mathematics?
- What is AEA’s involvement in this effort? (Participants were aware of Rodney Hopson’s activities in AEA, and wondered if there were other efforts.)

The discussant concluded with several reflections. First, the discussant pointed out similarities between the issues related to the identification and training of minority evaluators and those faced by many educational reform efforts. She stressed the importance of identifying potential future minority evaluators early in their careers since, if students do not make it through the pipeline, they cannot become evaluators later on. She emphasized the importance of using a multi-prong approach to identify evaluators so that specific programs target a critical mass of people, while other interventions focus on work with individuals. She suggested that program designers look beyond incentives when making decisions about ways to identify and train minority evaluators; for instance, redesign program structures within higher education institutions and other organizations. She urged the participants to look beyond NSF to other agencies that could also address this important problem of the evaluation field.

Papers/ Presentations

New Look at an Old Question*

Stafford Hood

This paper provides a starting point for our conversation about several provocative and critically important questions related to the evaluation of educational programs in general, with special attention given to mathematics and science programs. The two primary questions that frame this paper are: What ought to be the motivation for increasing the number of minority evaluators with advanced training and experience in the field of educational evaluation? And what might be the benefits of including minority evaluators in the evaluation of science and mathematics education programs? The paper further suggests the merit of looking beyond the dismal failure of African American, Hispanic and Native American students on standardized measures of mathematics and science achievement by mounting a systematic study of students from these groups who have achieved or are presently achieving in mathematics and science in secondary schools. I turn now to my first point.

In my opinion, the motivation for increasing the number of minority evaluators with advanced training in educational evaluation is a practical one. We live in a time when most of our major urban school districts have predominant enrollments of students of color, and evaluators of color are generally absent in the evaluation of educational programs that serve these students. To make this point I do not think it is necessary to provide numbers, but rather ask the reader to rely on personal recollection. You are simply asked to recall the number of African American, Hispanic and Native American evaluators with whom you have come in contact at research and evaluation units in central school district offices, state departments of education, the U.S. Department of Education and private foundations. How many African American, Hispanic and Native American evaluators have you seen as members of external evaluation teams evaluating educational programs that target students of color or even directing such evaluations? If your experience parallels mine, your answer to these questions will be "very few."

One of the major reasons for the limited number of trained program evaluators of color is that those graduate programs with the capacity to train program evaluators have not done enough to rectify this situation. The most telling symptom is the dearth of doctoral degrees awarded to African Americans and other groups of color by institutions capable of doing so. My on-going monitoring of the National Center on Education Statistics Integrated Postsecondary Education Data System's data, reporting doctoral degrees awarded by institution, race, and program areas within the field of education at major research universities, supports my observations (see Hood and Freeman, 1995). These data are not perfect for making this point but they are enlightening. The need for more trained program evaluators of color is what they can contribute to "understanding" in the evaluation of programs serving students from this population.

In program evaluation the production of clear, useful and objective knowledge and the pursuit of understanding seeks to determine program worth. In this case, I emphasize the importance of the evaluation resulting in an "understanding" of the program, its

* Two previous papers have articulated many of the views included in this paper. They are: Hood, S. (1999). Creating Evaluation Strategies Appropriate for African Education. Paper presented at the First Annual Meeting of the African Evaluation Association. Nairobi, Kenya. September 1999; and Hood, S. (1998). Responsive Evaluation Amistad Style: Perspectives of One African American Program Evaluator. Paper presented at the Invitational Retirement Symposium for Robert E. Stake. University of Illinois at Urbana-Champaign. Urbana, Illinois. May 1998.

value for those who are intended to be served and its refinements to improve the benefits. I would argue that an evaluator's understanding of a program, as it functions in the context of culturally diverse groups, is the most critical dimension for evaluating programs that serve these populations.

We must honestly assess whether current evaluation practice concerning diverse people does not systematically ignore potentially important aspects of diversity. When evaluators attempt to derive meaning from data gathered in cultural contexts they do not understand, their efforts are limited at best and potentially hurtful at worst. We must safeguard against producing evaluative knowledge "that seems counter intuitive to the [culturally diverse stakeholders] and seems to contribute little to our understanding of the people" (Gordon, 1998) and the programs which intend to serve them. Gordon referred to the work of anthropologist Michael Jackson, who queried "whether the lived experience is a necessary condition for valid observations." It was his view that "there was a possibility of our inability to understand the experience of the other." In my opinion, central to the observation is the meaning of what has been observed. Who is doing the looking is central to the process of evaluation.

When observing African American participants, for example, much in the program under evaluation can be lost long before reaching a summary "understanding" of the merit and worth of what has been observed. For example, too often nonverbal behaviors are treated "as error variance" by the observer and therefore ignored. Akbar (1975) asserts that African Americans "[rely] on words that depend upon context for meaning and that have little meaning in themselves.. [while also]...using expressions that have meaning connotations." Therefore the review of interview transcripts without the ability to interpret meaning based on these unwritten rules would likely result in interpretations that are more frequently wrong than right, thereby limiting communication and ultimately understanding between the African American participant/ stakeholder and the evaluator. I would expect that these concerns regarding "observations" and "translation" are equally true for other groups of color. A brief look at the program evaluation standards may also be instructive.

The second edition of the Program Evaluation Standards: How to Assess Evaluations of Educational Programs by the Joint Committee on Standards for Educational Evaluation (1994) has been offered to provide guidance to effective evaluation. The standards are organized around four major areas: Utility; Feasibility; Propriety; and Accuracy. The utility standards suggest that evaluators are required to "acquaint themselves with their audience, define the audience clearly, ascertain the audience's information needs, plan evaluations to respond to these needs and report the relevant information clearly and in a timely fashion." One of the ways to achieve this objective is through the identification of stakeholders.

The standards suggest that: 1) it is necessary to include "less powerful groups or individuals as stakeholders, such as racial, cultural or language minority groups;" 2) determine how the respective stakeholders view the evaluation's importance, would like to use the results and what information will be particularly useful; and 3) include the clients and stakeholders in designing and conducting the evaluation. Very few of the program evaluations with which I am familiar in American urban schools and other settings have shown a significant level of participation by the less powerful stakeholders. Additionally, I found it interesting that the standards did not consider the importance of these stakeholders assisting the evaluator to interpret the evaluation's results.

I find the standard on evaluator credibility to be important for culturally responsive evaluation. This second utility standard requires that: 1) the evaluators be trustworthy and competent to conduct the evaluation; 2) they be knowledgeable of the social and political forces affecting the less powerful stakeholders and utilize this information in the design and conducting of the evaluation; and 3) the work plan and composition of the evaluation team be responsive to key stakeholders.

Very fine guidelines — I have found little evidence of them being implemented in the evaluation of American urban schools. There may in fact be different views regarding evaluator credibility. Grace (1992) offers a different perspective on evaluator credibility when working with African American communities. She indicates that “in many cultures the age, race, sex, and credentials of the evaluators may have a significant impact on the evaluation process...[in the case of African Americans] — all things being equal — the most influential and respected members of the evaluation team are likely to be older individuals with academic credentials related to their expertise as evaluators.”

Grace and I agree that every effort should be made to include African American evaluators who are positively identified with the black community among the members of the external evaluation team, but racial similarity should not be the only requirement considered beyond competency. Grace also suggests that potential evaluators should be interviewed using questions designed to tap culturally relevant knowledge, attitudes and skills as the best indication of a candidate’s suitability for the job as an evaluator. Something for us to think about.

I trust that the need for increasing the supply of trained program evaluators of color has been established. I trust further that the benefits of including program evaluators of color throughout the evaluative process have been established to the reader’s satisfaction. If this is the case, then we may turn to the ultimate question for mathematics and science educators, and evaluators as well.

What do we know about the antecedents and conditions that are needed for meritorious academic achievement in mathematics and science by children of color? I’ll provide the answer: very little. Every experienced evaluator has discovered children of color who have distinguished themselves in either mathematics or science. Why these and not others? Further, even a modest historical search will reveal that African American educators, for example, have achieved remarkable success as scholars in science and mathematics. How was this accomplished? Similarly, there is a documented trail of African American educators who have contributed to the evaluation literature. In a word: America has produced high achievement in mathematics and science by people of color. How? Rather than participate in yet another examination of the correlates of failure by learners of color in the mathematical, biological and physical sciences, I envision a study of the covariates of success. To that end, we can begin by increasing the supply of trained people of color in educational evaluation and their participation in the evaluation of educational programs in urban schools.

References

- Akbar, N. (1975). Address Before the Black Child Development Institute Annual Meeting. October 1975. San Francisco.
- Gordon, E.W. (1998). Producing Knowledge and Pursuing Understanding: Reflections on a Career of Such Effort. AERA Invited Distinguished Lectureship. Presented at the Annual Meeting of the American Educational Research Association. San Diego, April 13, 1998.
- Grace, C.A. (1992). Practical Considerations for Program Professionals and Evaluators Working With African American Communities. In M.A. Orlandi (Ed.). Cultural Competence for Evaluators: A Guide for Alcohol and Other Drug Abuse Prevention with Practitioners Working with Ethnic/Racial Communities. Rockville, MD: U.S. Department of Health and Human Services.
- Hood, S. and Freeman, D. (1995). Where do students of color earn doctorates in education? The "Top 25" Colleges and Schools of Education. *Journal of Negro Education* 64, 4.
- The Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards*, second edition. Thousand Oaks, CA: Sage Publications.

Toward Participation and Liberation in Educational Evaluation: Developing Educational Pipelines to Increase Minority Evaluators

Rodney K. Hopson

Liberation is a value worthy of science. That should be the perspective from which the minority scientist seeks to advance knowledge, always in the spirit of respect for logical canons, multiple perspectives, and methodological rigor; not for the purpose of simply predicting, controlling and understanding, but for the purpose of emancipating (liberating) the bodies, minds, communities, and spirits of oppressed humankind (Gordon, Miller and Rollock, 1990:19).

Introduction and Problem Identification

The comments in this paper related to participation of minority professionals in educational evaluation directly address the mechanisms — or more accurately, lack thereof — available to identify the current population of minority evaluators*. This paper takes the position that sustainable, substantive interventions and mechanisms are needed to increase the number of minority evaluators with advanced training and experience in the field of educational evaluation, particularly those with expertise in science and mathematics education. The theme of this paper is that any proposals garnered to address the paucity of minority participation in educational evaluation are necessary but insufficient unless they both provide structural and adequate nurturance to a pipeline of scholars and create conditions whereby minority evaluators seek to advance knowledge production in the field for the purpose of emancipation and social justice.

The lack of mechanisms to confront the participation of minority professionals in educational evaluation, especially those with experience in science and math education, are symptomatic of larger and more systemic issues. That is, understanding the dearth of minority evaluators means putting into perspective the condition of education for blacks, Latinos, Native Americans and other underrepresented groups much earlier in the educational pipeline.

Reporting on the situation for blacks and Latinos, the most recent National Center for Education Statistics (NCES) edition of *The Condition of Education* highlights well-recognized disparities (USDOE, 1999). Despite a pattern of declining gaps in National Assessment of Educational Progress (NAEP) performance scores reported between the early 1970s and mid to late 1980s, the average performance for both blacks and Latinos in math, science and reading have remained lower than those of whites throughout elementary and secondary schooling. By the time these same groups are ready to enter the next level of schooling, gaps between whites and blacks and between whites and Latinos in their rates of enrolling in college immediately after graduating from high school have increased. College completion rates, moreover, depict a similar story as black and Latino rates are lower than their white counterparts.

* Much of the thinking reflected in this paper is based on two other papers; one paper in a recent issue of the *American Journal of Evaluation* (Hopson, 1999) and one co-authored paper presented during a presidential plenary at the most recent American Evaluation Association annual conference (Hopson and Rogers, 1999).

These patterns of achievement that manifest throughout the educational pipeline for blacks and Latinos are hardly novel; they continually speak to the worsening crisis in the education of these particular Americans. In fact, as Claude Steele has pointed out (1992), these disparate trends are the rule in most grade, high, undergraduate and graduate schools (including the most elite) as they pertain especially to African Americans. His message is alarming, but clear: despite beginning school with test scores fairly close to the test scores of white students their age, the longer blacks stay in school, the further they fall behind.

Framing the Problem of the Participation and Pipeline of Minority Evaluators

The problem of the participation and pipeline of minority professionals in educational evaluation cannot be fully explained and understood by the supply problem (i.e., the lack of “qualified” minority professionals is due to the small numbers of graduate level students) nor by deficit theories (i.e., the problem of academic achievement among poor and underrepresented groups stems from disadvantaged home and environmental backgrounds). Rather, the problem of the participation and pipeline of minority evaluators most directly lies within the arena and structure of schooling in the United States.

Within this schooling structure, as many critical educationists suggest (Apple, 1996; Bourdieu and Passeron, 1977; Giroux, 1983; House, 1999; MacLeod, 1995; Spring, 1997; Yeakey, 1981; Yeakey and Bennett, 1990), mechanisms and ideologies are constructed to prevent lower and working class individuals (and marginalized or underrepresented groups) from advancing into the upper strata of the social class structure. The very history of educational program evaluation surrounding achievement tests, as Edmund Gordon reminds us, has long promoted deficit-oriented explanations pertaining to minority and disadvantaged populations (1977:31).

When we first turned attention to the problems of educating educationally and socially disadvantaged children, a great deal of attention was given to the special characteristics of this population. The notions that dominated were largely determined by conceptions of this population as homogenous with respect to conditions of life and behavioral characteristics — we assumed a pervasive “culture of poverty.” The population was largely identified by its deficits in comparison to characteristics assumed to be typical of the white middle class.

That an inescapable racial devaluation is faced by children, students and scholars of color who attempt to matriculate through schooling is relevant to our discussion, particularly as we offer ways to counteract the societal preconditioning to see and expect the worst in certain minority groups.

Building Mechanisms to Increase Participation and Pipeline of Minority Evaluators: Focus on the American Evaluation Association

Increasing the participation and pipeline of minority professionals in educational evaluation inevitably involves increasing the number of minority scholars, particularly from underrepresented groups, in the discipline of educational evaluation. The role of professional associations, such as the American Evaluation Association (AEA), in

addressing participation and pipeline issues by increasing the number of minority evaluators and building the capacity of all evaluators to work cross-culturally is now only beginning to take shape (AEA, 2000).

Due perhaps to the impetus from several important persons in key leadership positions in AEA, including the attention to greater inclusiveness of marginalized groups by major plenary speakers at recent annual conferences (Chelimsky, 1998; Kirkhart, 1995; Mertens, 1999; Stanfield, 1999; Weiss, 1998), the major evaluation association in the United States is taking an active role in building diversity among the evaluation community. The Building Diversity Initiative Phase I proposal, to extend from June 2000 – June 2001, suggests a number of tasks upon completion of its first year, including a directory of minority evaluators, a survey of evaluation training programs, the development of a diversity-building plan, publication of guiding principles for evaluators to work across cultures, and other expected deliverables. Questions remain, however, over the sustainability of the Building Diversity Initiative project after its first year of Kellogg funding, the nature of Phase II activities, and articulated efforts to address the educational pipeline for minority evaluators.

Another potentially promising mechanism to address the participation and pipeline of minority professionals in educational evaluation includes the nascent development of a mentoring initiative within the Minority Issues in Evaluation Topical Interest Group (TIG) of AEA*. With AEA board support to provide three \$300 travel award scholarships for students to attend this year's annual conference, the executive committee of the Minority Issues in Evaluation TIG has begun steps to organize a mentoring activity pairing the travel award winners (i.e., protégés) with more experienced members of the TIG (i.e., mentors). Questions here concern the one-time nature of the travel award scholarships from the AEA Board, the support and/or endorsement of the Board for a coordinated, mentoring effort, and the exact nature and activities of the mentoring program. As Hood and Boyce (1997) point out from experience in developing mentoring activities within professional organizations, structured efforts to mentor future educational researchers must be multifarious, from awarding fellowships and other financial awards to developing organizational structure and policies to reflect a level of commitment in achieving a culturally diverse membership within the association and the profession.

Summary

Promoting the active participation of minority professionals in educational evaluation is an important goal for professional associations and the larger field of evaluation. Inherent in the efforts to build mechanisms to increase minority evaluators should be the concomitant realization of nurturing budding evaluators toward sustainable diversity and leadership within the ranks as well as promoting social justice and liberating evaluation frameworks.

* The Minority Issues TIG, estimated to be roughly 7% of the AEA membership, is one of a host of topical interest groups. Largely made up of evaluators of color, the TIG aims to address minority interests in evaluation. Currently, exact numbers of the racial and ethnic demographics of the AEA membership are unknown; a revised membership form to include information about members' ethnicity is expected to be presented before the Board — July, 2000.

References

- American Evaluation Association (2000). Proposal for the Initiative for Building Diversity Among the Evaluation Community, Phase I.
- Apple, M.W. (1996). *Cultural Politics and Education*. New York: Columbia University Teacher's College Press.
- Bordieu, P. and Passeron, J.C. (1977) *Reproduction in Education, Society, and Culture*. London: Sage.
- Chelimsky, E. (1998). The role of experience in formulating theories of evaluation practice. *American Journal of Evaluation*, 19, 35-56.
- Giroux, H.A. (1983). *Theory and Resistance in Education: A Pedagogy for the Opposition*. Sough Hadley, MA: Bergin and Garvey.
- Gordon, E.W. (1977) Diverse human populations and problems in educational program evaluation via achievement testing. p. 29-40. M.J. Wargo and D.R. Green (Eds.), *Achievement Testing of Disadvantaged and Minority Students for Educational Program Evaluation*. New York: McGraw-Hill.
- Gordon, E.W., Miller, F., and Rollock, D. (1990). Coping with communicentric bias in knowledge production in the social sciences. *Educational Researcher*, 19, 14-19.
- Hood, S. and Boyce, J. (1997). Refining and expanding the role of professional associations to increase the pool of faculty researchers of color through mentoring. *Diversity in Higher Education* 1, 141-159.
- Hopson, R.K. and Rogers S.J. (1999). Building a Diverse Evaluation Community of Sustainability through Mentorship: Towards Generational Success and Reproduction. Paper Presented at the American Evaluation Association Annual Conference: Orlando, FL.
- Hopson, R.K. Minority Issues in Evaluation Revisited: Re-Conceptualizing and Creating Opportunities for Institutional Change. *American Journal of Evaluation*, 20(3), 445-451, 1999.
- House, E. (1999). Race and policy. Educational Policy Analysis Archives. Available at <http://epaa.asu.edu/epaa/v7n16.html>.
- Kirkhart, K.E. (1995). Seeking multicultural validity: a postcard from the road. *Evaluation Practice* 16, 1-12.
- MacLeod, J. (1995). *Ain't No Makin' It; Aspirations and Attainment in a Low-Income Neighborhood*. Boulder, CO: Westview Press.
- Mertens, D.M. (1999). Inclusive evaluation: implications of transformative theory for evaluation. New York: *American Journal of Evaluation* 20(1), 1-14.
- Spring, J. (1997). *The American School: 1642 - 1996*, 4th ed. New York: McGraw-Hill.
- Stanfield, J.H. (1999). Slipping through the front door: relevant social sciences in the people of color century. *American Journal of Evaluation* 20(3), 415-431.

- Steele, C.M. (1992). Race and the schooling of black americans. *Atlantic Monthly*. Available at: <http://www.theatlantic.com/unbound/flashbks/blacked/steele.htm>.
- United States Department of Education (1999). *The Condition of Education*. Washington, DC: The National Center for Education Statistics. Available at <http://nces.ed.gov/pubs99/condition99/index.html>.
- Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation* 19,21-34.
- Yeakey, C.C. (1981). Schooling: a political analysis of the distribution of power and privilege. *Oxford Review of Education* 7(2), 173-191.
- Yeakey, C.C. & Bennett, C.T. Race, Schooling, and Class in American Society. *Journal of Negro Education*, 59, 3-18, 1990.

SESSION THREE: Evaluator Training Issues: Barriers, Recruitment and Outreach Strategies and Training Mechanisms

Session Chair

Floraline I. Stevens

Floraline Stevens and Associates

Presenters

Henry T. Frierson

Professor

School of Education

University of North Carolina, Chapel Hill

Sandra Fox

Retired and Adjunct Professor

University of New Mexico

Discussant

Costello L. Brown

Division Director

Educational System Reform

EHR/NSF

Guiding Questions

- How does the academic environment influence career choice and support or deter the entry and persistence of underrepresented minorities in the evaluation field? What are other barriers?
- Does this population have specific training needs? If so, how do we meet them? The discussion should include three levels of training — preparation to enter the field; professional development for current evaluators to transfer into the area of math and science program evaluation; expanding credentials of practicing evaluators who lack advanced degrees.

Discussion Highlights

Floraline I. Stevens

Prior to the presentations, the session chair reviewed salient points from Sessions One and Two that might lead to the final recommendations made by the group as a whole. Next, the Session Three presenters brought forward their suggested strategies to address the training barriers to increasing the number of underrepresented minority evaluators. Some of the strategies had emerged earlier in Sessions One and Two; however, the Session Three presentations further emphasized the importance of the issues and provided some new information.

Floraline Stevens, the chair of the session and a former Director of Research and Evaluation for the Los Angeles Unified School District, presented a summary of interviews of current underrepresented minority evaluators. Those interviewed emphasized that their presence was required on evaluation teams to make sure that all viewpoints were at least examined before opinions were formed and findings were determined.

Barriers to Becoming an Evaluator. The minority evaluators reported that there was a lack of professional development or classes provided by school districts to enhance the skills of non-traditional evaluators who may not have formal university education and training in evaluation. The cost of university classes and a lack of time to attend the classes were deterrents. To remedy some of these barriers, fellowships should be offered to encourage minorities to become evaluators and to ensure that they will get jobs once their training is completed.

Recruitment and Outreach. The minority evaluators reported that the evaluation profession is generally unknown because of lack of communication about it and lack of outreach activities by colleges and universities. Presently, evaluation is not touted to mainstream and minority university students as a desirable profession. The minority evaluators indicated that better career counseling at secondary and university levels and more outreach activities on the part of university evaluation departments were needed to give evaluation a higher profile as a profession.

Specific Training Mechanisms Needed. Skilled and talented minority evaluators can be developed with the proper framework for education and training. University course work should include statistics, research methods, evaluation theories, tests, measurement, etc. Practical evaluation training should include some knowledge and experience in teaching and learning as well as conducting evaluations under the supervision of a trainer/mentor.

The second presenter, Henry T. Frierson of the University of North Carolina at Chapel Hill, felt that there had not been a concerted effort to remedy the small numbers of underrepresented minority evaluators. He indicated that the January NSF Workshop where concern was voiced about the lack of underrepresented minorities in program evaluation served as a beginning.

Although there is a need to educate individuals in program evaluation, short-term workshops are insufficient for proper training and production of more evaluators. The focus must be on in-depth preparation of evaluators through formal education and training. NSF can make this possible by supporting participants in graduate programs and seriously seeking to enroll minority students. When funds are available and

minority students are targeted, most university graduate programs will be proactive in their recruitment activities to identify and enroll prospective minority graduate students.

Barriers. There is a general lack of awareness of what evaluation is, and the important role it can play in improving promising programs. To address this problem information should be disseminated about the field of evaluation, and underrepresented minority groups should be targeted for recruitment.

Recruitment. Increasing awareness of program evaluation by minority students should be one of the major efforts undertaken. Moreover, there must be opportunities to participate in evaluation projects. Since the need to give back or contribute to the community still runs strong in many minority individuals, when program evaluation is seen as improving programs, one would predict that the interest in program evaluation by minorities would expand tremendously. Evaluators without advanced degrees or non-traditional evaluators would benefit professionally from acquiring advanced degrees. They could serve in part-time and full-time paid positions at such agencies as NSF, NIH and other federal departments while they are completing or as part of their formal graduate studies.

Preparation. Quick-fix efforts in the form of short-term in-service training have not been successful. NSF should follow its model for increasing the production of PhDs in science, engineering, mathematics and technology from underrepresented minorities in the production of well-prepared program evaluators. Training grants can be awarded to doctoral programs that can train and prepare evaluators and such programs should ensure and demonstrate that they will enroll a sufficient number or proportion of minority students.

Sandra Fox from the University of New Mexico was the third presenter. She described her experience as an evaluator for the Bureau of Indian Affairs, Office of Indian Education Programs. The evaluation process for the 185 schools in the program was based upon the Effective Schools with its 10 correlates. The correlates were used to evaluate and monitor a school's efforts. One-fourth of the schools were visited annually to determine correlate implementation and to acquire information on attendance, achievement, and other specified outcomes. Monitoring and evaluation teams of four to five persons used a correlates implementation checklist so that the evaluation system was uniform. Prior to the team visits, schools used the correlates to conduct self-evaluations and provided this information to the team. This evaluation process lasted from 1990 to 1995. Major results from this process resulted in improved academic achievement, increased attendance, and increased enrollment in the system. Another outcome was the inclusion of Native Americans trained in the Effective Schools evaluation process.

Responses to the Issues

In the discussion that followed the presentations, the issues of barriers (e.g., money and time), recruitment and outreach, and mechanisms for training were found to be interrelated. For example, when the discussion centered principally on recruitment and outreach, the location of the sites selected for training and funding possibilities were also elements thought to be important to this issue. All of the presenters and the group agreed that little has been done to make graduate students — particularly

underrepresented minority graduate students — aware of program evaluation as a graduate specialization and as a profession. It was agreed that an evaluation training program funded by NSF should be established and should follow the model for increasing the number of underrepresented minority PhDs.

In-service projects have not been successful. In-service programs make their attendees more aware of program evaluation, but do not produce evaluators. When recruitment and outreach efforts are attached to the funding of student fellowships, most universities with the capabilities to provide quality education and evaluation training will make concerted efforts to recruit and have outreach activities. In addition, the group agreed that non-traditional evaluators who are underrepresented minorities can be the nucleus for forming a potential pool of individuals seeking advanced degrees in program evaluation. However, these evaluators would need financial assistance to increase their educational knowledge and skills at universities because they are usually older and have jobs. Fellowships should be offered at their current salary range to encourage this group to earn advanced degrees, while fellowships should be offered to traditional graduate students at the appropriate rate.

The group was particularly concerned that university training efforts be placed at sites where large pools of interested underrepresented minority students may be located. Dr. Sandra Fox, a Native American, further emphasized this point. She stated that if Native American graduate students were part of focused recruitment efforts, then at least one training site should be on a college or university campus for Native American students.

Recommendations from Session Three:

- NSF should fund training where the greatest pool of potential evaluators can be located.
- NSF should establish collaboratives between university training sites and consortiums of school districts with teachers and administrators interested in becoming evaluators. These persons have the essential knowledge of teaching and learning.
- NSF should fund internships or fellowships at a level of current salaries, so financial barriers are eliminated.
- Evaluation training must include traditional evaluation courses such as theories, methodology; non-traditional courses, such as thinking systematically, multicultural education, socio-cultural and linguistic cognition; practical experiences of conducting evaluations; and support from mentors.

Papers/Presentations

Reflections and Interviews: Information Collected about Training Minority Evaluators of Math and Science Projects

Floraline I. Stevens

Introduction

Equity and equal access are seen by educators and policymakers as bridging the gap between the haves and have-nots. Unfortunately, in many of our public schools the have-nots include a disproportionate number of underrepresented minority students. The issues of how best to address educational opportunity inclusion versus exclusion among students on many occasions bring about emotional debates. However, Federal, State, and local governments responded with numerous educational programs to enhance the educational opportunities for racial minorities, people with disabilities, and students from low socioeconomic status. In particular, the NSF, through its science and technology education programs responded to the continuing concern that the numbers of underrepresented minorities need to be increased in careers of science and technology. When evaluating these programs to determine their quality and impact, we must be careful to really hear the voices of underrepresented students, and not have these voices screened through channels that may not always be reflective of the students. We do this by having multi-ethnic evaluation teams. Through this process we help ensure that the voices and their nuances have a chance to be heard. Just as equity and equal access are necessary for the students, equity and equal access are necessary for the professionals — minority evaluators — that interact with these students.

Therefore, when deliberating about whether there is a need to train more minority evaluators of math and science projects, the issue of the need for multiple voices in evaluation became real. In response, I felt that instead of synthesizing previous research findings, two types of qualitative information should be gathered because this might prove to be useful and insightful.

The information came first through the reflections of an experienced minority evaluator of math and science projects; and second, from interviews of currently working public school system minority evaluators who are at different stages of evaluator development.

Data Sources

Information from Reflections was composed of my experiences in becoming an evaluator. These experiences included: receiving on-the-job and university training; networking with other evaluation professionals at meetings; and serving as the administrator responsible for securing in-service training for evaluators.

The interviews were conducted with minority staff members of the Los Angeles Unified School District's Program Evaluation and Research Branch (formerly Research and Evaluation Branch/Program Evaluation and Assessment). The intent was to gather their opinions of whether or not minority evaluators are needed to evaluate math and science projects in public schools. There was a professional dichotomy among the

members of the evaluation staff: some were educators (teachers and school administrators) with little/some technical knowledge while others had much technical knowledge but no school experience.

The information from these two types of data-gathering efforts provide some insight into such issues as diversity in evaluation teams, academic and other barriers, recruitment and outreach, and need for specific evaluation training.

Reflections of an Experienced Minority Evaluator of K-12 Math and Science Education Projects

In 1965, I became an “evaluator” when the ESEA Title I legislation was enacted and Federally funded education programs were being implemented in the Los Angeles Unified School District and other public school districts. One of the requirements for receiving the money from the Federal government was that the funded programs had to be evaluated. At that time, I was a demonstration training teacher in an elementary school. I had recently received my master’s degree in educational psychology in counseling from UCLA.

There were no evaluation types in the school district. However, in response to the Federal guidelines, the school district recruited a cadre of persons who had training in counseling because of their coursework in tests and measurements, statistics and research design. After a review of our resumes and an interview, we were selected to be ESEA Title I evaluators. We knew nothing about evaluation theories and evaluation procedures. In the beginning, we did evaluations by rote. We were given a format for data collection and writing the reports. Some of the early reports were awful because there was little or no intellectual decision-making when planning and conducting the evaluations. After three to four years of on-the-job experience and reading a lot, most of us became adequate in our attempts to provide quality information about the ESEA Title I projects. In this group of 10 evaluators, four were African Americans.

What helped us to survive was our extensive knowledge of teaching and learning in classrooms and, for some of us, our familiarity with the predominantly minority students who were in schools where the ESEA Title I programs were operating. We knew how to gain access to people and information in schools, a critical element in evaluation. We knew when the information provided did or did not make sense.

From those early experiences, I later became an evaluator of my first science education project, an ESEA Title III, K-12 ecology-biology-focused project. I used the evaluation project for my dissertation when I investigated the application of a specific evaluation theory. Since the early evaluation of the Ecology Program, I have evaluated many K-12 science education and science-focused professional development programs. Currently, I am the evaluator of two science education programs funded by NSF and NIH. Although I do not have a degree in science, mathematics or technology, I have taken many science courses. I found that extensive or “deep” science knowledge was not necessary to evaluate these K-12 science projects and programs effectively. What was most helpful was my graduate training at UCLA in research methods and evaluation and the on-the-job training.

Training Needed: University Coursework and Real Life Experiences

Statistics, research methods, evaluation theory, norm-referenced testing, criterion-referenced testing, learning theory, evaluation seminars, evaluation fieldwork and conducting research for the dissertation were required to receive a professional degree in research methods and evaluation at the doctoral level at UCLA. All of this information proved to be valuable to me when conducting project evaluations. In addition, the university training gave me the confidence to know that the evaluation procedures that I selected had educational and theoretical foundations. On-the-job training was also invaluable because through doing the actual evaluation work, receiving professional development on specific topics such as writing evaluation reports, and receiving critical remarks from evaluation experts improved the professional quality of my day-to-day evaluation work. It is the combination of these experiences that continues to guide me in my work as an evaluator.

An Administrator Responsible for the Training of Evaluators

When I became the Director of the Research and Evaluation Branch, I developed an ongoing program of professional development to assist the evaluation staff to become better qualified. The technical staffers, while having knowledge of research design, statistics and sometimes tests and measurement, were not always good writers. Educators needed reports to be clear and not burdened with technical and statistical terminology. In addition, technical staff had to learn about teaching and learn to relate their evaluation activities to findings that were meaningful and substantive. On the other side of the issue were the data collectors who were teachers and administrators with few technical skills. Classes were provided that taught evaluation and research design, statistics, and also report writing. For practical evaluation information, these persons were paired with evaluators such as Drs. Marvin Alkin, Winston Doby and Joan Herman from UCLA, and other evaluators from universities and research firms. In particular, Marvin Alkin spent a lot of time teaching qualitative evaluation procedures through what could be described as "clinical learning."

Interview Responses from Five Underrepresented Minority Evaluators

Interviews were conducted with five minority staff members of the Los Angeles Unified School District's Program Evaluation and Research Branch. The five interviewees were each asked the following questions:

- Have you ever evaluated a math, science or technology project?
- Is there a need to train and increase the number of minority evaluators of math and science projects? Why?
- Do you think an evaluator of math and science projects is different from other evaluators? Yes. No. Why?
- Do/did you have to meet special needs/requirements to become an evaluator? Yes. No. What?
- What professional development have you received to become a better or more efficient and effective evaluator?
- What barriers or obstacles have you encountered in receiving training to become an evaluator?

- What do you think are the requirements to receive an advanced college/university degree in evaluation?
- What barriers or obstacles have you encountered in seeking an advanced college/university degree in evaluation?
- What conditions would encourage you or other minority persons to seek more training to become an evaluator and/or seek an advanced degree in evaluation?

Minority Evaluators Feel There Is a Need for Training More Minority Evaluators

Four of the five minority evaluators (underrepresented in the fields of evaluation) who were interviewed were educators and one was not. None of the five evaluators had evaluated a math, science or technology project. They agreed that there was a need for minority evaluators of these types of projects. Selected explanations were:

"There is an absolute need for minority evaluators for every project. This is a multiethnic, multilinguistic society. Diversity of an evaluation team is paramount to paint a true picture for program staff. Unfortunately, most evaluation is from the white perspective. This is not meant to be interpreted as a negative comment. It is just that various cultures/lifestyles perceive the same scenario through different eyes, differing experiences. Having a diverse team ensures that all viewpoints are at least examined before opinions are formed."

"Yes absolutely. All segments of our society must be included in the data gathering and analysis of information pertaining to math, science and technology. This equalizes the division of labor and promotes knowledge and information parity throughout every strata of our country's ethnic and social groups."

"There is a serious shortage of minority evaluators all over the country. In my opinion, there are even fewer in the specialized fields of math and science. Last fall, LAUSD Program Evaluation Branch advertised nation-wide for evaluators and the results were sad relative to minorities. Out of more than 100 respondents, less than five percent were minorities (Hispanic and African American). There was a fair amount of Asian representation. The African Americans who responded and who were interviewed had very limited academic preparation in evaluation. Sadly to say, only whites and Asians made the list. This experience suggests that there is a serious need for African American evaluators. The training is needed."

Evaluators of Math and Science Projects Need Not Be Different

None of the evaluators felt that an evaluator of math and science projects was different from other evaluators. They indicated that a trained evaluator can evaluate any program. However, it helps to have knowledge of the subject area. Rather than content-focused, evaluators' orientation is toward being quantitative or qualitative methodologists and those who are able to successfully combine both perspectives. Precision was believed to be most important in any type of project evaluation.

Need to Meet Special Needs and Requirements to Become an Evaluator

The four certificated evaluators started out as data collectors who had easy access to schools, and not as program evaluators who could design an evaluation and plan evaluation activities. However, they worked with and were trained by experienced evaluators, attended in-service classes within the branch, attended conferences/seminars and took other professional development courses and continuing education units to help them to become more proficient. They indicated that they still needed such coursework as SPSS data analysis, statistics, report writing and technology/computer classes, and to attend evaluation conferences.

Barriers and Obstacles Encountered to Receive Training to Become an Evaluator

The evaluators stated that in the 1980's and early 1990's, professional development was provided but now there was no professional development support in the district to become a better evaluator. They reported that they must seek training external to the district. They said:

"Currently, there are no professional development or classes offered by the school district. To receive additional training would require finding the resources."

"Unfortunately, I have found very little 'formal' training available to those not in a terminal degree program (besides statistics classes). Most of my training has been self-initiated and self-sought. Heretofore, I had been very fortunate to have been granted most requests to attend training conferences which have been available locally."

"Currently, one must find one's own way, one's own training, professional development, etc. There is precious little time to do this adequately."

"Not enough money to take other courses and attend other conferences."

Barriers to College/University Degree in Evaluation

Lack of finances and time to attend a college/university were the barriers identified by most of the evaluators. One evaluator will be attending UCLA in Fall 2000 to earn a doctorate but not in evaluation. This evaluator will be using subsidized and unsubsidized loans along with grants to finance her graduate degree program.

Conditions That Would Encourage Minority Persons to Seek More Training to Become an Evaluator or Seek an Advanced Degree in Evaluation

According to the evaluators, evaluation as a profession is unknown because of lack of communication about it and lack of outreach activities by colleges and universities. The evaluators indicated:

"Better counseling in secondary and university levels. More outreach activities on the part of university evaluation departments. Higher profile of evaluation as a professional pursuit."

"I think that this is a relatively non-traditional field for many minority persons. It is not one that is touted by many institutions (at least not to the mainstream population) and there is not a lot of publicity given to it. Many of us 'find' this field, much like myself, and became self-taught. I am continually seeking training because I like this type of work, but there needs to be much more said about this profession to those entry-level undergraduate and graduate students who are at the onset of their careers."

"More awareness. Most minorities are not aware of the opportunities in evaluation. Also, more fellowships should be offered to encourage minorities to become evaluators."

"Higher salary, more responsibilities and knowing that one will get a JOB as a professional evaluator. Right now, this field is becoming less diverse."

Summary

Knowledge of teaching and learning is important when evaluating math and science projects. Some knowledge of K-12 math and science content is necessary but university-level training in research methods and evaluation is essential. On-the-job training when there are opportunities for "clinical" interactions with evaluation experts and attending inservice training are valuable but should be combined with university level training.

According to the five interviewees, minority evaluators are needed so that the viewpoints of various cultures and diversity will be at least examined before evaluative opinions are formed.

There is a dearth of African American and Hispanic evaluators to serve on evaluation teams.

Technical staff need procedural knowledge and experience about schools to become effective evaluators.

Educators who serve as evaluators indicated that they needed additional training in SPSS, technology/computers, report writing, statistics and data analysis. They did not mention the need to learn evaluation theories.

Lack of minority evaluators was attributed to multiple factors: lack of support to improve qualifications within school systems; lack of finances; and sometimes, inadequate time to seek external professional development in colleges and universities.

Evaluation as a profession is unknown to many undergraduate and graduate college students because of lack of communication about the profession and lack of outreach activities, particularly for minorities, by the colleges and universities.

Conclusion

The two qualitative information-gathering approaches proved to provide pertinent information about increasing the number of underrepresented minority evaluators to evaluate mathematics and science projects. Access to several opinions about the meaning of the data gathered from multi-ethnic students help to ensure that meaningful results will be reported. Skilled and talented minority evaluators can be developed with the proper framework for training — university course work and practical evaluation experiences, particularly those with skilled evaluation mentors. Aspiring non-traditional minority evaluators should be sought through outreach activities and provided with financial assistance for them to acquire university graduate degrees in evaluation. Although, multi-ethnic evaluation teams are desirable for a number of reasons, right now there is a dearth of these evaluators. The National Science Foundation should step forward to train minority evaluators of science and technology projects.

The Need for the Participation of Minority Professionals in Educational Evaluation

Henry T. Frierson

Introduction

If there were means to accurately determine the number of individuals from underrepresented minority groups who are professionally involved in program evaluation, the obvious hunch is that the numbers would be dismally small. Moreover, regarding those who have received formal graduate education and training in program evaluation, one can further assume that as for essentially every other academic discipline, the number of minorities is scarce among those formally trained in program evaluation. One can merely look at the numbers of underrepresented minorities who are receiving doctorates in the social sciences and research-based educational fields to quickly surmise that the number of individuals of color who might be formally trained in program evaluation is and will be low (see Sanderson and Dugoni, 1999). In the social sciences, the combined percentage of 1997 doctoral degrees received by African Americans, American Indians, Mexican Americans and Puerto Ricans was only 6.6 percent. In education, when the administrative degrees were removed, the percentage was just 7.2 percent. The social sciences and education are the general fields from which the majority of professional evaluators tend to come, and given the proportion of minority doctoral recipients, their underrepresentation in the field of evaluation will continue unless concerted efforts are made to address the situation.

If the democratization of program evaluation is to occur and if the field is to be truly pluralistic, then the field should apply assiduous efforts to ensure the involvement of more individuals who have been historically and traditionally underrepresented. Currently on the face of it, however, little concerted effort appears to have taken place. For example, one begs the question concerning the number of underrepresented minorities who have participated in the NSF/American Educational Research Association Evaluation Training Program (ETP). Although I have not seen the data, my hunch is that the number is significantly small. Importantly, however, it was noted in the Proceedings from the January NSF Workshop on Training Mathematics and Science Evaluators, that there is a concern about the lack of underrepresented minorities in program evaluation. If this holds true, it is a start.

Needed Effort

It is widely recognized that there is a clear need to educate and train individuals in program evaluation (see, for example, Altschuld et al., 1994, and Worthen, Sanders, and Fitzpatrick, 1997). Short-term workshops or institutes to expose individuals to evaluation principles and applications are insufficient as a measure to involve more minorities in evaluation. The same holds true for minority principal investigators (PIs) of programs supported by NSF and other funding agencies. Efforts to train PIs in evaluation through two- or three-day workshops are not going to result in the production of more evaluators. PIs have their own disciplines and with the added responsibility of running their programs are unlikely going to find the time to learn a new discipline such as program evaluation.

As there is a clear need to educate individuals in program evaluation, the focus should be on in-depth preparations, not brief exposures. Consequently, they should have extensive experience in the evaluation process through formal education and training. This would be most appropriately gained via graduate programs, and the participation of individuals from underrepresented minority groups should be a purposeful objective.

NSF can indeed play a considerable role in the formal education and training of increased numbers of minorities in evaluation. If NSF provides support for individuals to participate in graduate evaluation programs and if there is a clear aim to increase minority participation, many graduate programs will accordingly and seriously seek to enroll minority students. With opportunities to gain funding if minority students are targeted, most graduate programs will be proactive in their efforts to enroll minority students in programs where they can receive education and training in evaluation.

Barriers

A major task is to inform minority students of the opportunities and excitement that exists related to program evaluation and then recruit them into the graduate educational programs. Special training needs are not an issue for minority students. However, a problem that needs to be overcome is the lack of awareness of what evaluation is and the important role it can play in the continued existence and improvement of promising programs and certainly for those programs that can be shown to be successful. Actually, making students aware of the concept of evaluation is a critical task. Few graduate students can discern the difference between evaluation and research. Many will probably suggest that evaluation is just some or another form of applied research. Many will feel that if they have had some statistics or research methodology classes, they can naturally conduct evaluation projects. As most of us know, for years that was the thinking of many who made forays, so often unproductive, into evaluation. Evaluation has evolved into a discipline and should be treated as such. Just as graduate students specialize in educational statistics, qualitative research methodology, quantitative research methodology and testing and measurement, there can also be a specialization in program evaluation. Thus, the dissemination of information about the field of evaluation, graduate programs and career opportunities should occur while ensuring that individuals from underrepresented minority groups are targeted.

Recruitment

How can minority students be recruited into graduate evaluation programs? One cannot assume that there is a quick fix or that, just because there is a pronouncement that minority students are being sought, individuals of color will flock to those programs. Moreover and not surprisingly, a compounding problem is that few individuals from underrepresented minority groups are familiar with or have been exposed to the concept of pursuing a graduate program in evaluation. Increasing awareness would be

It will even be worth the effort to expose minority undergraduate students to evaluation and opportunities to participate in evaluation projects — just as there is an increasing opportunity to provide undergraduate students with more opportunities to participate in research projects. Of course, this exposure and experience should be gained in a setting where quality work is occurring to ensure meaningful learning

occurs. Such opportunities would lay the groundwork for a flow of individuals seeking graduate programs and advanced degrees related to evaluation. As the increase in research opportunities and programs for minority undergraduate students appear to be occurring, similar effects may accrue for programs providing exposure and real-life opportunities in evaluation for undergraduate minority students. A carrot to entice students into evaluation is the concept that a main goal of evaluation is program improvement. If that is the philosophy that is transmitted, rather than that evaluation is essentially about program judgment per se, one would predict that the interest in program evaluation by minorities would expand tremendously. The sense of giving back or contributing to the community still runs strong in a number of minority individuals. This is clear in the types of graduate programs that attract significant numbers of minority students.

On the other hand, there may well be a number of practicing evaluators from underrepresented minority groups that do not have advanced degrees. The January proceedings cited the difficulty in providing training programs for individuals who may be considered "non-traditional" because of their experiences and economic needs. Such individuals could certainly benefit professionally by having the desired credentials an advanced degree allows, but importantly, the more advanced formal training and education should add considerably to their knowledge and skills. Support for such individuals to participate effectively in graduate programs should be made available and the graduate programs should seek to accommodate and complement the individuals' background and experiences. It is important that the education program be meaningful for such individuals. Practicing evaluators could probably more readily serve in evaluation positions where they can apply their old and newly acquired skills and knowledge. These individuals might be great candidates to serve in temporary full-time and fully paid evaluation positions at agencies such as NSF, NIH, the Departments of Education and Defense, and GAO either while they are completing or as part of their graduate studies.

Preparation

As indicated in the proceedings from the January meeting, there appear to be relatively few formal graduate education programs in evaluation despite the need for professionally educated evaluation. Interestingly enough, Worthen, Sanders, and Fitzpatrick (1997) predict that graduate programs in evaluation are unlikely to expand. Ironically, although Worthen et al. made such a prediction, they noted that the need for more trained program evaluators exist exactly for the list of reasons they enumerated in their chapter on the future of evaluation. They listed, for example, an increase in the opportunity for careers in evaluation; an increasing institutionalization of evaluation; and the need for more trained evaluators. They further predicted that this need for more evaluators will result in quick-fix efforts to provide more in-service training of evaluators. An example is the National Institute of General Medical Sciences of the National Institutes of Health initially employing the American Physiological Society to train Minority Access to Research Careers and Minority Biology Research Support programs principal investigators and program directors to become evaluators. So far, this effort appears to be less than successful. Despite signs of increased demands for evaluation, Worthen et al. believe that there will be little growth in graduate evaluation programs because of the lack of understanding of its importance by university administrators. If Worthen et al.'s predictions are even partially accurate, then it is even

more important to redouble efforts to bring about the involvement of minority individuals in program evaluation because under normal circumstances, their access will continue to be limited.

An agency such as NSF can play a major role in serving as a catalyst to increase the number of minorities in graduate evaluation programs. An example is the role agencies such as NSF and the NIH have played in the production of science PhDs from underrepresented minority groups. Granted the numbers are less than impressive, but if one thinks that the situation is bad now, just imagine what the situation would be if NSF and the NIH did not provide support to encourage the enrollment and training of minority PhD students. The availability of support often galvanizes the call to action. Training grants can be awarded to doctoral programs that can train and prepare program evaluators. Selected programs should have to ensure and demonstrate that they will enroll a sufficient number or proportion of minority students. Additionally, minority graduate students with training in evaluation can be targeted for internships at NSF and other governmental funding agencies where they might assist and eventually serve as evaluators to projects funded by these agencies.

The funding of training grants should serve to increase the number of quality graduate programs. The University of California–Berkeley model, for example, appears to be quite impressive, although one might expand the evaluation activities beyond education, the focus of the Berkeley model. The idea of allowing students to do dissertations based on evaluation studies should prove quite beneficial for those who will continue on with careers in evaluation. This approach should allow students to strengthen their skills as evaluators, just as the research dissertation serves to strengthen research skills. The increase in minority students participating in quality graduate programs in evaluation will produce evaluators able to carry out well developed evaluation studies or conduct and manage high quality evaluation projects. A number may hold academic positions and teach program evaluation theory and methodology. Further, we can expect some of these individuals to become leaders in the field of evaluation.

Conclusion

Clearly, there is a need for highly trained evaluators. Evaluation is too important to be left in the hands of individuals with less than adequate preparation. Major problems arise regarding the small number of minorities pursuing graduate education and careers in evaluation and limited awareness of the field. As mentioned earlier, minorities do not have special training needs and it would be a folly to think so. What is really needed, however, is the opportunity to gain exposure to and experiences in evaluation. Moreover, informing individuals of the usefulness and value of evaluation in program development, in addition to program improvement would make the field even more attractive and gain greater attention from a broader population. Further, to tap practicing evaluators from underrepresented minority groups who lack advanced degrees, opportunities for those individuals should be made more readily available and educational programs should be structured to accommodate their backgrounds. Because of their nontraditional status and the economic constraints those individuals are more likely to face, adequate fellowships and other funding opportunities should exist for such individuals.

References

- Altschuld, J.W., Engle, M., Cullen, C., Kim., and Macce, B.R. (1994). The Directory of Evaluation Training Programs. *New Directions for Program Evaluation*, 62, 71-94.
- Sanderson, A. and Dugoni, B. (1999). Summary Report 1997: Doctorate Recipients from United States Universities. Chicago: National Opinion Research Center.
- Worthen, B.R., Sanders, J.R., and Fitzpatrick, J.L. (1997). *Program Evaluation: Alternative Approaches and Practical Guidelines*, 2nd Ed. White Plains, NY: Longman.
- National Science Foundation (2000). Proceedings from the NSF Workshop on Training.

An Effective School Evaluation and Training Program

Sandra Fox

In 1988 the Bureau of Indian Affairs, Office of Indian Education Programs, began a program of school improvement for its 185 schools. The Effective Schools improvement process was utilized and training and technical assistance for schools was provided on the implementation of ten correlates or characteristics of effective schools. These characteristics were based upon research on schools that were in high poverty areas but were achieving results. The 10 correlates were: 1) sense of mission, 2) monitoring and feedback of school and student progress, 3) challenging curriculum and appropriate instruction, 4) access to resources for teaching and learning, 5) high expectations for success, 6) safe and supportive environment, 7) strong instructional leadership, 8) home/school/community partnerships, 9) participatory management/shared governance, and 10) cultural relevance.

A second part of the process included a monitoring and evaluation system to determine the implementation of the 10 correlates of excellence and to gather and report process and outcome data. One-fourth of the 185 schools were visited annually to determine correlate implementation and to acquire data regarding attendance, achievement and other outcomes. A checklist was developed for each of the correlate areas to provide for uniform site visits. Extensive school reports were written outlining site visit findings.

The reports outlined schools' major strength and weakness areas in regard to each of the 10 correlates of excellence and indicated outcome data. Follow-up visits to schools were based upon findings outlined in the school reports.

The outcome data gathered were highlighted in the school reports and were aggregated for the whole system. Extensive databases were maintained. This data-gathering effort led to the school report cards required by the Improving America's Schools Act. Annual reports for the system were submitted to Congress utilizing aggregated process and outcome information gathered through the monitoring and evaluation effort.

Monitoring and evaluation teams were made up of four or five members, depending upon the size of the school. A cadre of team leaders for the monitoring and evaluation teams was created. The team leaders were Indian educators from outside the Bureau system who were recognized as distinguished educators. They were usually from colleges or universities, were private consultants or were superintendents or principals from public schools. Team members were administrators from public schools with large Indian populations or administrators of Bureau schools and education specialists from the Bureau's Central Office or the Line (district) Offices representing Title I and special education programs.

The team leaders made arrangements with the schools so that the school staff would be ready for the visitations. Using the checklists for each correlate area, the schools did self-evaluations prior to the teams' visits and this information was made available to the teams and provided starting points. The schools also made many reports and other documents available for the teams to review as documentation of correlate

implementation and data gathering. On-site, the teams first met with the administrators and school boards to explain the process before going separate ways to evaluate the implementation of the various correlate areas and to gather outcome data.

Team members were assigned specific correlate areas to investigate during the on-site visits. Each evening, during on-site visits that lasted at least three days, team members met to share information from their assigned correlate areas and for other correlates if they observed something pertinent. The team leaders gathered information from each of the team members to write extensive reports on the school's implementation of the correlates and outcome data. The reports were written after the teams left the school, but executive summaries of the reports were crafted by the teams and presented to the administrators and school boards before the teams left. Sometimes the school asked that the executive summaries be presented to as many staff members as possible in exit presentations. The final written reports were provided to the schools no longer than two weeks after the visitations.

Prior to each school year, a training session was provided for team leaders and team members. Training included in depth information on each correlate area with the latest research findings and recommendations, on outcome data gathering, and on all aspects of the monitoring and evaluation process including such things as following local protocol and filing for consultant fees and travel expenses. Team leaders and members from outside the Bureau system were paid consultant fees and travel expenses. Team members from inside the system were provided only travel expenses.

At the end of each school year, the team leaders and team members were brought together for a debriefing session. At that time, successful practices and problems with the process were discussed. Modifications were made, if necessary. The checklists were revised for the next year if it was found that items were not providing reliable or valid information. New items were added if it was found that certain variables were affecting the success of implementation of the correlates but were not addressed in the checklists.

This monitoring and evaluation process was started in 1990 and ended in 1995 when the Bureau of Indian Affairs' Central Office received a 50 percent budget cut from Congress.

Major results of the Bureau's Effective Schools project included increased academic achievement, increased attendance and increased enrollment in the system. Another outcome was the provision of many Indian educators trained in the Effective Schools process and the evaluation portion of it. Many of these individuals are still using their training to improve their individual schools or to provide training for others.

The number of team leaders involved was 12. Two hundred different team members were involved in the process over the five-year period. The monitoring and evaluation process was coordinated by a staff of four professionals and one clerical staff. This group secured and trained the team members, scheduled the on-site visitations, received the written reports and disseminated them to the schools, maintained data bases and summarized the process and outcome data for annual reports for the system and to Congress.

At present, the need for many more Indian evaluators is great. The Department of Education, as a result of President Clinton's executive order of 1998, has made research in Indian education a priority. The National Science Foundation needs more Indian

evaluators to assist in that evaluation process. The Bureau of Indian Affairs wants to resurrect the Effective Schools monitoring and evaluation process. Many of the evaluators who were involved in the Effective Schools process from 1990 until 1995 have moved on to positions that make them less available to serve on evaluation teams, although a handful of them are still actively utilizing the process to evaluate schools. The process has influenced their work, wherever they are, however.

Most Indian evaluators are not trained in evaluation but have learned by being involved in some evaluation process. The need for formal training is critical, however, as schools serving Indian students participate in the results-oriented school reform process and the outcome data and sound evaluation practices become more and more important. As the smallest minority, statistics on Indian education are often not reported even by the National Center for Education Statistics. It will be up to Indians to ensure that evaluation of Indian education is carried out and that it is done in the most fair and effective way.

Workshop Recommendations

At the close of Session Three, the Session Chair asked participants to reflect upon two days of discussion and provide guidance to NSF. Participants identified broad themes, outreach mechanisms and products for further study. The following is a list of recommended actions.

1. NSF's User Friendly Handbook for Project Evaluation should be updated to encompass and respond to the points specified below:
 - Cultural awareness of the environment from which the participants are drawn must be emphasized.
 - Test results must be reported with context data.
 - Disaggregation of program data should include, as appropriate, factors such as, but not limited to, race, gender, socioeconomic status and opportunity to learn.
 - The level of implementation of a program/intervention must be coupled with achievement results.
 - Evaluations must recognize that the culture of students influences how they respond to the assessment process and assessment items.
 - Consideration must be given to the impact of teachers' attitudes, beliefs and behaviors on student achievement because low-income and minority students are more teacher dependent.
2. Non-minority evaluators should be trained to evaluate programs that target minority students.
3. NSF should fund training where the greatest pool of potential minority evaluators can be located and trained.
4. Training of minority evaluators should be conducted by a team that includes minority education evaluation faculty/trainers.
5. NSF should fund evaluator internships to eliminate barriers to advanced training:
 - Regular students should be funded at the prevailing wage level for internships.
 - Non-traditional students should be funded at the level of their current salary.
6. Evaluation training must include traditional evaluation courses such as: methods and statistics; non-traditional courses such as multi-cultural, socio-cultural and linguistic cognitive factors; a holistic and systemic approach that considers both internal and external factors that influence the process and focuses on components and the relationship between them; practical experiences of conducting evaluation; and support from mentors.

7. NSF should fund the identification and development of a database of practicing minority evaluators, to be added into an existing database.
8. NSF should establish collaboratives between university evaluation training sites and consortiums of school districts having administrators and staff interested in becoming evaluators.
9. NSF should fund a research study that captures from minority evaluators those experiences that led them to become evaluators.
10. NSF evaluation training program sites should have a critical mass of trainees so that support mechanisms can be planned and implemented.
11. NSF evaluation training should be provided at locations conducive to reaching Native Americans, perhaps at Haskell Indian Nations University in Kansas.
12. NSF should fund evaluations/studies of successful programs that are encouraging involvement of minorities in math and science, such as MESA.

CLOSING REMARKS

Elmima C. Johnson

Staff Associate

EHR/NSF

This two-day meeting has raised more questions than it answered. However, it did clarify the questions to be asked and sharpen our focus. As a result, I can list several potential activities that would respond to the discussion and recommendations emanating from the Workshop.

Directorate actions will begin with the compilation of the proceedings of the workshop. This document, which will include the papers and presentations as well as summaries of the discussions and recommendations, will serve as a reference and blueprint for designing a strategy to respond to participant concerns regarding capacity-building and NSF's role in evaluator training and education. To date, EHR capacity-building activities have been primarily ad hoc and none have focused specifically on the issue surrounding culturally relevant evaluation. These have been relatively small efforts suggested by the field, and informal feedback suggests they were successful. We recognize that a specific rationale and framework should guide future efforts. Additionally, any proposed efforts should be defined as part of a more comprehensive approach.

Before we can develop the framework we need additional information in several areas. First, we need a comprehensive picture of NSF-supported training opportunities. To this end we are supporting a project to provide a detailed description of our funded efforts and their success. We are also exploring other formal evaluation training approaches and related efforts that can be adapted for our needs. This is being accomplished through a broad-based literature review to identify other evaluation models and potential models, i.e., from disciplines other than mathematics and science education.

Second, we need to determine the demographics of the current population of "evaluators," including minority evaluators. Two manpower surveys are under consideration. One would survey graduates of formal evaluation training programs, i.e., university-based efforts. The second survey would target practicing evaluators without formal training in evaluation methodology. The first task will be to define the two populations and there probably will be some overlap. Both would include a special effort to identify minority evaluators.

At some point we will need feedback from our stakeholders. We will utilize the recommendations offered by workshop participants regarding evaluator training and updating NSF evaluation publications to include a focus on the multicultural context of evaluation. We also plan to talk with representatives of national evaluation associations. For example, we could hold discussions with officials of the American Evaluation Association (AEA) regarding their efforts to diversify the pool of evaluation professions. Another issue to consider is how to attract more of the minority social science graduates into the field of evaluation.

These steps, when completed, should prepare us to set goals and priorities and define the parameters of future NSF efforts. The framework developed must be flexible enough to accommodate new practices and new Federal mandates regarding accountability stemming from the Government Performance and Results Act (GPRA). The question to be answered is what does the directorate want to accomplish in addressing the need for culturally relevant evaluation within the GPRA and diversity contexts.

APPENDIX A

Workshop Agenda

June 1

- | | |
|-------------|---|
| 8:30-9:00 | Continental Breakfast |
| 9:00-9:10 | Welcome – Dr. Judith Sunley, Interim Assistant Director, Education and Human Resources, EHR/NSF |
| 9:10-9:20 | Greetings – Dr. Eric Hamilton, Interim Division Director, Research, Evaluation and Communication, EHR/NSF |
| 9:20-9:35 | Introduction to the Workshop – Dr. Elmima Johnson, Staff Associate, Division of Research, Evaluation and Communication, EHR/NSF |
| 9:35 -9:45 | Remarks – Dr. Conrad Katzenmeyer, Senior Program Director, Division of Research, Evaluation and Communication, EHR/NSF |
| 9:45 -10:30 | <p>Session 1: Evaluation of Educational Achievement of Underrepresented Minorities. Group discussion: Chair, Dr. Beatriz Clewell; Presenters – Dr. Carlos Rodríguez; Dr. Gerunda Hughes; Discussant – Dr. Jane Butler Kahle</p> <ul style="list-style-type: none"> • Several Federal agencies offer support for science and mathematics education reform, i.e., NSF, NASA, DOE and ED. Evaluation of these efforts includes attention to student academic achievement. What are the issues surrounding the evaluation of science/ mathematics achievement, especially the academic assessment of underrepresented populations? The discussion should highlight the cultural context of this area of evaluation with reference to relevant literature. |
| 10:30-10:45 | Break |
| 10:45-11:30 | Continue discussion |
| 11:30-12:00 | Emergent Issues—Chair and Discussant |
| 12:00-1:30 | Lunch |

1:30-3:45

Session 2: Participation of Minority Professionals in Educational Evaluation. Group discussion: Chair, Dr. Stafford Hood; Presenters – Dr. Stafford Hood, Dr. Rodney Hopson; Discussant – Dr. Norma Dávila

- What is the motivation for increasing the number of minority evaluators with advanced training and experience in the field of educational evaluation? (Minority evaluators are defined as persons from those groups underrepresented in the field of evaluation.) What is the importance of including minority evaluators in the evaluation of science and mathematics education programs?
- What mechanisms are available to identify the current population of minority evaluators, in particular those with expertise/experience in science/mathematics education? (Information sources include a survey of professional organizations, university programs, etc.)

3:45-4:00

Break

4:00-4:30

Emergent Issues—Chair and discussant

4:30-6:00

Reception at hotel

June 2

8:30-9:00

Continental Breakfast

9:00-11:30

Session 3: Training Issues Group discussion: Chair, Dr. Floraline Stevens; Presenters – Dr. Floraline Stevens, Dr. Henry Frierson, and Dr. Sandra Fox; Discussant – Dr. Costello L. Brown

- How does the academic environment influence career choice and support or deter the entry and persistence of underrepresented minorities in the evaluation field? What are other barriers?
- Does this population have specific training needs? If so, how do we meet them? The discussion should include three levels of training — preparation to enter the field; professional development for current evaluators to transfer into the area of math and science program evaluation; expanding credentials of practicing evaluators who lack advanced degrees.

11:30-12:00

Emergent Issues—Chair and Discussant

12:00-12:30

Break

12:30-1:30

Lunch

1:30-3:00

Next Steps – Dr. Elmima Johnson

APPENDIX B**Invited Participants****Dr. Beatriz C. Clewell**

Principal Research Associate and Director
 Evaluation Studies and Equity Research
 Program
 Education Policy Center
 The Urban Institute
 2100 M Street NW, Suite 500
 Washington, DC 20037
 202-261-5617 – phone
 202-823-2477 – fax
 tclewell@ui.urban.org
 And *former* Executive Director
 Commission on the Advancement of
 Women and Minorities in Science,
 Engineering and Technology (CAWMSET)
 National Science Foundation

Dr. Norma Dávila

Co-PI, Puerto Rico Statewide Systemic
 Initiative
 University of Puerto Rico
 Box 23334
 San Juan, PR 00931-3334
Delivery Address
 Resource Center for Science and
 Engineering
 Facundo Bueso Building, Office 304
 Rio Piedras, PR 00931
 787-765-5170 – phone
 787-756-7717 – fax
 n_davila@upr1.upr.clu.edu

Dr. Sandra Fox

Retired and Adjunct Professor
 University of New Mexico
 4200 Spanish Broom Avenue NW
 Albuquerque, NM 87120-2589
 505-890-5316 (phone and fax)

Dr. Henry Frierson

Professor, School of Education
 University of North Carolina, Chapel Hill
 121 Peabody Hall
 CB #3500
 Chapel Hill, NC 27599
 919-962-7507 – phone
 919-962-1533 – fax
 ht_frierson@unc.edu

Dr. Anthony L. Hill

Senior Evaluator – General Government
 Division
 Administration of Justice Issues
 U.S. General Accounting Office
 441 G Street NW, Rm. 2B-41
 Washington, DC 20540
 202-512-9604 – phone
 202-512-4516 – fax
 hilla.ggd@gao.gov

Dr. Stafford Hood

Associate Professor
 Division of Psychology in Education
 Arizona State University
 Tempe, AZ 85287-0611
 480-965-6556 – phone
 480-965-0300 – fax
 stafford.hood@asu.edu

Dr. Rodney K. Hopson

Assistant Professor
 School of Education
 Center for Interpretive and Qualitative
 Research
 Duquesne University – Canevin Hall
 Pittsburgh, PA 15282
 412-396-4034 – phone
 412-396-1681 – fax
 hopson@duq.edu

Dr. Gerunda B. Hughes

Howard University, School of Education
 CRESPAR
 2441 4th Street, N.W.
 Washington, DC 20059
 202-806-7343 – phone
 ghughes@howard.edu

Dr. Francine Jefferson

Telecommunications Policy Analyst
 National Telecommunications and
 Information Administration (NTIA)
 U.S. Department of Commerce
 1400 Constitution Avenue NW, Room
 4096
 Washington, DC 20230
 202-482-5560 – phone
 202-501-8009 – fax
 fjefferson@ntia.doc.gov

Dr. Vinetta C. Jones

Dean, School of Education
 Howard University
 2421 4th Street NW
 ASA Building
 Washington, DC 20059
 202-806-7340 – phone
 202-806-5302 – fax
 v_jones@howard.edu

Dr. James M. Patton

Professor, School of Education
 The College of William and Mary
 P.O. Box 8795
 Williamsburg, VA 23187
 757-221-2318 – phone
 757-221- 2988 – fax
 jmpatt@wm.edu

Dr. Carlos Rodríguez

Principal Research Scientist
 American Institutes for Research
 1000 Thomas Jefferson Street NW
 Suite 400
 Washington, DC 20007
 202-944-5240 – phone
 202-944-5454 – fax
 crodriguez@air.org

Dr. Thelma L. Spencer

Educational Consultant
 109 Linden Hall Lane
 Gaithersburg, MD 20877-3461
 301-330-3493 – phone

Dr. Floraline I. Stevens

Floraline Stevens and Associates
 707 S. Orange Grove Blvd, Unit H
 Pasadena, CA 91105
 626-403-5752 – phone
 626-403-8318 – fax
 stevensfi@webtv.net

Dr. Sheila D. Thompson

Center for Research on the Education
 of Students Placed at Risk (CRESPAR)
 Howard University
 2900 Van Ness Street NW
 Washington, DC 20008
 202-806-8484 – phone
 202-806-8498 – fax
 sdt@crespar.law.howard.edu

NSF Staff**Dr. Costello L. Brown**

Division Director
Educational System Reform
EHR/NSF
Room 875
4201 Wilson Blvd
Arlington, VA 22230
703-292-8690 – phone
703-292-9047 – fax
clbrown@nsf.gov

Dr. Eric R. Hamilton

Interim Division Director
Division of Research, Evaluation and
Communication
EHR/NSF
Room 855
4201 Wilson Blvd
Arlington, VA 22230
703-292-8650 – phone
703-292-9046 – fax
ehamilto@nsf.gov

Dr. Elmima C. Johnson

Staff Associate
Division of Research, Evaluation and
Communication
EHR/NSF
Room 855
4201 Wilson Blvd
Arlington, VA 22230
703-292-5137 – phone
703-292-9046 – fax
ejohnson@nsf.gov

Dr. Jane Butler Kahle

Division Director
Elementary, Secondary and Informal
Education
EHR/NSF
Room 885
4201 Wilson Blvd
Arlington, VA 22230
703-292-8628 – phone
703-292-9044 – fax
jkahle@nsf.gov

Dr. Conrad G. Katzenmeyer

Senior Program Director
Division of Research, Evaluation and
Communication
EHR/NSF
Room 855
4201 Wilson Blvd
Arlington, VA 22230
703-292-5150 – phone
703-292-9046 – fax
ckatzenm@nsf.gov

Dr. Judith S. Sunley

Interim Assistant Director
Directorate for Education and Human
Resources
National Science Foundation
Room 805
4201 Wilson Blvd
Arlington, VA 22230
703-292-8600 – phone
703-292-9179 – fax
jsunley@nsf.gov

APPENDIX C**Biographies of
Invited Participants****BEATRIZ CHU CLEWELL, PhD**
The Urban Institute

Dr. Clewell served as the Executive Director of the Commission on the Advancement of Women and Minorities in Science, Engineering, and Technology Development at the National Science Foundation. Her appointment ran until June 2000, at which time she returned full-time to the Urban Institute.

Before joining the Urban Institute in 1994, she was a Senior Research Scientist at the Educational Testing Service (ETS) in Princeton, NJ where she was employed for 13 years in the Educational Policy Research Division. She has directed over thirty research studies, many of them studies of factors affecting access of underrepresented groups to high quality mathematics and science education. In the late 1980's, she directed a study of 168 intervention programs for middle school minority and female students; this study, funded by the Ford Foundation, later became a book, *Breaking the Barriers: Helping Female and Minority Students Succeed in Mathematics and Science*. More recently, she completed a research project funded by the Sloan Foundation to identify factors that influence the non-SEM career choices of high-ability African American and Latino undergraduates on the math/science career pathway.

Dr. Clewell was PI for the evaluation of NSF's Program for Women and Girls. She has had almost 20 years of experience as a program evaluator and has directed a number of other large-scale evaluations, among them the Evaluation of the Pathways to Teaching Careers Program, currently underway, which involves 41 institutions, and the Ford Foundation Minority Teacher Education Program Evaluation, which involved 50 institutions. In the mid-1970's she was a middle school teacher for two years and credits this experience with inspiring her life-long interest in education and equity.

In addition to her research, she is affiliated with many professional organizations, including serving as an associate editor for the Educational Evaluation and Policy Analysis journal and as a member of the National Science Foundation's Committee on Equal Opportunities in Science Engineering (CEOSE).

NORMA DÁVILA, PhD
University of Puerto Rico

Education

- 1991 PhD in Psychology, Committee on Human Development, University of Chicago, Chicago, Illinois.
- 1988 MA in Behavioral Sciences, Committee on Human Development, University of Chicago, Chicago, Illinois.
- 1985 BA in Psychology, Yale University, New Haven, Connecticut.

Research Experience

- 1998 – Present Project Director, Puerto Rico/New York City Educational Linkages Demonstration Project, sponsored by the US Department of Education.
- 1997 – Present Co-Principal Investigator, Puerto Rico Statewide Systemic Initiative, sponsored by the Puerto Rico Department of Education, the Resource Center for Science and Engineering, and the National Science Foundation.
- 1996 – Present Evaluation and Assessment Coordinator, Puerto Rico Statewide Systemic Initiative, sponsored by the Puerto Rico Department of Education, the Resource Center for Science and Engineering, and the National Science Foundation.
- 1993 – 1995 Evaluation Coordinator, Puerto Rico Statewide Systemic Initiative, sponsored by the Puerto Rico Department of Education, the Resource Center for Science and Engineering, the General Council on Education, and the National Science Foundation.
- 1992 – 1993 Evaluation Coordinator, Puerto Rico Scope, Sequence, and Coordination Project, sponsored by the National Science Teachers' Association at the Resource Center for Science and Engineering, Río Piedras, Puerto Rico.

Publications

- In Press Century, J. R., Clune, B., Dávila, N., Heck, D., Osthoff, E., and Webb, N. Evaluating Systemic Reforms and Initiatives. New York, NY: Teachers' College Press.
- 1999 Dávila, N. Assessing Student Outcomes. In N. L. Webb (Ed.), Evaluation of Systemic Reform in Mathematics and Science: Synthesis and Proceedings of the Fourth Annual NISE Forum. Workshop Report No. 8. Madison, Wisconsin: National Institute for Science Education.

Dávila, N. Charting the Future of Assessment in Systemic Educational Reform: Teacher and School Principal Involvement in Evaluation and Assessment Use. ERIC Document TM030314.
 Dávila, N. Measuring and Documenting Outcomes: Going Beyond Tradition in Program Evaluation. ERIC Document 432 585.

1996

Dávila, N., Gómez, M. and Vega I. Y. Evaluating the Transformation of the Teaching/Learning Culture of Schools Involved in Systemic Science and Mathematics Educational Reform. ERIC Document 395 803.

1995

Dávila, N. and Gómez, M. Evaluation of School-Based Regional Dissemination Centers as Scale-Up Mechanisms for Systemic Educational Reform in Science and Mathematics. ERIC Document 310 702.

1994

Dávila, N. and Gómez, M. Assessment of the Impact of a New Curriculum on Systemic Change. ERIC Document 390 907.

Teaching Experience

1997 – Present

Associate Professor, Department of Psychology, University of Puerto Rico, Río Piedras.

1991 – 1997

Assistant Professor, Department of Psychology, University of Puerto Rico, Río Piedras.

SANDRA J. FOX, PhD
University of New Mexico

Sandra J. Fox is a member of the Oglala Lakota Nation of the Pine Ridge Indian Reservation in South Dakota. She grew up, however, on the Fort Berthold Indian Reservation in North Dakota where her mother was a teacher in a Bureau of Indian Affairs (BIA) school. She attended BIA schools from grade one through twelve. Upon high school graduation, she attended Dickinson State College and received a bachelor's degree in English education. Dr. Fox and her husband taught for two years in a public school in North Dakota before joining the Bureau of Indian Affairs school system, starting at Cheyenne-Eagle Butte High School in Eagle Butte, SD, on the Cheyenne River Indian Reservation where they taught for three years. At that time an Indian administrators' program was started at Pennsylvania State University, and she and her husband attended that program receiving master's and doctoral degrees.

Dr. Fox became an education specialist for the Aberdeen Area Office, Bureau of Indian Affairs, in Title I and language arts. From there, she and her husband transferred to the Washington, DC office of the Bureau of Indian Affairs where she was in charge of the Eisenhower math and science program for Bureau schools and developed and coordinated the Effective Schools monitoring and evaluation process. She is retired now after working for the Bureau for 24 years. Her last assignment for the Bureau was School Reform Team Leader. She was named 1998 Indian Educator of the Year.

She presented a paper on school reform and American Indian education at a national Indian education research conference in Albuquerque on May 31. A recent article is on the use of performance-based assessment with Indian students.

HENRY T. FRIERSON, PhD

University of North Carolina, Chapel Hill

Henry T. Frierson received his B.S. in psychology and master's in educational psychology from Wayne State University, and he received his Ph.D. in educational psychology from Michigan State University in 1974. Currently, he is a Professor of Educational Psychology, Measurement, and Evaluation at the University of North Carolina at Chapel Hill. From 1988 to 1996, he was associated with the University's Graduate School and served as the Associate Dean for six years. From 1974 to 1993 he was a faculty member in the UNC-CH School of Medicine. At the School of Medicine he was the founder and director of the Learning and Assessment Laboratory, an academic support unit for Medical School and other UNC-CH students. During his tenure at the Graduate School he was successful in obtaining considerable funding for graduate student support and special research programs. He has continued those efforts as a full-time professor in the School of Education. In the School of Education, he teaches program evaluation and research methods courses, and an additional course titled, The Psychology of Adult Learning. He also directs a major research support program, the Research Education Support Program, that is largely funded by NIH and NSF grants. The program provides support for minority undergraduate students to become involved in quality research, graduate students to complete their research for their PhD degrees, medical and dental students to have extensive research experiences, and for undergraduate students from other colleges and universities to have full-time summer research experiences at the University of North Carolina at Chapel Hill. His current interests rest in program evaluation and in increasing the number of individuals of color in doctoral programs and research careers.

ANTHONY L. HILL, PhD
U.S. General Accounting Office

Dr. Hill is a Senior Evaluator at the U.S. General Accounting Office (GAO) where he has led several major domestic and international program evaluations and audits. Currently, he is engaged in a GAO-wide effort to assess how well Federal agencies have complied with the Government Performance and Results Act in developing and adhering to management performance plans. Over the last two years, he has served as the principal instructional design specialist for developing program evaluation and performance management training curricula for GAO's professional evaluation and audit staff. With extensive adult training experience in government, private sector consulting, and academia, Dr. Hill has been instrumental in the agency's efforts to identify the core competencies of its professional management and evaluation staff, and to design and implement learning strategies to enhance these competencies. He is highly skilled in instructional design and delivery of adult education and technical assistance to internal and external clients. For example, he has extensive experience teaching graduate level courses in social research, psychological and educational testing, and cross-cultural counseling. Further, he has provided technical assistance to GAO evaluation teams in designing and conducting audits and evaluations of government programs.

Over the last 19 years, he has held a variety of technical and project management positions at GAO that have required extensive expertise in program evaluation and organizational development. He has consulted with congressional staff, U.S. and international government and private sector officials. His duties have included extensive international travel, including several visits to the republics of the former Soviet Union, North and West Africa, and Western Europe.

He has over 25 years professional experience in psychotherapy and educational and psychological testing. He is licensed as a psychologist and as a professional counselor. In 1996, he was appointed by Governor Glendening to the Maryland State Board of Examiners of Professional Counselors and is the former president of the Maryland Association of Measurement and Evaluation.

STAFFORD HOOD, PhD
Arizona State University

Academic Background

PhD in Education (emphases in administration, program evaluation, and policy analysis), University of Illinois at Urbana-Champaign (1984)

Selected Professional Experiences

Arizona State University (1992 to Present)

Associate Professor of Psychology in Education (tenured) where he teaches graduate courses in Counseling/Counseling Psychology (e.g. Psychological Testing, Program Evaluation, and Multicultural Counseling). He is also the co-director of the annual national conference, Relevance of Assessment and Culture in Evaluation hosted by Arizona State University.

Northern Illinois University (1988-1991)

Assistant Professor of Educational Psychology where he taught undergraduate and graduate courses (e.g. Measurement and Evaluation in Teaching, Standardized Testing, Test Construction, and Program Evaluation).

Acting Assistant Dean – Graduate School (1989-1990)

Illinois State Board of Education (1984-1988)

Assessment Specialist – Student Assessment Section (1987-1988) with responsibilities for the technical activities associated with the development and validation of the Illinois Student Assessment Program (standardized tests across five subject areas administered to third, sixth, eighth, and eleventh grade students in Illinois Public Schools) and Illinois Certification Testing System (tests across 54 areas to certify teachers, administrators, and other educational personnel in Illinois Public Schools). Designed and monitored the implementation of the bias review components for these two programs.

Program Evaluator III – Program Evaluation and Assessment Section (1984-1987) with responsibilities for conducting evaluations of state and Federally funded special education programs and special projects.

Selected Evaluation and Training Consulting Activities

Arizona Supreme Court, Foster Care Review Boards. Co-Director. Responsible for designing qualitative component of project and analyses of qualitative and quantitative data for the Foster Care Review Board's Annual Report. September 1998 to December 1998 and June 1999 to December 1999.

Chicago State University, College of Education. Project Director. Responsible for the design, implementation, and report on the evaluation of the Field Based Teacher Preparation Program. July 1997 to present.

National Center for Urban Partnerships, Center for Educational Evaluation. Responsible for serving as an evaluation facilitator to three partnerships of local school districts, universities, and community colleges in three cities (Seattle WA, Denver CO, and Memphis TN) as they implement program evaluation activities related to increasing the attainment of baccalaureate degrees for at risk students. August 1994 to July 1997.

RODNEY K. HOPSON, PhD
Duquesne University

Rodney K. Hopson is an assistant professor in the School of Education, Department of Foundations and Leadership and faculty member in the Center for Interpretive and Qualitative Research (CIQR), Duquesne University. In 1997, upon completion of his dissertation in the Department of Foundations, Leadership, and Policy in the Curry School of Education, University of Virginia, he completed a year as a Postdoctoral Research Fellow in the Department of Social and Behavioral Sciences at the School of Hygiene and Public Health, Johns Hopkins University.

His research interests include social politics and policies, foundations of education, sociolinguistics, and ethnographic evaluation research. Forthcoming publications include guest editor of a special issue of the American Evaluation Association journal, *New Directions for Evaluation*, that will address how language shapes the evaluation of social programs and policies, senior author of a paper addressing HIV/AIDS ethnographic intervention in Baltimore City, MD, and author of a paper analyzing transformation of higher education in the post-apartheid context of the Republic of Namibia, (where he will continue researching and lecturing during the 2001 calendar year as a Fulbright Scholar on issues pertaining to educational language politics and policies).

He is currently working on two projects in southwestern Pennsylvania exploring and advancing the notion of educational resilience among high and low performing schools with colleagues at the University of Pittsburgh, the Educational Policy and Issues Center, and Duquesne University. One of the projects, funded by the Heinz Endowment, is investigating factors that contribute to understanding exemplary elementary schools for replication of sustainable, educational policies and practices in the region.

At Duquesne University, Dr. Hopson teaches courses in the Department of Foundations and Leadership and is a part-time faculty member in a new master's program in Program Evaluation and Planning in the same department. Select courses include: *Society, Politics, and the Teaching Profession*; *Philosophical, Historical, Sociological Foundations of Education, Society and the Individual*; and *Educational Language Politics and Policies*.

During the 1999-2000 academic year, Dr. Hopson was named as Fellow to the Pennsylvania Education Policy Fellowship Program where he is part of a cohort that is influencing children's and educational policy in the commonwealth. He recently participated in the 1999 Teaching with Technology Summer Institute and received a Duquesne University Presidential Scholarship awarded to faculty for demonstration of research promise. Select service accomplishments include: charter member of the Western Pennsylvania Evaluation Network, local affiliate of the American Evaluation Association, member of the Duquesne University Charter Schools Project Advisory Board, and chair of the academic curriculum and development committee at Ethnan Temple SDA Christian Elementary School.

Dr. Hopson is married to Wabei Siyolwe and they live in Pittsburgh, Pennsylvania with their two children, Hannibal and Habiba.

GERUNDA B. HUGHES, PhD
Howard University

Gerunda B. Hughes is Assistant Professor of Curriculum and Instruction in the School of Education where she teaches mathematics methods courses for elementary and secondary pre-service teachers. Dr. Hughes serves as a Co-Principal Investigator of the "Assessment and Evaluation Innovations Project" at the Center for Research on the Education of Students Placed At Risk (CRESPAR) which is funded by the Office of Educational Research and Improvement (OERI). She received her BS in Mathematics from the University of Rhode Island; MA in Mathematics from the University of Maryland, College Park; and PhD in Educational Psychology from Howard University.

At CRESPAR, Dr. Hughes collaborates with other researchers on projects in which she plans and conducts basic research and development activities with the aim of aligning classroom instructional practices with assessment practices that maximally develop students' skills, abilities and talents. She also serves as a Guest Co-Editor of the Journal of Negro Education.

Prior to joining the faculty in the School of Education in 1995, Dr. Hughes taught mathematics courses in the Department of Mathematics and developmental mathematics courses to underprepared college students in the Center for Academic Reinforcement (CAR) for twenty years.

Dr. Hughes has served as Project Coordinator of the WBHR Alliance for Minority Participation Teacher Preparation (AMP-TP) Program (1996-98) funded by the National Science Foundation (NSF); research/evaluation consultant for the Calculus Reform Project at Howard University; Co-Principal Investigator of a NSF funded project to develop and evaluate the use of performance assessments in college pre-calculus; a member of a reverse site visit panel for the Collaboratives for Excellence in Teacher Preparation (CETP) program; and as Principal Investigator of small-scaled Howard University -funded projects entitled, "Transforming Professors into Teachers," "Effective Teacher Preparation: How Are We Doing?", and "Developing Technology-Proficient University Professors and Prospective K-12 Teachers."

Dr. Hughes is a member of the National Council of Teachers of Mathematics (NCTM), the Benjamin Banneker Association, the American Educational Research Association (AERA), and the National Council on Measurement in Education (NCME).

She also serves on the National Assessment of Educational Progress (NAEP) Validity Studies (NVS) Panel as a mathematics and test/item bias consultant. The NVS Panel is coordinated by the American Institutes for Research (AIR) and reports to the National Center for Education Statistics (NCES).

FRANCINE E. JEFFERSON, PhD

U.S. Department of Commerce

Dr. Francine E. Jefferson is a Telecommunications Policy Analyst with the National Telecommunications and Information Administration (NTIA) in the U.S. Department of Commerce. Her primary role is that of evaluation specialist for NTIA's Technology Opportunities Program (TOP).

Since coming to NTIA, she has been responsible for the design of a Web-based performance reporting system and the development of evaluation guides for TOP grant recipient. She conducts yearly research and evaluation sessions for new grantees and technical assistance workshops on evaluation for prospective applicants.

Dr. Jefferson came to NTIA after having spent six years at Cheyney University of Pennsylvania. During that time, she served as director of distance learning and was responsible for the design and implementation of Cheyney's Telecommunications Center and the Cheyney Education and Research Telecommunications Network (CERTNet). Although Dr. Jefferson has been engaged in the uses of telecommunications and information technologies for educational purposes for about 30 years, she has been a researcher and evaluator for most of her professional life.

Dr. Jefferson received her PhD in sociology, and a certificate in applied sociology, from the University of Pittsburgh. She has been a senior evaluator with the U.S. General Accounting Office, Dean of Graduate and External Programs, and served on the Mayor's Telecommunications Task Force for the City of Philadelphia.

JAMES M. PATTON, EdD

The College of William and Mary

James M. Patton is Professor of Leadership and Special Education at the College of William and Mary in Virginia. He was formerly Associate Dean of Academic Programs and Director of Project Mandala, a Federally funded research and development project aimed at identifying and serving selected students and their families who exhibit at-risk and at-promise characteristics. He directed professional development, teacher education and evaluation programs for the Commonwealth of Virginia for three years. He has also served as Dean of the School of Education and Chairperson of the Department of Special Education at Virginia State University and Chair of the Special Education Program at Hampton University. Dr. Patton has taught special education in the public schools of Louisville, Kentucky, where he also directed the Career Opportunities program, a Federally funded effort to increase the number of indigenous inner-city teachers in the Louisville Public Schools.

Dr. Patton has served extensively as an evaluator employing both quantitative and qualitative methodology in his evaluation efforts. He has served as an evaluator of 15 major programs with total funding of \$11.5 million. Some selected examples of his evaluation efforts include Hampton University's Title III Program, Project Teagle in the School of Nursing at Hampton University, Third Party Evaluator at Virginia Polytechnic Institute and State University, Kaiser Family Foundation Funded Project at the School of Medicine, Morehouse College, Pre-Service Teacher Institute, NASA, Langley, United States Office of Special Education funded projects, ILIAD and ASPIIRE, Newport News Alliance for Youth Family Preservation and Family Support Program, The Parental Involvement Program of the Norfolk Public Schools, Norfolk, Virginia, and Project Success at South Carolina State University.

A member of the Executive Committee of the Council for Exceptional Children (CEC), Dr. Patton also serves as a Senior Scholar in the Shaklee Institute, a special education think tank. His major research interests include the educational and psychosocial development of African-Americans, particularly those with gifts and talents, the holistic development of African-American males, the social, political and economic correlates of mild disabilities, curriculum and pedagogical issues around multi-cultural education, and analysis of policies that affect people of color and those from low socioeconomic circumstances. His funded grants approximate \$4.7 million.

CARLOS M. RODRÍGUEZ, PhD

American Institutes for Research

Expertise

Primary research interests focus on issues of equity, access and educational attainment of minority populations from K-12 through higher education, adult education and educational partnerships as mechanisms for educational reform.

Projects

Currently serves as Project Director for two national studies: the longitudinal study of three cohorts of students involved with the Equity 2000 Project of The College Board, and the Partnerships for Health Professions Education of the Health Resources and Services Administration. He also serves as a task leader related to bias and sensitivity reviews of test items in the development of the Voluntary National Tests. Additionally, he serves as a task leader for a national study to identify the effective or "what works" educational practices with adults of limited literacy skills. Other recent projects include advising on bias and sensitivity issues related to the assessment of special needs and limited English proficient students and the development of a special analysis of trends and challenges confronting education development in Latin American countries.

Career

Serves as a key advisor to the White House Initiative on Educational Excellence for Hispanic Americans and the National Council for Community and Educational Partnerships. He delivered prepared remarks on August 2, 1999 at the White House for the First White House Conference on Hispanic Children and Youth convened by First Lady Hillary Rodham Clinton. He also holds an appointment as Associate Professor and Scholar-In-Residence at American University in Washington, DC.

Selected Publications

- "A Practitioner's Perspective" in *Access Denied: Race, Ethnicity and the Scientific Enterprise*, Oxford Press, December 1999.
- *A Profile of Policies and Practices for Limited English Proficient Students: Screening Methods, Program Support, and Teacher Training (SASS 1993-94)*, Statistical Analysis Report, NCES, January 1997.
- *America on the Fault Line: Hispanic American Education*, for the Presidential Commission on Educational Excellence for Hispanic Americans, Washington, DC, 1996.

Honors**Associations**

- Fellow, The Spencer Foundation of the Woodrow Wilson Foundation
- American Association for the Study of Higher Education
- Higher Education Group of Washington, DC.
- NABE
- TESOL
- AERA

Education

Received his Ph.D. from the University of Arizona in 1993 in the field of Higher Education Administration and his Masters of Arts in Bicultural and Bilingual Studies from the University of Texas at San Antonio in 1989.

THELMA L. SPENCER, EdD
Educational Consultant

Education

EdD University of Colorado-Boulder
MA Case Western Reserve University
BA Case Western Reserve University

Professional Experience

Current Independent consultant on institutional and program evaluation, performance assessment, faculty and staff development, curriculum and test development, test-taking skills and strategies, interdisciplinary critical thinking skills, instructional improvement

1988-1989 Eminent Scholar, Norfolk State University (VA) Graduate School of Education, adjunct professor

1969-1988 Educational Testing Service, Princeton, NJ: Assistant Program Director, National Teacher Examinations (NTE) and Program Director, Teacher Education Examination Program (TEEP) 1969-1972; Executive Associate for School and College Relations 1972-1988

Previous Experience

Teacher, Cleveland (OH) public schools; deputy probation officer, Los Angeles County (CA); casework visitor, Cuyahoga County (OH) Child Welfare Department

Related Experiences

Visiting Professor, Princeton University, 1976-77; Urban Mass Transit Administration (UMTA, DOT) Advisory Group, 1970-79

Affiliations

American Educational Research Association
American Psychological Association
Association for Supervision and Curriculum Development
National Council on Measurement in Education
University of Colorado, Graduate School Advisory Council

FLORALINE I. STEVENS, EdD
Floraline Stevens and Associates

Floraline I. Stevens received her Bachelor of Science degree from the University of Southern California, and Master of Education and Doctor of Education degrees from the University of California, Los Angeles (UCLA). She held the following positions in the Los Angeles Unified School District (LAUSD): teacher, evaluation specialist, testing coordinator, assistance director for research and evaluation, and director of research and evaluation from 1979 to 1994. She was the 1991-92 American Educational Research Association's Senior Research Fellow at the National Center for Education Statistics (NCES), U.S. Department of Education in Washington, DC; and from 1992-94 was a Program Director at the National Science Foundation, Division of Research, Evaluation and Dissemination. She retired from LAUSD in 1994 and currently serves as an independent evaluation and research consultant. Also, Dr. Stevens is a research associate at Temple University's Laboratory for Student Success (LSS), Mid-Atlantic Regional Educational Research Laboratory. She serves on several evaluation advisory committees including the U.S. Department of Education and National Education Association. She is a former vice-president for Division H (School Evaluation and Program Development), American Educational Research Association, and is the chair-designate of the Research into Practice Committee.

SHEILA D. THOMPSON, PhD
Howard University

Dr. Sheila Thompson has been a Senior Research Associate with the Center for Research on the Education of Students Placed At Risk (CRESPAR) at Howard University since 1995. She is currently working with CRESPAR's Assessment and Evaluation Innovations Project and its Talent Development Elementary School Project. Dr. Thompson's previous research and evaluation experience includes her employment with the Maryland State Department of Education; the District of Columbia Public Schools; A Better Chance, Inc.; the University of Maryland at Baltimore, and Research and Evaluation Associates, Inc.

As a professional consultant, Dr. Thompson has completed various projects for the U.S. Department of Education's Office of Educational Research and Improvement (OERI); the National Institute on Student Achievement, Curriculum, and Assessment; and the National Center for Education Statistics (NCES). These activities have included her work with studies related to large-scale assessments such as the National Assessment of Educational Progress (NAEP). She also provided program evaluations and reviews for the General Educational Development (GED) Testing Service of the American Council on Education; Metis Associates; A Better Chance, Inc.; and the Human Resource Research Organization (HumRRO).

Dr. Thompson is a member of the American Educational Research Association (AERA): Divisions D (Measurement and Research Methodology) and H (School Evaluation and Program Development) and the Special Interest Groups: Research Focus on Black Education and the Talent Development of Students Placed At Risk. She is also a member of the National Council on Measurement in Education (NCME). She was formerly the Chair of the Minority Issues and Testing Committee of NCME.

Dr. Thompson received a BS degree in Psychology from Morgan State University in 1980, a MA degree in Educational Psychology from Michigan State University in 1981, and a PhD in Educational Psychology from Howard University in 1989. While at Howard University, she participated in the Summer Fellowship Program in Research for Graduate Students, sponsored by the Educational Testing Service. She is an inaugural inductee of the Morgan State University Psychology Department Hall of Fame. Dr. Thompson is listed in *Outstanding Young Women of America* and has been initiated into Promethean Kappa Tau and Psi Chi Honor Societies.

Dr. Thompson volunteers in local public and private schools in activities ranging from read-a-thons to science fairs. Her additional community service involvement includes her work as Secretary of the National Board of Directors of Choice Services International, Inc.; as a member Delta Sigma Theta Sorority, Inc.; and as the Church School Superintendent of Payne Memorial A.M.E. Church in Jessup, Maryland.

About the National Science Foundation

The National Science Foundation (NSF) funds research and education in most fields of science and engineering. Awardees are wholly responsible for conducting their project activities and preparing the results for publication. Thus, the Foundation does not assume responsibility for such findings or their interpretation.

NSF welcomes proposals from all qualified scientists, engineers and educators. The Foundation strongly encourages women, minorities and persons with disabilities to compete fully in its programs. In accordance with Federal statutes, regulations and NSF policies, no person on grounds of race, color, age, sex, national origin or disability shall be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving financial assistance from NSF (unless otherwise specified in the eligibility requirements for a particular program).

Facilitation Awards for Scientists and Engineers with Disabilities (FASSED) provide funding for special assistance or equipment to enable persons with disabilities (investigators and other staff, including student research assistants) to work on NSF-supported projects. See the program announcement/solicitation for further information.

The National Science Foundation has Telephonic Device for the Deaf (TDD) and Federal Information Relay Service (FIRS) capabilities that enable individuals with hearing impairments to communicate with the Foundation about NSF programs, employment or general information. TDD may be accessed at (703) 292-5090, FIRS at 1-800-877-8339.

The National Science Foundation is committed to making all of the information we publish easy to understand. If you have a suggestion about how to improve the clarity of this document or other NSF-published materials, please contact us at plainlanguage@nsf.gov.