

DOCUMENT RESUME

ED 398 079

SE 058 799

AUTHOR Haug, Carolyn; And Others  
 TITLE Dilemmas in Evaluating Professional Networks: The Case of the Mathematicians and Education Reform Forum.  
 SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.  
 PUB DATE Apr 96  
 CONTRACT RED-9255347  
 NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association and the National Council on Measurement in Education (New York, NY, April, 1996).  
 PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)  
 EDRS PRICE MF01/PC01 Plus Postage.  
 DESCRIPTORS Demonstration Programs; Higher Education; \*Mathematicians; Mathematics Education; Program Evaluation

ABSTRACT

The purpose of this paper is to describe experiences evaluating the effectiveness of the Mathematicians and Education Reform (MER) Forum, a network of research mathematicians interested in educational issues. After a brief description of MER, three traditional evaluation models are considered to illustrate shortcomings found with using these models for network evaluation: quasi-experiments, objectives-based evaluation, and responsive evaluation. Finally, brief scenarios are constructed of how each of these models might be applied to the MER evaluation. In each of these three hypothetical scenarios, conclusions are proposed that seem viable. (MKR)

\*\*\*\*\*  
 \* Reproductions supplied by EDRS are the best that can be made \*  
 \* from the original document. \*  
 \*\*\*\*\*

PERMISSION TO REPRODUCE AND  
DISSEMINATE THIS MATERIAL  
HAS BEEN GRANTED BY

*C.A. Haug*

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

This document has been reproduced as  
received from the person or organization  
originating it  
 Minor changes have been made to improve  
reproduction quality

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official  
OERI position or policy

ED 398 079

## **Dilemmas in evaluating professional networks:**

### **The case of the Mathematicians and Education Reform Forum**

Carolyn Haug, Scott F. Marion & Ernest R. House

University of Colorado, Boulder

Paper presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, New York, April 1995. Research on Evaluation Special Interest Group.

This research was supported by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation and the National Center for Education Statistics (U.S. Department of Education) under NSF Grant #RED-9255347. Opinions reflect those of the authors and do not necessarily reflect those of the granting agencies.

The purpose of this paper is to describe our experiences evaluating the effectiveness of the Mathematicians and Education Reform (MER) Forum, a network of research mathematicians interested in educational issues. (The technical report contains more complete information concerning the evaluation; see Haug & Marion, 1996.) While conducting this evaluation, we found ourselves modifying standard evaluation approaches to such an extent that we felt it important to present these issues to a community of professional evaluators. Our objective is to stimulate discussion about how to evaluate voluntary professional networks. In this paper we explain why evaluators might be interested in voluntary networks and then use our evaluation of MER as an example of how evaluations might best capture the effectiveness of these networks. After our brief description of MER, we consider three traditional evaluation models -- quasi-experiments (Campbell & Stanley, 1966), objectives-based evaluation (Tyler, 1993) and responsive evaluation (Stake, 1993) -- to illustrate the shortcomings we found with using these models for network evaluation. We constructed brief scenarios of how each of these models might be applied to the MER evaluation. In each of these three hypothetical scenarios, we propose conclusions that seemed viable, based on our understandings of these models and MER. Undoubtedly, there are other possible outcomes.

#### Why an interest in evaluating voluntary professional networks?

Systemic reform of mathematics, science and engineering education, especially increasing the participation and capabilities of US citizens in mathematics, science and technology is considered crucial to the national interests of the United States. In fact, Goal 4 of the recently authorized Educate America Act states that: "By the year 2000, the

United States students will be first in the world in mathematics and science.” (Educate America Act of 1994). Means to these desired ends have taken many forms. One of the most popular is standards-based reform where content and performance standards are sanctioned by some central authority (e.g., state or Federal) and assessments are designed to measure progress towards these standards. Usually the results are made public in order to increase the “stakes” attached to these assessments based upon the assumption that if the stakes are high enough teachers will teach (and students will learn) the specific content standards. While this movement has many influential supporters (e.g., Smith & O’Day, 1991; Resnick & Resnick, 1992), others have been calling for a more decentralized approach to educational reform (e.g., Darling-Hammond, 1994; Lieberman & McLaughlin, 1992).

While much of this debate has been focused on K-12 education, there is interest also in improving the education of postsecondary students in the United States. In 1986, the National Association of Governors cited the quality of higher education as one of its seven priorities for educational areas needing improvement (National Association of Governors, 1988). Unlike the K-12 reforms, however, the push for improvement in higher education was decentralized and directed almost exclusively toward new assessment policies. That is, the Governors allowed each institution to develop its own assessment and accountability plans. It is unlikely that any centralized efforts will be implemented in the near future that meaningfully improve college education for a variety of reasons, including the relatively high professional status of most university educators, the “academic freedom” of both individual faculty members and university departments,

differences in the governance structure of K-12 systems compared to university systems (even public systems), and varied missions across universities. Therefore, voluntary professional associations appear to be the most likely vehicle to promote systemic change at the postsecondary level.

“Systemic reform” is a popular goal of educational reform but the meaning of this term is often vague. One definition which we found particularly useful comes from Jenness & Barley (1995) who wrote

*Systemic reform is characterized by (1) development of new or reformulation of existing systems rather than simply conducting new programs or activities, (2) the engagement of key partners in effective relationships to promote change, and (3) the identification of new roles and relationships for all players inside and outside the educational system. ... A systemic approach to reform differs considerably from the more traditional ‘project mentality’ approach of the past. Systemic change is about building infrastructure for reform; replacing/revising system elements rather than adding new ones; focusing change on oneself first rather than on others; building on system strengths rather than fixing deficits; understanding that reform is a long-term, evolving process, not a ‘quick fix’; addressing all dimensions of the system rather than focusing on one or two; building capacities and marshaling resources; being part of mainstream reform efforts, not focusing on specific projects; focusing on lessons learned rather than concentrating efforts on creating models for replication; placing power in the hands of those in the system rather than relying only on funders or top-down change agents.*

Another view of systemic reform comes from the National Science Foundation. It is consistent with Jenness’ and Barley’s ideas but more specific. NSF’s Statewide Systemic Initiatives (SSI) project is based on the premise that

*[M]eaningful reforms in schools are most likely to be achieved through state initiatives that set clear and ambitious learning goals and standards, align all of the available policy levers in support of reform, stimulate school-level initiatives, and mobilize human and fiscal resources to support these changes. (Consortium for Policy Research in Education, 1995)*

In this paper, we focus on voluntary -- often "bottom-up" -- approaches to systemic reform. In doing so, we discuss concerns about how to best evaluate the effectiveness of one of these approaches.

### Network Evaluation Literature

There has been much theory development regarding the evaluation of programs, products, and personnel, yet little has been written about evaluating voluntary professional networks. Attributing effects or describing relationships is difficult with standard program evaluations; in the case of a relatively loose association of professional voluntary networks, it is even more complex. Networks are becoming increasingly popular in educational reform. According to Lieberman & McLaughlin (1992)

*The experience of diverse networks suggests that policy can lever change more effectively if it takes an indirect approach -- concentrating on the environments available to support and stimulate teachers' professional growth -- than if it directly tackles concerns about teachers' knowledge base and classroom competencies. In this period of intensive school reform, when traditional inservice training and staff development have been shown to be inadequate, networks can provide fresh ways of thinking about teacher learning. (p. 677)*

Although evaluators have written little about professional networks, some discussion has evolved regarding related issues, namely the evaluation of research and the use of evaluation results. In the context of evaluating the effectiveness of research, House (1995) notes problems paralleling those encountered in evaluating the effectiveness of networks. House cautions that both intended and unintended impacts, which are often elusive, must be considered. One of the problems with tracking the results of a particular piece of research is that new ideas often are "reworked or combined with other ideas to

produce 'second-generation' ideas" (p. 5). In our view, this is similar to tracing communication through a grapevine. Many results are hard to trace because they take so long to emerge and, even then, may be indirect or obscure. Ideas spawned by particular research may influence people in ways in which they are unaware of and unable to attribute to a particular source. As such, it is very difficult for evaluators to trace these impacts.

Weiss has noted similar problems with assessing the long term usefulness of evaluation results. She refers to the impact of evaluation results as "knowledge creep" (1991, p. 221) noting that incremental changes to programs as a result of evaluations are valuable because small changes lead to cumulative change. Weiss writes "Even modest progress should be cherished...Incremental improvements have been made, and cumulative increments can add up to pretty significant change after all." (p. 227) At the same time, she warns that cumulative effects are hard to measure and even harder to attribute to particular incremental changes. Not only is policy hard to change, radical changes may not be desirable after all because they may cause unanticipated negative side effects.

Sociologists describe a particular type of social network, an affiliation network, which consists of a *group* of actors and a *set* of events in contrast to a smaller network composed of a dyad or triad of actors or events (Wasserman & Faust, 1994). In describing the attributes of affiliation networks, Wasserman & Faust state

*Joint participation in events not only provides the opportunity for actors to interact, but also increases the probability that direct pairwise ties (such as acquaintanceship) will develop between actors. For example, belonging to the same club (voluntary organization, boards of directors, political party, labor union, committee, and so on) provides the opportunity for people to meet and interact, and thus constitutes a link between individuals. Similarly, when a person (or a number of people)*

*participates in more than one event, a linkage is established between the two events. Overlap in group membership allows for the flow of information between groups, and perhaps coordination of the groups' actions ... The fact that events can be described as collections of actors affiliated with them and actors can be described as collections of events with which they are affiliated is a distinctive feature of affiliation networks. (pp. 293-294)*

As can be gleaned from this quote, social network analysis is centrally concerned with relations and interdependence among actors and views linkages as potential conduits of information. In such analyses, the network is the unit of analysis rather than the individual. Characteristics of the individual actors could be part the analysis in so much as they describe network composition. However, the primary concern is measurement of the ties among actors and events.

Properties of affiliation networks include reachability and connectedness, concepts which allow us to consider whether paths exist between pairs of actors and/or events. Once paths are established, one can also study the value or strength of these paths, most likely finding that some actors or sub-groups will be more strongly connected than others. Alternatively, one can examine the situation in terms of reachability and address the number of intermediaries it takes to reach a particular person.

As a rationale for studying affiliation networks, these authors claim that social groups provide opportunities for interpersonal influence and the formation of personal ties between actors, "such as becoming acquainted or becoming friends." (p. 297) McPherson (1982, cited in Wasserman & Faust) has noted that voluntary organizations with face-to-face contact time provide members with the benefit of "heightened probability of contact." (p. 226)

While our ideas are closely aligned with the sociologists concerning the important aspects of social networks, our choice of methods to examine these issues distinguishes us from them. By and large, the methods described by Wasserman & Faust and others are complex mathematical modeling techniques. In contrast, we prefer more qualitative methods for assessing the transfer of information among network nodes, as described later in this paper. (Further citations for literature of network analysis can be found in House, Marion, Rastelli, Aguilera, Weston & Min, 1996.)

### An Example: The MER Forum Evaluation

The MER Forum is a network composed primarily of research mathematicians with subgroups of mathematics educators and kindergarten through 12th grade math teachers. Supported by National Science Foundation (NSF) since its inception in 1988, the MER Forum has evolved from an organization primarily targeted at individuals to now including a component for entire mathematics departments. Its current mission is "to facilitate the institutionalization of mathematics education reform within the mathematics community." The Forum does not have a formal membership (e.g., no dues or formal application to stay involved with the network), rather it consists of those people who attend its numerous workshops or receive its publications. Currently, there are over 730 MER members. Although the membership is informal and fluid, the governance is more structured: Four Co-directors, one of whom serves as an executive director-type role, make most of the organizational decisions; thirteen Advisory Committee members share responsibility for the individual component of the network; and ten Task Force members contribute to decisions concerning the departmental component of MER.

The beliefs currently driving MER, as stated in their most recently funded proposal, are that

*[a]s disciplinary experts in the content, methods, and directions of modern mathematics, mathematicians have a great deal to contribute to mathematics education reform; as teachers of mathematics, mathematicians have a great deal to learn from other professionals involved in mathematics education. Further, there is growing awareness that many issues in mathematics education are systemic in nature and must be addressed at all levels and by all constituencies (MER funding proposal, 1993).*

The Forum has recently initiated a "Departmental Network" which consists of thirteen Research-I university mathematics departments to foster systemic reform of mathematics education. MER purposely invited outstanding research departments that are supportive of educational reform efforts. MER restricted participation to research universities because it believes these function as role models for smaller institutions and if research universities "buy into this" then other institutions will as well. MER required that several faculty and administrators from each mathematics department be supportive of math education reform in order to participate in the Departmental Network.

As evaluators for NSF, we conducted a formal evaluation of the MER Forum with the intent of providing information to two separate audiences: (1) the NSF Program Officer expected summative information useful for future funding decisions and (2) the MER Forum Co-Directors, Advisory Committee and Task Force members and participants wanted formative information, assuming that program funding was to continue. In this paper we consider the following subset of our evaluation questions (the complete listing of evaluation questions can be found in Haug & Marion, 1996):

1. Is the MER network an effective means of mathematics education reform?

- A. Do MER participants engage in different and effective classroom, outreach, development, and/or research behaviors as a result of their participation in MER?
  - B. How does MER affect workshop participants? Do MER participants feel that they are more aware and involved in mathematics reform as a result of their participation in MER workshops?
2. Is MER an example of systemic reform?
- A. If MER is systemic, how does it spread?
    - 1. How do the partnerships with other organizations (AMS, MAA, etc.) allow MER to spread its reform agenda?
    - 2. Does MER work with and/or contribute to other reform efforts in order to further mathematics education reform? If so, how?
  - B. How does MER affect the community of mathematicians? Has MER had an impact on the way mathematics education is viewed within this professional group?
  - C. How does MER affect the community of mathematics educators? Has participation in MER encouraged mathematicians to make connection with mathematics education faculty?
  - D. Is the departmental network an effective means of encouraging systemic reform?

### Scenarios based on various evaluation models

We developed scenarios that we think reflect how our evaluation might have proceeded if we decided to use another evaluation model to evaluate MER. We are particularly interested in different conclusions we might have reached using other evaluation designs. As mentioned earlier, the three other possible approaches to judging MER we examine are quasi-experimental, objectives-based and responsive evaluation.

*Quasi-experiments.* In the case of quasi-experimental evaluations, we did not have to create an hypothetical situation for an evaluation of MER because a previous MER evaluator used this approach. Berger (1992) used a quasi-experimental design where the evaluator "matched" the group of MER participants to a control group of other mathematicians interested in educational issues. A random sample of subscribers to a Mathematical Association of American (MAA) publication, "UME Trends," was chosen as the control group because this journal is targeted toward an audience of mathematicians interested in education. In this evaluation, MER participants were considered a treatment group. Following standard quasi-experimental methods, the two groups were compared on numerous dependent variables such as the types of K-12 activities, students (urban, rural, minority, female and gifted) and teachers with which they were involved and awareness of issues in mathematics education reform. Data were collected from both groups regarding these things before MER began (1988) and after three years of "treatment." This design resulted in very positive conclusions regarding MER's effects on participants. In his conclusions, the evaluator commented that self-selection of participants may have had a significant impact on these results.

We do not believe that this approach is viable for the MER evaluation because there was no way to control for self-selection into the MER network. By virtue of the voluntary nature of MER participation, there is no fair way to establish a meaningful comparison group. Further, this evaluator did not interview any participants, observe any program activities, or seek enough contextual information to describe MER.

*Objectives-based.* Another possible method for the evaluation would have been to follow Tyler's objective model (Tyler, 1993). Had we adopted an objectives-based model, we would have determined MER's intended program objectives or outcomes, and then assessed the degree to which MER achieved them. The intent of our evaluation conclusions would be to improve the program so that intended outcomes are better met. Using this approach, we would have started the evaluation by reading the program's mission statement and interviewing the Co-directors about expected outcomes. For example, two goals of this organization are to facilitate participation of mathematicians in educational reform and learn from other educators.

*As disciplinary experts in the content, methods, and directions of modern mathematics, mathematicians have a great deal to contribute to mathematics education reform; as teachers of mathematics, mathematicians have a great deal to learn from other professionals involved in mathematics education... (MER funding proposal, 1993)*

Based on this information, MER may have adequately achieved the first goal of encouraging mathematicians to participate in reform activities, but would certainly fail on its other objective requiring learning from professionals involved in mathematics education, which would seem to include professional educators. To judge MER by this criteria would be to hold it to its own claims and nothing else.

In our opinion, this would be a better approach to the evaluation than a quasi-experiment but would not be adequate. Implicit goals or goals of the participants (one important group of stakeholders) would not be considered. For example, many of the mathematics educators (not research mathematicians) who participate in MER do not feel valued by the organization. Positive or negative side effects would not be addressed. Further, MER would receive no recognition that it legitimizes education activities within the mathematics community. Using this model, the combination of narrow program objectives and rather lofty goals put forth in the funding proposal would have led us to conclude that MER is not effective. One reason for this conclusion is that MER includes as a major goal the involvement of mathematicians in K-12 mathematics reform and it has clearly fallen short in this area. However, its influence in postsecondary mathematics education reform is notable and successful. In fact, we have come to believe that these mathematicians and this organization would be most effective if they directed their attention to higher education rather than K-12 education. These types of evaluation decisions and conclusions would not be permitted in the fairly strict protocol dictated in an objectives-based approach.

*Responsive evaluation.* Stake's (1993) responsive evaluation model comes the closest of the three to our approach. Stake advocates qualitative evaluation methods for the purpose of describing individual situations at the local level so that they can be understood by others. Stake's emphasis is on program activities rather than program goals. Instead of determining early in the evaluation the types of data to be collected and analyzed, Stake's theory allows questions and methods to emerge from observations.

Of the two methodologies which are consistent with Stake's beliefs, naturalistic inquiry and case studies, we initially adopted a case study approach to the MER evaluation. We felt that naturalistic inquiry was not appropriate for this evaluation (perhaps not for any evaluation) because it stops short of judging the program, placing the burden of drawing conclusions on the reader. Consistent with our belief that the essential element in any program evaluation is determining the worth or value of the program, we are skeptical of methodologies that do not include this component. However, the case study approach appealed to us because it does not have this "value-phobic" (Scriven, 1983, p. 230) flaw and allowed us to investigate both intended and side effects of MER. The case study model allowed us to spend a significant amount of time becoming familiar with MER and generate questions while we were in the field.

Although Stake's model does not suffer from the same problems as the other two and we suspect we would have reached some of the same conclusions about MER (e.g., benefits to active participants), we believe Stake's model would have failed to detect important information. With this case study model we perhaps would have stumbled upon some non-participants and interviewed them as well but this is less certain. We believe we would have failed to detect the extent to which MER operates by word-of-mouth beyond the participants to non-participants who may never have heard of MER. Stake's case study model would have not indicated that we were working with a communication web and needed to determine its boundaries. Adopting Stake's case study approach was our initial thought which later evolved into a much more complex activity. We found, after beginning to collect data, that in addition to a standard case study approach we needed to

(1) incorporate an extensive written survey of the 750 participants and conduct the corresponding quantitative analyses, and (2) extend the interviewing and observations beyond those people directly connected to the Forum, as discussed in the following section.

*Our approach.* While we adopted a case study methodology -- an approach consistent with Stake's beliefs -- we did not completely share Stake's exclusively local-level focus of the case study. Instead, we combined the local level information with quantitative survey data, a modification to Stake's approach which is certainly not new to practicing evaluators. We believe that the novelty we used was our conceptualization of the target group for interviewing and observations which we expanded beyond participants to include non-participants associated by direct and indirect ways to participants. To illustrate our point, following is a typical scenario with which we were faced during this evaluation:

*One mathematician attended a MER workshop on how to improve undergraduate math courses for those math majors intending to teach math at elementary and secondary levels. Upon returning to her university mathematics department, she shared this information with a colleague but failed to mention that she obtained these ideas from the MER workshop. This colleague may have adopted and shared these ideas with other colleagues but, at this point, the connection to the MER Forum had been lost. (Interview with mathematician)*

Through in-depth interviews with a diverse group of department personnel, we hoped to trace MER-generated ideas from the core group of MER participants to other faculty in university mathematics departments. We sought to find a way to detect subtle lines of communications among people so that the impact of the network could be reasonably assessed. In particular, we paid attention to uncovering and making explicit

underlying communication patterns which involve the MER Forum. With these beliefs, we employed a modified case study approach. Our primary data collection methods included written surveys of participants, participant-observations of workshops and advisory board meetings, direct observation of special conference presentations, in-depth interviews with the Co-directors, document review and four, week-long site visits to university mathematics departments. During these site visits, we observed mathematics and mathematics education classes and interviewed mathematics and mathematics education faculty (more than 20 faculty at each university), departmental and university-level administrators, and graduate and undergraduate students.

As a result of our adoption of Jenness' & Barley's conception of systemic reform, our criteria for judging MER's effectiveness in influencing systemic change consisted of the extent to which it encourages development of new working relationships within the mathematics community and among disciplines, and broad changes within the community of research mathematicians in addition to local or individual types of change. We do not mean to discount the local and individual changes that the Forum may facilitate and, in fact, have quite a bit to say about these positive changes on this level. However, individual changes are a slow route toward systemic change if departmental and university structures are ignored.

Once we began looking for indirect MER influences instead of narrowly focusing on changes directly attributable to MER, we discovered several possible MER successes and failures which otherwise we might have missed. For example, we found that MER provided support to mathematicians in efforts to improve their own teaching, leadership to

mathematics departments, and legitimization of educational interests. Although mathematicians generally could not attribute changes in their teaching directly to MER and believed that these changes were inevitable, the following comments suggest that these participants still attribute at least an indirect effect to MER.

*“Even Professor X, our big-shot hire from Harvard, mentors undergrads ... When people like Professor X do those kinds of things, it kind of sends a message to all of us.” [Interview with Mathematics Professor]*

*“Probably at least 1/2 of the tenured faculty are interested in teaching but somewhat ashamed about showing it because that is not how you earn points in a major research university. As a result, most of the effort to talk about teaching and the reform of teaching comes from a fairly small group of people. The small group of people that try to lead the discussion sort of act like a prick on the side of a hippopotamus or something like that.” [Interview with Mathematics Professor & former Undergraduate Chair]*

*“What’s happening is there are certain programs that are run by people who are dedicated to teaching in the university. Other faculty see this happening and that its OK to do. Its a generational change; its broadening out.” [Interview with Mathematics Professor]*

Although the majority of MER’s impact was reported to be on an individual level, to some extent MER also influenced broader changes within mathematics departments. Since many institutions of higher education sought to improve the education they offer to undergraduates during this same time frame, it is difficult to attribute these changes in departmental attitudes solely to MER. These shifts are equally likely to be the result of

external pressures on universities. Nonetheless it appears that MER helped provide some leadership as mathematics departments have been called upon to address educational issues. The following quotations illustrate these points:

*"Because of MER, many [faculty] who were once vehemently opposed to any reform are beginning to at least listen and lean toward change."*

*"I've been able to cite instances where the use of technology has proved useful to bolster moving my department in that direction."*

*"My ability to cite peer institutions to my administration was improved."*

*"It has given me ideas for submitting my own grants to improve the teaching/learning of university-level mathematics."*

In order for educational reform to become sustained in university mathematics departments, the typical reward structure of higher education institutions will have to support faculty members' participation in educational initiatives. We looked to see whether institutional rewards had been changed at any of the universities associated with MER. At some departments we found institutional support for educational endeavors had improved to include consideration of these activities in promotion and tenure decisions, release time for educational involvement, becoming more respectful of educational involvement and general administrative support.

*"Faculty have been promoted for national educational efforts; faculty have been hired for education work alone."*

*"Promotion to full professor for math education types."*

*"Teaching now an important part of tenure decisions."*

*"The department chair has become very aware of reform and is very supportive of it."*

*"Establishment of 90 minutes of faculty development time each week for in-service development. Support for a system of five faculty development workshops for in-service development."*

*"Support of release time for K-12 work."*

*"I, no longer, am criticized for my involvement. Much more education innovation is taking place."*

Another strength of the Forum that emerged from this line of inquiry is that MER helps legitimize mathematics education reform within the mathematics research community.

*"Affecting positively mathematics education reform mainly by legitimizing mathematics education and get rid of its second-class status."*

*"Give higher standing to consideration of teaching as part of serious work with university math departments."*

*"The 'respectability' of MER so that participants feel empowered to return to their campuses and push for changes."*

*"They have made mathematics education issues socially acceptable in research math departments."*

*"That it exists! That it is trying to create the awareness of how students learn mathematics."*

*"Mathematicians with their hearts in the right places -- paying attention to crucial educational issues."*

This open-minded approach to the evaluation also uncovered some negative attitudes toward MER which otherwise we might have missed. Some participants and non-participants felt that MER duplicated activities of the major professional organizations (e.g., MAA) and that most MER activities could be incorporated by them. Some participants say the benefits of the network could be realized over the Internet while

others say the face-to-face interactions are crucial to successful networking and dissemination of ideas. We were told:

*"Do we need another organization? In the beginning MER brought a fresh perspective and brought new players to the discussions. This is less the case now."*

*"Overlap in programming with MAA; overlap in policy issues with AMS."*

*"MER is 'nice' but probably not essential. MER's main function, I believe, is to try to get a significant number of faculty around the country involved in math education in an attempt to break down anticipated departmental resistance to change."*

An additional consideration in judging the value of MER is its cost/benefit ratio. Some mathematicians believed that if the program is not too expensive it should be funded so that some people can have personal interactions; on the other hand, they felt that if the program is very expensive it should not be funded. Of course, the next question that must be addressed is then, "How much is 'expensive?'" These same cost/benefit issues arise in other program evaluations as well but it seems that the impact of the network is so nebulous that it is even harder to answer them.

### Conclusions

It is not surprising that mathematicians (or academics in general) tend to work individually to change their own practice rather than combine efforts to change the system in which they work -- a system which is steeped in centuries of tradition and in which they have succeeded. Changing educational beliefs and practices is very difficult. It is the culture of academia to work on your own research, perhaps collaborating with other experts in your specific area but not necessarily with your neighbors next door. Systemic

reform will require mathematicians and others to begin to adopt the attitude that changing your own practices is important but is not enough to change the entire system of mathematics education. We support MER in its efforts to bring about systemic change and hope that this organization will put even more effort into changing entire university mathematics departments so that they value and reward educational activities of their faculty.

Broad-based reform, called for in much educational rhetoric and encouraged by the national standards and assessment movement, probably will require the creation of professional communities with specific interests such as the MER Forum. If this occurs, then we, as evaluators, must be able to address questions related to whether such professional networks are effective, how they work, and whether they contribute to systemic change. Unless the effects of voluntary networks can be evaluated fairly and accurately, it is unlikely that they will be able to compete for federal or foundation funding. We found that standard evaluation methods may not be sufficient for such research. Initially we hoped to discover, through our interviews and following leads, recognizable paths from the voluntary network through several intermediaries to actual mathematics education reform activities. In some cases we found such paths, however, in most cases the path was uncertain and landmarks unclear. Revealing the communication patterns through the network and beyond became one of the main foci of the evaluation.

The present study reports the findings of one evaluation addressing these questions with a view toward beginning or continuing a dialogue about network evaluations. Our research is limited to one professional network. We hope to stimulate interest in

examining other voluntary networks. As the Forum's best effort at institutionalizing education reform -- MER's Departmental Network -- matures, we will be better able to judge its progress.

### References

- Berger, T. (May, 1992). Evaluation of the Mathematicians in Education Reform Network. Technical report.
- Campbell, D. & Stanley, J. (1966). *Experimental and quasi-experimental design for research*. Chicago: Rand McNally.
- Consortium for Policy Research in Education, May 1995. Reforming Science, mathematics, and technology education: NSF's statewide systemic initiatives. New Brunswick, NJ: Author.
- Haug, C. & Marion, S.F. (1996). Professional Networks for Educational Change: An evaluation of the Mathematician and Education Reform Forum. Technical report.
- House, Ernest R., Marion, S., Rastelli, L., Aguilera, D., Weston, T. & Min, K. (1996). The Impact of CRESST R&D. Technical report.
- House, E. R. (1995). Evaluating Research and Development Impact: The case of the CRESST R & D Center. (unpublished document, February 1995)
- Jeness, M. & Barley, Z (1995). Using cluster evaluation in the context of science education reform. *New Directions for Program Evaluation*, no. 65, Spring, pp. 53-69.
- Lieberman, A. & McLaughlin, M.W. (May, 1992). Networks for educational change: Powerful and problematic. *Phi Delta Kappan*.
- Scriven, M. (1983). Evaluation Ideologies. In Madaus, G. & Stufflebeam, D. (Eds.) *Evaluation models: Viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff Publishers, pp. 229-260.
- Shadish, W., Cook, T. & Leviton, L. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.

Stake, R. E. (1993). Program evaluation, particularly responsive evaluation. In Madaus, G. & Stufflebeam, D. (Eds.) *Evaluation models: Viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff Publishers, pp. 287-310.

Tyler, R. W. (1993). A rationale for program evaluation. In Madaus, G. & Stufflebeam, D. (Eds.) *Evaluation models: Viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff Publishers, pp. 117-142.

Wasserman, S. & Faust, K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.

Weiss, C. H. (1991). Evaluation research in the political context: Sixteen years and four administrations later. In McLaughlin, M.W. and Phillips, D. C. (Eds.) *Evaluation and education: At quarter century*. Chicago: University of Chicago Press, pp. 211-231.