

# **UCLA**

## **UCLA Previously Published Works**

### **Title**

Program Staff Perceptions of Barriers to Evaluation Implementation

### **Permalink**

<https://escholarship.org/uc/item/3886273x>

### **Journal**

American Journal of Evaluation, 24(2)

### **ISSN**

1098-2140

### **Authors**

Taut, Sandy M  
Alkin, Marvin C

### **Publication Date**

2003-06-01

### **DOI**

10.1177/109821400302400205

Peer reviewed

# Program Staff Perceptions of Barriers to Evaluation Implementation

SANDY M. TAUT AND MARVIN C. ALKIN

## ABSTRACT

This study investigated the nature of barriers to the implementation of (external) program evaluation. The purpose of the study was to expand the scarce empirical research on this topic by adding program staff perspectives. Eighteen staff members of a university outreach program that had been subjected to external evaluation were interviewed (1) about their general thoughts on barriers to evaluation implementation, and (2) about their views on the explanatory value of factors found in research related to evaluation utilization and which we believed to be theoretically relevant for barriers to evaluation implementation. These factors are summarized as human, evaluation, and context factors. Asked what they considered to be barriers to evaluation implementation, the interviewees mentioned human factors most frequently. They talked about evaluation and context factors much less frequently. With regard to human factors, interviewees focused on the evaluator's social competence and program staff's lack of trust in evaluators and evaluation process. Taking into account the context of the study, particularly the participants' prior experiences with evaluation, the findings reinforce our understanding that in order to avoid barriers to evaluation implementation, it is important to create a trusting relationship with those affected by the evaluation through continuous participation and communication, and to conduct carefully planned, methodologically appropriate evaluations.

## INTRODUCTION

Empirical research can and should guide the practice of program evaluation. In tune with this assertion, some evaluation experts have noted the paucity of empirical research on program evaluation (see [Alkin, in press](#); [Chen, 1994](#); [Henry & Mark, in press](#); [Mark, 2001](#); [Rogers, 2001](#)). Particularly, there is not a noteworthy body of research on the impediments to effectively implementing an evaluation. In this study we considered perceived barriers to evaluation implementation in the context of a university outreach program. The study examined, from the perspective of a program staff (1) spontaneous, general perceptions of

---

**Sandy M. Taut** • Graduate School of Education and Information Studies, University of California, 2027 Moore Hall, Los Angeles, CA 90095-1521, USA; Tel: (1) 310-206-6944/825-4800; E-mails: [staut@ucla.edu](mailto:staut@ucla.edu), [alkin@gseis.ucla.edu](mailto:alkin@gseis.ucla.edu).

---

**American Journal of Evaluation**, Vol. 24, No. 2, 2003, pp. 213–226. All rights of reproduction in any form reserved. ISSN: 1098-2140 © 2003 by American Evaluation Association. Published by Elsevier Inc. All rights reserved.

---

barriers to evaluation implementation, and (2) the applicability of factors previously found to explain barriers to evaluation utilization. In this paper, when we say *barriers to evaluation* we refer to the problems faced by evaluators when trying to implement an evaluation. Such barriers can exist from the inception to the completion of a program evaluation cycle.

### STUDY CONTEXT

Given the potential importance of contextual influences on the nature of our findings, we begin by describing the setting of the study and the participants' background. The findings of this study depend particularly on the participants' frame of reference regarding evaluation, that is, their previous and current experiences with program evaluation. [Alkin, Daillak, and White \(1979, p. 238\)](#) noted that users' expectations for the evaluation and feelings based upon experiences with previous evaluations, as well as their conversations with others about evaluation and reading they may have done on the subject, combine to establish a user frame of reference for the specific evaluation. Therefore, throughout the entire interview, the participants were repeatedly encouraged to situate their statements in their experiences with program evaluation and to provide illustrative examples.

The university outreach program that provided the setting for this study was designed to better prepare underrepresented middle and high school students for college. The program consisted of two major components. One was a student-centered component, including such activities as counseling, academic enrichment, personal enrichment, and parental involvement. In addition, the program contained a second component that aimed at fostering longer-term systemic change related to college readiness. This was attempted through so-called School–University Partnerships. These partnerships were meant to support certain conditions in schools that are seen as contributing to students' preparedness for university (for an overview of university outreach programs see [U.S. Department of Education, 2001](#)). The program received university funding and was obligated to spend a certain amount on external evaluation. Requirements for the evaluation were minimal, so that much of the resources could be spent in whatever way the program staff and the evaluators they hired found most beneficial.

The participants of our study varied to a large extent regarding their level of experience with and knowledge about program evaluation. Some could reflect back on a lifetime of work experiences in the educational system, whereas for others this program provided their only experiences with evaluation. Despite these differences, most interviewees provided examples from the current program context. The program had recently experienced two very different evaluation approaches. The previous evaluator had conducted quasi-experimental studies. Program staff voiced their doubts about the usefulness of this approach for program improvement. In addition, some noted the lack of social competence of this evaluator. Staff pointed out the remarkable contrast to the current evaluation team that had been the evaluators for two years at the time of this study. The current team engaged in both qualitative and quantitative studies, depending on the evaluation purpose and questions as determined by the primary users of the evaluation. Therefore, program staff had been involved in the evaluation process. Relationship building had been a primary concern of the new evaluation team.

To frame the study, we asked the participants at the beginning of the interview to tell us about what they perceived to be others' attitudes toward evaluation, followed by a question

about what were their own attitudes, and a question about how they thought attitudes toward evaluation develop. The responses to these initial questions further helped put into context and appropriately interpret the study's main findings. General attitudes (attitudes of others) were said to be evenly split between positive and negative, while personal attitudes were depicted as very largely positive. Of course, it seems difficult to admit personal negative attitudes toward evaluation when talking to a member of the current evaluation team. According to the interviewees, positive attitudes often stemmed from experiences that resulted in personal benefits from the evaluation, whereas negative attitudes were often based on the fear of negative results, as experienced in prior evaluations. The perceived methodological quality of previously experienced evaluations (approach, design, methods, findings) also seemed to considerably influence attitudes toward evaluation. The interviewer noted that the higher the position of the participant in the program, the more knowledge they had about evaluation, and the more contact they had had with the current evaluation team, the more positive were their attitudes toward evaluation.

## THEORETICAL BACKGROUND

To some extent there are both prescriptive and descriptive (empirical) writings that address the opposite of barriers—evaluation facilitating factors. The prescriptions are, in essence, to be found in textbooks and evaluation conceptualizations that suggest to readers the proper ways of conducting effective evaluations. The closest empirical base is to be found in the extensive research on evaluation utilization (see Cousins & Leithwood, 1986; Hofstetter & Alkin, 2003; Leviton & Hughes, 1981). This body of research identifies the factors empirically found to be associated with a high degree of evaluation *utilization*. Presumably, the absence of these utilization-enhancing conditions would lead to a lack of utilization. Could these same factors, summarized in Alkin (1985) as *human, evaluation and context factors*, explain the existence of barriers to the effective *implementation* of evaluation? In conceptualizing this study we made this assumption. In other words, we took empirical research in one area, evaluation utilization, and tried to apply it to another, related area, evaluation implementation. We first asked program staff in general terms to share what they spontaneously thought to be barriers to evaluation implementation (an inductive approach in order to find rival explanations and unanticipated factors). Then we asked the participants to share their views on the relevance and applicability of the human, evaluation, and context factors for explaining such barriers (in an attempt to contribute to more cumulative knowledge about evaluation).

The human (or personal) factor describes evaluator and user characteristics—for example, previous experiences with evaluation, user knowledge about evaluation, and perceptions about the credibility of the evaluator.

The second of the three factors, the evaluation factor, concerns the way the evaluation itself is conducted. This includes procedural issues such as ethical conduct, but also the appropriate choice of an evaluation design, data collection methods, and the overall quality of outcome information provided by the evaluation.

The third factor refers to the context in which the evaluated program exists. Here, the question is how contextual variables, including the broad political and organizational background as well as specific program features and administrative structures, influence the barriers an evaluation has to face.

## STUDY PROCEDURES

Because we were interested in personal insights that reach well beyond placing checkmarks or even writing a few lines on a questionnaire, we decided to conduct a qualitative study based on face-to-face interviews as the data collection method. The first author conducted eighteen semi-structured, confidential, 45-minute interviews. The interview protocol enabled the interviewer to be adaptive to each individual at the expense of fully comparable results. The interviewees included staff from different hierarchical levels of the program: four academic and four administrative staff in high-level university positions, one administrative staff in a lower-level university position, six program implementation personnel, and three school district employees. An original list of interviewees was slightly altered and expanded as the study proceeded. In the end, we succeeded in completing all the intended interviews.

Seventeen out of eighteen interviews were tape-recorded and transcribed. The interviewees all received the interview transcriptions of their individual interviews for validation. Approximately half of the interviewees replied to the validation request. When respondents asked for changes or additions, they were very minor. Although all identifying information was removed from the data, anonymity could not be ensured because the principal investigator was engaged in both data collection and data analysis.

The analysis consisted of a cycle of categorizations, interpretations, and revisions of the category systems. All interviewee statements were categorized under their corresponding categories and subcategories. A few times one statement fit in two different categories or subcategories and was therefore double-coded (for details, see [Taut & Alkin, 2002](#)). As qualitative methodologists suggest (e.g., [Erickson, 1998](#)), we counted the frequencies of emerging themes in order to determine their prevalence, while being aware of the limited nature of our sample. We report these frequencies in the tables in the following section (FINDINGS).

We need to point out that the number of statements in one category depends on the scope of the category. A very broadly defined category (e.g., *evaluator competence*) has the potential of encompassing many statements, while a very specific category (e.g., *expectation of positive results*) can, by definition, only contain few. In general, when designing a category system, the data analyst should keep in mind the following guidelines ([Miles & Huberman, 1994](#)):

1. Categories of the same “order” should have approximately the same level of abstraction.
2. The data should be described completely by the category systems; no relevant statements should be left uncoded. “Other” categories should be avoided or kept at a minimum.
3. Categories need to be clearly defined (or labeled) so that other people arrive at the same categorizations. This includes minimal overlap between the categories within one category system.

We did our best to adhere to these guidelines. To check for potential bias in our data analysis due to the prior involvement of the principal investigator with the topic of the study, we employed two strategies: (1) the first author suggested category systems and specific categorizations of responses which were critiqued and modified by the second author and (2) the first part of the interview transcripts (unprompted statements on barriers) were given to an evaluation colleague unfamiliar with the study. We asked him to derive main themes from the

data. We then compared his results with our own analyses. The authors' categories were more detailed, that is, closer to the interviewees' statements than the themes derived by the evaluation colleague, but all of the latter's themes closely corresponded to the authors' categories.

## FINDINGS

### Spontaneous Perceptions of Barriers

After some introductory remarks and a few questions related to the participants' attitudes toward evaluation (see section on STUDY CONTEXT), the interviewer made the following statement: "Could I now ask you to share with me your thoughts on the following question: What do you think are the barriers to evaluation getting implemented? Think about your experiences with evaluation and try to provide some examples." The interviewer did not provide any further prompting because this part of the interview focused on participants' spontaneous remarks on and explanations of barriers. This approach allowed us to assess the salience of certain barriers from multiple staff perspectives. We first extracted themes from these statements and then used the three factors (human, evaluation, and context) as an organizing structure.

Table 1 shows the category system and the response frequencies for this part of the interview. Without any prompts, the participants talked most about human factors as explanations for barriers to evaluation (33 statements). Participants mentioned evaluation and context factors with far less frequency (20 and 15 statements, respectively). The importance of the human factor is also supported by prior research done in other settings (see Alkin, 1985; Alkin et al., 1979; Monsen, 2002; Patton et al., 1977; Taut, 2001). In addition, although we cannot present frequencies to support our claim, we noted differences between higher-level and lower-level staff. The latter tended to focus on *bad evaluation* and *bad evaluators* to explain barriers, whereas the former often considered a broader range of factors, indicating a possible correlation with the respondents' extent of knowledge about the process of program evaluation.

We differentiated the human factor themes into two subcategories: *evaluator competence* and *program staff issues*. Regarding the evaluator, participants talked most often about the importance of his or her *social competence*, particularly relationship-building skills. The following are examples of the social competence category:

When we started to work with [. . .], it was good that this team went out of their way to create the kind of working relationship with the staff that was necessary to get the evaluation going (Interviewee 3).

The personalities of the evaluators could be a barrier . . . (Interviewee 6).

Complementing this view by focusing on program staff issues, barriers related most often to *lack of trust*. The following are examples of this category:

Winning the stakeholders' confidence. You have to convince them that you are there to help and not to judge them or grade them. You have to get their trust (Interviewee 8).

The evaluator and the program people must have conversations about the purpose, the process, and the outcomes of the evaluation. A trusting relationship must be built (Interviewee 16).

The predominant evaluation factor subcategory was *gaining access to correct data*, followed by *inappropriate methods, procedures, instruments, etc.* Regarding the context,

**TABLE 1.**  
**Spontaneous Perceptions of Barriers**

<i>Categories</i>	<i>Frequency</i>
Human factor	
1. Evaluator competence	
(a) Social competence (relationship building)	8
(b) Context and program knowledge	4
(c) Technical competence	1
2. Program staff	
(a) Lack of trust	7
(b) Lack of knowledge about evaluation	4
(c) Threat of negative results	4
(d) Lack of appreciation of evaluation	4
(e) Prior negative experiences	1
Total	33
Evaluation factor	
7. Gaining access to correct data	7
8. Inappropriate methods, instruments, etc.	5
9. Lack of cooperation/detachment of evaluation	4
10. Lack of utility of evaluation	4
Total	20
Context factor	
11. Lack of resources of stakeholders (especially time)	7
12. Staff turnover (schools)	3
13. Program setting (schools)	3
14. Other program characteristics	2
Total	15

participants mainly held *lack of stakeholders' resources*, especially time, responsible for creating barriers to evaluation.

In summary, for this part of the interview we can highlight:

1. Program staff most frequently discussed the competence of the evaluator(s) as a barrier to effective evaluation (whereas evaluators often see program staff perceptions as being the source of resistance to evaluation; see Taut, 2001). They considered social competence, along with context/program knowledge, as much more important than technical competence.
2. Participants regarded program staff's lack of trust as the second most prevalent barrier to evaluation. Therefore, building trust seems an important objective at the onset of an evaluation.
3. Gaining access to correct data (inadequate data systems) was most often mentioned as an evaluation factor hindering the evaluation process.
4. Program staff consider lack of resources, especially lack of program staff time, as an important context factor hindering the conduct of evaluations.

**Prompted Perceptions of Factors Related to Barriers**

This part of the interview dealt with the three general factors (human, evaluation and context): “Now I would like to ask you to comment on some of the factors that evaluation researchers have used in the past to explain barriers faced by evaluation. Please feel free to dispute the relevance of the factors based on your experiences. Again, please provide examples to illustrate your views.” The interviewer introduced the three factors to the participants one at a time. For each factor, the interviewer gave a short, general definition and then asked the interviewees to share their thoughts on the topic. At this point, the interviewer generally had not mentioned any particular examples of the factors.

**Human factor.** “The first factor I would like to explore with you is called the human factor. This factor describes characteristics of people involved in an evaluation. What characteristics of people or psychological mechanisms could lead to potential barriers to evaluation?” From the interviewees’ statements following this general introduction of the human factor, we extracted eight subcategories (see Table 2). The following are examples of statements categorized under the most frequent theme, labeled *trust/fear issues*:

- There is the issue of fear, what are you trying to find out about me, and how is this going to impact me and be used against me [ . . . ] (Interviewee 2).
- Fear, people are afraid of participating. And the other part is lack of trust in the people doing the evaluation, which relates to fear, in terms of “What are they going to do with the information when they get it” (Interviewee 9).
- People think evaluations are reflecting them, they might feel threatened by the evaluation findings because it is perceived as their work (Interviewee 17).

This factor elicited the most responses from the participants. Seven out of the eighteen interviewees mentioned these trust/fear issues first when commenting on the question, indicating their high salience. For three interviewees, the social skills of the evaluator were the first, most salient response. Of course, the first and second categories are related: program staff’s trust (lack of fear) is based on a good relationship with an evaluator who has strong interpersonal skills.

**Context factor.** “The second factor used in evaluation research to explain barriers relates to the organizational setting of the program that is being evaluated, called the context factor.

**TABLE 2.**  
**Analysis of Human Factor**

<i>Subcategories</i>	<i>Frequency</i>
1. Trust/fear issues	10
2. Relationship building and interpersonal skills of evaluators (inadequate)	7
3. Communication about the evaluation (inadequate)	5
4. (Lack of) staff resources	5
5. (Lack of) stakeholder knowledge about evaluation	3
6. Personalities of stakeholders	3
7. Bad experiences with evaluation	2
8. Unclear benefits of evaluation	1



**TABLE 3.**  
**Analysis of Context Factor**

<i>Subcategories</i>	<i>Frequency</i>
1. Influences of program context and setting	5
2. Purpose and use of the evaluation (lack of clarity of)	3
3. Political structures, decisions, mandates	2
4. Accountability requirements	2
5. Organizational hierarchical structures	2
6. Organizational culture	2
7. Staff turnover	1
8. Level of implementation of the program	1
9. Evaluator context knowledge (not sufficient)	1

How could organizational or program features explain the barriers evaluation has to face?" The participants' comments following this definition resulted in nine categories (see Table 3). Please note that some of these subcategories, namely numbers 2 and 9, overlap with subcategories of the human factor and the evaluation factor. In practice, differentiation of the three factors was not always as clear as might be indicated theoretically. The following quotations illustrate the most frequently mentioned subcategory, *influences of program context and setting*:

Outreach is this big, vast component. Every school is different. [. . .] The school environment really has a big impact and that makes the evaluation difficult (Interviewee 7).

For Outreach in general, from a programmatic standpoint, there probably would not be many barriers [. . .] At the schools it might be different (Interviewee 14).

Overall, the context factor elicited little input from the study participants; in some instances, interviewees asked for clarification on its definition. Research on the distinct but related topic of evaluation utilization found that the program context has a strong influence on use (e.g., King & Pechman, 1984). Why we did not get many responses in this regard might be due to the participants' lack of knowledge of the context, or it might be easier to identify contextual influences from an outside perspective than as an insider. It would be insightful to investigate the context issue comparatively, across programs, instead of applying our approach of asking one program's staff to share their views on the influence of the context on barriers to evaluation implementation.

**Evaluation factor.** "The third and last factor used in evaluation research to explain barriers concerns the way the evaluation is conducted; it is called the evaluation factor. How could this factor explain the barriers evaluation has to face?" Seven subcategories summarize the statements following this definition (see Table 4). As in the discussion of the context factor, some of these subcategories (particularly numbers 2, 6, and 7) overlap with context and human factor subcategories. The following two statements exemplify the most frequent evaluation factor subcategory *good methods and meaningful data*:

The kinds of programs we are involved in are so complex. If the participants don't feel that the evaluation is getting at all the issues in a meaningful way, then they will see the evaluation as not helpful and meaningful (Interviewee 9).

Whenever you do an evaluation, the data are a major barrier (Interviewee 11).

**TABLE 4.**  
**Analysis of Evaluation Factor**

<i>Subcategories</i>	<i>Frequency</i>
1. Good methods and meaningful data (lack of)	8
2. Participation and information of staff (inadequate)	7
3. Context sensitivity of evaluation design (poor)	4
4. Justifiable and context-specific criteria/conclusions (poor)	4
5. Reporting issues	3
6. Issues with time	3
7. Competence of the evaluator (inadequate)	3

Many participants shared advice at this point regarding important features that the evaluation should display—as the category labels suggest: for example, *participation and information of staff*, *context sensitivity of evaluation design*, or *justifiable and context-specific criteria/conclusions*.

### DISCUSSION

In many respects, the study was both a research study and a part of an on-going evaluation process. Therefore, the principal investigator played a dual role, as both an evaluation researcher and a program evaluator. On the one hand, we are aware that this might present a particular challenge to the validity of the findings because the interviewer was not uninvolved and distant to the program staff as a more traditional research paradigm would demand. On the other hand, similar to some qualitative research, this dual role facilitated access to program staff. The positive relationship that at least some staff members had established with the interviewer might also have enhanced the validity of the findings by supporting an open and natural conversational atmosphere. In our view, if evaluation research insisted on an outsider role of the investigator, not much evaluation research would get done, because of the applied nature of our field. In addition, the extent of potential validity problems due to the dual role of the investigator depends on the types of questions asked. While we acknowledged that this dual role was problematic in regard to the validity of the attitude statements at the beginning of the interview (see *STUDY CONTEXT* section), this does not apply to the same extent to the barriers questions, which constituted the main focus of the study (see *FINDINGS* section).

We would like to reiterate that the findings of this study are context-specific. We only talked to program staff of one specific program, who had specific frames of reference for their comments. Thus, we do not purport to present generalizable findings. However, findings might be indicative of potential responses in other settings. We have provided a description of the program and evaluation context to let the reader decide whether the findings could be applied to settings they themselves are familiar with. We need more empirical studies in diverse settings in order to be able to judge the generality of our findings and, ideally, to develop a matrix of potential barriers to evaluation implementation in relation to relevant program and context characteristics.

In this study we examined barriers to evaluation implementation from general to specific. First we examined barriers directly, but provided interviewees with no specific structure. Then,

**TABLE 5.**  
**Overlap Between Unprompted and Prompted Discussion of Barriers to Evaluation**

<i>Main Categories</i>	<i>Unprompted Barriers Discussion Subcategories</i>	<i>Frequency</i>	<i>Prompted Barriers Discussion Subcategories<sup>a</sup></i>	<i>Frequency</i>	<i>Total</i>
Human factor	1. Evaluator competence:	13		11	24
	(a) Social competence	8	Interpersonal skills of evaluator	7	15
	(b) Program knowledge	4	Evaluator context knowledge	1	5
	(c) Technical competence	1	–	–	1
			Evaluator competence (general)	3	3
	2. Program staff:	20		19	39
	(a) Lack of trust	7	Trust/fear issues	10	17
	(b) Lack of knowledge	4	Staff evaluation knowledge	3	7
	(c) Threat of negative results	4	Unclear benefits of evaluation	1	5
	(d) Lack of appreciation	4	–	–	4
(e) Prior experiences	1	Bad experiences with evaluation	2	3	
		Personalities of staff	3	3	
Total		33		30	63
Evaluation factor	3. Gaining access to correct data	7	–	–	7
	4. Inappropriate methods, design, etc.	5	Good methods and meaningful data	8	17
			Context-sensitive evaluation design	4	
	5. Lack of cooperation/detachment	4	Communication about evaluation	5	16
			Participation and information of staff	7	
	6. Lack of utility of evaluation	4	Purpose and use of the evaluation	3	7
		Reporting issues	3	3	
		Justifiable and context-sensitive criteria and conclusions	4	4	
Total		20		34	54

Context factor	7. Lack of resources of stakeholders	7	Lack of staff resources	5	15
			Issues with time	3	
	8. Staff turnover	3	Staff turnover	1	4
	9. Program setting	3	Program context and setting	5	8
	10. Other program characteristics	2	Political structures, decisions, mandates	2	10
			Accountability requirements	2	
			Organizational hierarchical structures	2	
			Organizational culture	1	
			Level of implementation of program	1	
	<b>Total</b>	<b>15</b>		<b>22</b>	<b>37</b>

<sup>a</sup> Please note that descriptions from [Tables 2–4](#) are nearly identical to what is presented here, but they have been slightly modified or combined for purposes of presenting this comparison.

we asked about perceptions of barriers within the categories found in the evaluation utilization literature (human, evaluation and context factors). Reviewing the categorizations of both parts of the interview, we observed substantial overlap (see Table 5). When guided by the brief introduction of the three types of factors, participants mentioned most of the subcategories that they had noted in the prior unprompted discussion. A few new themes emerged, for example *reporting issues* and *justifiable and context sensitive criteria and conclusions* under the evaluation factor heading. The overall frequencies of the statements relating to the main categories in the prompted discussion compared to the unprompted discussion were as follows: (1) human factors (30 statements-33 statements), (2) evaluation factors (34 statements-20 statements), and (3) context factor (22 statements-15 statements). Thus, when interviewees were questioned in an unprompted structure, discussion of human factors dominated, whereas in the prompted section, interviewees had slightly more input about evaluation factors. It is important to note that in the first part of the interview, participants were not presented with the human, evaluation, or context factors as explanations for barriers to evaluation. Only later did we classify their responses using these categories. This analysis of the overlap between unprompted and prompted discussion clarifies that, even without any guidance, the participants identified the same explanations for barriers to evaluation implementation that were found in prior research on the associated area of evaluation utilization.

We began the interview with some questions about attitudes toward evaluation, in order to frame the interview and become more familiar with the background of the participants. We attempted to make a clear distinction between these introductory questions and the main part of the interview. However, we cannot rule out that the attitude questions influenced the nature of the responses to the subsequent questions. To investigate this effect, we would like to see a replication of this study that switches the order of the questions for some of the interviews. Another caveat is that some interviewees might have talked repeatedly about *their* issues throughout the interview, others might have raised certain points in the first part of the interview, and might not have repeated them later, even though, from a theoretical point of view, these issues would have *belonged* there.

During the course of the study it became apparent that barriers to *evaluation* implementation are often entangled with barriers to *program* implementation. Therefore, it is not surprising that a few participants talked more about obstacles they were facing in their daily outreach work than about barriers specific to evaluation. Although we could not use these statements as data in the current study, we were interested in hearing about these struggles in our roles as program evaluators.

## CONCLUSION

This study set out to investigate program staff perspectives on barriers to evaluation implementation. In the program context where we conducted this study, program staff inductively emphasized human factors over evaluation or context factors in explaining barriers to evaluation implementation. On the background of their recent experiences with program evaluation, staff highlighted the social competence of the evaluators, particularly their trust and relationship building competence. These findings are supported by utilization evaluation literature reflecting experiences accumulated in other settings. For example, Patton (1997, p. 34) summarized his experiences with barriers to evaluation as follows: "Many of the problems encountered by

evaluators, much of the resistance to evaluation, and many failures of use occur because of misunderstandings and communication problems.” Patton’s list of “threats to utility” (1997, p. 263) includes most of the points our interviewees perceived as barriers to effective implementation of evaluation: failure to focus the evaluation on intended use by intended users; poor stakeholder understanding of evaluation generally and the findings specifically; low user belief and trust in the evaluation process and findings; low evaluator credibility; and failure to keep stakeholders adequately informed and involved along the way (also see Cousins & Leithwood, 1986; Greene, 1988).

Thus, training in evaluation should incorporate skill building (both theoretically and practically) in the human factor area, besides focusing on methodological competence. Training should cover questions such as “How do you build trust in an evaluation-skeptical program staff?”; “How do you effectively communicate with your evaluation clients throughout the evaluation process?” or “How do you address staff’s negative attitudes toward evaluation?” How can evaluators improve their knowledge and skills in these respects? Especially if we conceptualize the field of program evaluation as inter-disciplinary (Rogers, 2001), it seems obvious that evaluation researchers (and practitioners) should more often engage in borrowing and adapting knowledge accumulated in other fields, for example, psychology, sociology, public policy, and management.

## REFERENCES

- Alkin, M. C. (1985). *A guide for evaluation decision makers*. Thousand Oaks, CA: Sage.
- Alkin, M. C. (in press). Evaluation theory and practice: Insights and new directions. In C. A. Christie (Ed.), *The practice-theory relationship. New directions for program evaluation*, 97. San Francisco, CA: Jossey-Bass.
- Alkin, M. C., Daillak, R., & White, P. (1979). *Using evaluations: Does evaluation make a difference?* Thousand Oaks, CA: Sage.
- Chen, H. T. (1994). Current trends and future directions in program evaluation. *Evaluation Practice*, 15, 229–238.
- Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
- Erickson, F. (1998). Qualitative research methods for science education. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education*. London: Kluwer.
- Greene, J. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, 12(2), 91–116.
- Henry, G. T., & Mark, M. (in press). Toward an agenda for research on evaluation. In C. A. Christie (Ed.), *The practice-theory relationship. New directions for program evaluation*, 97. San Francisco, CA: Jossey-Bass.
- Hofstetter, C. H., & Alkin, M. C. (2003). Evaluation use revisited. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation*. London: Kluwer.
- King, J. A., & Pechman, E. M. (1984). *The process of evaluation use in local school settings. Final report*. New Orleans Public Schools: Louisiana Department of Research and Evaluation.
- Leviton, L. A., & Hughes, E. F. X. (1981). Research of utilization of evaluations: A review and synthesis. *Evaluation Review*, 5(4), 525–548.
- Mark, M. M. (2001). Evaluation’s future: Furor, futile, or fertile? *American Journal of Evaluation*, 22(3), 457–479.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

- Monsen, L. (2002). School-based evaluation in Norway: Why is it so difficult to convince teachers of its usefulness? In D. Nevo (Ed.), *School-based evaluation: An international perspective* (pp. 73–88). Oxford: JAI Elsevier Science.
- Patton, M. Q. (1997). *Utilization-focused evaluation* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making*. Lexington, MA: Lexington Books.
- Rogers, P. J. (2001). The whole world is evaluating half-full glasses. *American Journal of Evaluation*, 22(3), 431–436.
- Taut, S. (2001). *Psychologische Erklärungsansätze für Widerstand gegen Programmevaluation* [Psychological explanations for resistance to program evaluation]. Unveröffentlichte Diplomarbeit [Unpublished master's thesis]. Psychologisches Institut, Universität zu Köln, Germany.
- Taut, S., & Alkin, M. C. (2002). *Barriers to program evaluation: UCLA outreach program staff perspectives* (Occ. Rep. No. 10). Los Angeles: University of California Los Angeles, Outreach Evaluation, Graduate School of Education & Information Studies; <http://www.gseis.ucla.edu/~outreachevaluation/reports.html>
- U.S. Department of Education, National Center for Education Statistics. (2001). *Paving the way to postsecondary education: K-12 intervention programs for underrepresented youth, NCES 2001-205*. Prepared by Patricia Gándara with the assistance of Deborah Bial for the National Postsecondary Education Cooperative Access Working Group. Washington, DC.