
**An Evaluation of an Evaluation:
CIRCE's Metaevaluation of the
Site Visits and Issue Papers of the
ATE Program Evaluation**

Christopher Migotsky and Robert Stake

CIRCE
University of Illinois at Urbana-Champaign

October 2001

Table of Contents

1. CIRCE Metaevaluation.....	1
2. ATE Program Evaluation Design.....	4
3. Evaluation of the ATE Program.....	6
4. Site Visits.....	8
5. Site Visit Reports.....	11
6. Technology Education	133
7. ATE Drivers	15
8. Issue Papers	167
9. Summary.....	19
10. References.....	200

1. CIRCE Metaevaluation¹

The Advanced Technological Education (ATE) program administered by the National Science Foundation (NSF) has been the focus of a 3-year evaluation grant awarded to The Evaluation Center at Western Michigan University (WMU).² Public Law 102-476 mandated the ATE program in October 1992. The Scientific and Advanced Technology Act of 1992 led to the creation of the national ATE program. The intention of the act was to expand the pool of skilled technicians in 20 advanced technology fields (such as semiconductor manufacturing, biotechnology, and information technology—among others). Planning grants were awarded in 1993, and the first full ATE grants were awarded in 1994. Overall, there have been over 400 grants since 1994 and there are currently over 150 active awardees across the country. The ATE program draws upon the resources and local educational missions of the nation's community colleges to deliver and implement most of the ATE grants. In this sense, it is unique among NSF programs.³

The Evaluation Center at Western Michigan received an NSF grant to evaluate the ATE program in April of 1999. The Evaluation Center's evaluation comes to an end in early 2002. As part of the NSF evaluation grant, the Western Michigan team's own evaluation efforts were scheduled to be metaevaluated. Robert Stake and Chris Migotsky, both from the University of Illinois, received the contract to conduct the metaevaluation of the site visit work performed by The Evaluation Center.⁴ This report is the result of the metaevaluation that spanned January to August of 2001. The following 14 questions guided our early metaevaluation activity, but as more information was gleaned, these questions were modified to reflect the emerging nature of our work. Short answers to these original questions are included, but do not form the full structure for our metaevaluative efforts.

1. Has the protocol/fieldwork been conducted in accordance with the contract?

Yes. It is our opinion that WMU fulfilled its major obligations in accordance with its grant.

2. Were the 13 sites chosen with an appropriate rationale?

Not entirely. The rationale was reasonable, but the final sites did not provide enough of an illustrative sample of ATE programs in general.

¹ Valuable interpretative and editorial assistance was provided by Wee-Haur Pek, graduate student, College of Education, University of Illinois at Urbana-Champaign.

² Throughout this report we will refer to the evaluation team as both "Western Michigan" and "The Evaluation Center." At times simply "WMU." We find these terms interchangeable, even though we realize The Evaluation Center is an independent unit within the infrastructure of Western Michigan University. For the sake of variety and readability we will take the liberty of treating these three terms as synonymous.

³ WMU has done a thorough job of describing and documenting the basic history and nature of the national ATE program in their Status Report I. We found that report consistent with the latest description from NSF in "Advanced technological education program solicitation" (NSF 01-52).

⁴ The University of Illinois metaevaluation did not require a deep look into the surveys administered by WMU. The metaevaluation contract stipulated that our focus would be on the site visit process, subsequent site reports, and first drafts of Issue Papers. During our document review, interviewing, and visits to ATE projects, we came across information that reflected upon the quality of the surveys. In this report we make some peripheral comments regarding the survey methodology and data, but our reflections should be interpreted with caution since our focus was never on the surveys. A separate metaevaluation of the surveys is currently being conducted.

3. Are team members individually competent and effective as a team?

Yes, that was a strength of the WMU evaluation, but matching a small team to broad project content left some issues unattended.

4. Were the hosts properly prepared to optimize use of site visit time?

Yes. WMU did an admirable job of preparing and organizing the site visitors and the sites themselves.

5. Has an effort been made to have the visit and report useful to the projects visited?

Yes, with mixed results. Projects generally found the reports useful, and at times extremely insightful, but ATE centers were less enthusiastic with the visits and subsequent reports.

6. Have site visitors been well prepared to carry out their responsibilities?

Yes, even possibly overloaded with pre-data. The 14 “critical questions” kept the visitors in sight of the “big picture.” The fully-structured protocols were so extensive and structured that they may have diminished the opportunity for the site visitors to provide the rounded, contextualized, interpretive commentary that later was needed for writing Issue Papers--even though they were explicitly asked to rely on their personal experience and expertise.

7. Is the protocol sufficiently oriented to actual activities, field operations, classrooms, and students?

It is sufficient, but perhaps too driver focused. Deeper issues that might aid understanding of local sites and their challenges seemed nudged aside for the sake of consistency and ease of aggregation.

8. Is the protocol sufficiently flexible, useful, and issue oriented?

Yes, if visitors are experienced (or trained) in protocol utilization.

9. Is the evaluation attendant to the state of the art in the relevant technical field?

Yes and no. It was considered in team staffing, but more attention might have been paid to risky practices or political compromises that might affect quality.

10. Are the reports useful, comprehensible, and issue oriented?

Yes, reports to date are comprehensible and usefully structured around drivers as issues. Issue Papers, still in draft form, are not fully developed and may require significant revision to achieve utility for sites and NSF audiences.

11. From team to team is there sufficient communication for coordination but not so much that it reduces the independence of the individual report?

There was little inter-team communication, but this was acceptable. The lack of extensive dialogue did not interfere with effective site work.

12. Is the site work cost effective? Does NSF get its money's worth from this part of the evaluation?

Yes, considering all the costs and benefits of face-to-face data collection. Several site visitors commented on the high value of these visits.

13. Has an appropriate effort been made to validate the description and merit of the local projects?

Yes, but not as fully as we feel was needed. Some unique qualities of the sites were not adequately described in the reports.

14. Can it be concluded that the brief visit provides sufficient data on the work at the site over the several years of their contract?

Yes, when combined with survey data. With WMU's budget, multiple visits were ruled out. A different kind of evaluation study was needed to make strong inferences about national program quality.

To address these questions and to pursue emerging issues, we employed the following primary data collection methods.

- Review of WMU correspondence with ATE sites and site visitors
- Review of site visit protocols, directions, and guidelines
- Review of 13 site visit reports
- Review of 9 Issue Papers (in draft stage)
- Interviews and visits with 3 WMU evaluators at The Evaluation Center
- Interview with ATE program officers
- Interviews with 7 WMU site visitors (some phone, some face-to-face)
- Visits to 3 ATE projects (interviewing PIs and “touring” the sites)
- Phone interviews with 6 ATE project directors/principal investigators
- Observation and participation at Issue Paper author meeting (July 2, 2001)

As we progressed with our metaevaluation, it was clear that larger issues were emerging that needed to be investigated. This report is designed around those larger issues, some of which relate closely to the original 14 questions, others are clearly emic entities that we had no intention of addressing, but felt obligated to discuss. Our metaevaluation style was heavily qualitative, fairly unstructured, and highly interpretive.⁵ This approach was in some contrast to that of our colleagues at Western Michigan who practice a more traditional, criterion-based evaluation.

⁵ We are aware that some authorities have reservations about a naturalistic approach to metaevaluation. See Smith (1990).

2. ATE Program Evaluation Design

The design of the ATE evaluation began shortly after receiving the NSF grant in April of 1999. After reviewing NSF and ATE documents and speaking with ATE program officials, the evaluators from Western Michigan conceived a plan centered on the following four evaluation questions:

- To what degree is the program achieving its goals?
- Is it making an impact and reaching the individuals and groups intended?
- How effective is it when it reaches its constituents?
- Are there ways the program can be significantly improved?

While these questions forged the focus for their evaluation, the WMU team remained open to new questions and opportunities. The initial phase of the evaluation was to develop surveys that could be administered, via the web, to all active ATE projects.⁶ Following the development and first administration of the surveys, their efforts were scheduled to shift to site visits at a small, but illustrative, sample of ATE projects. Through these two main data collections, along with meetings composed of PIs and the evaluation advisory group, the evaluators planned to deliver four primary products to NSF:

- Status Report 1 (description and overview of the ATE program)
- Status Report 2 (aggregate survey data—description and impact)
- Status Report 3 (final evaluation report)
- Site Visit Guidebook (guidebook for ATE sites and evaluators)

It was intended that these reports would have a strong organizational element based upon the ATE drivers.⁷ The drivers were developed from the ATE Program Solicitation with the assistance of the advisory panel and the ATE staff. The eight drivers were intended to both assist in the capture and representation of ATE project activities and to allow coordination and continuity of survey and site visit work. As the evaluation progressed, it became clear to the WMU team that a closer look at the drivers might prove useful to both ATE projects and NSF. After negotiations with NSF, it was determined that Status Report 3 would be replaced by a series of topical Issue Papers relevant to the ATE program (primarily the ATE “drivers”).⁸ An additional evaluation report was provided in 2001, comparing survey responses with those of the previous two years and including selected findings from the site visits.

⁶ As is the custom with concurrent WMU writings, we will use the term “project” for both ATE projects and ATE centers unless there is a specific reason to discriminate between the two.

⁷ The eight drivers were: collaboration, standards development, course/curriculum/materials development, professional development, program improvement, recruitment, student services, and sustainability/ transportability/ dissemination. The drivers were used to structure the annual surveys, the site visit protocols, and the site reports themselves.

⁸ We take up the issue of drivers in section 7 of this report. The Issue Papers are dealt with in section 8. While we agree that a switch to an alternative final report was a wise decision, we are not convinced that the current Issue Paper format digs deeply enough.

One foregone element in the overall design is the provision of a control group. There is no explicit comparison to see if the ATE-sponsored technological education programs are better than other technology education programs at community colleges. We realize a true control group does not exist. Devising one is sometimes attempted by evaluators--often without satisfactory results. We do not fault the WMU team for not having a control group, but given the strong expectations of providing evidence on program impact, a stronger design than used was needed. In a similar vein, we were discouraged that the evaluation design did not include a look at “old” ATE programs—those that have ended their grant term, suggesting a certain maturity of these sites. A close look at these matured ATE sites could have provided insight into sustainability issues and long-term student outcomes (skills obtained, satisfaction with program, job placement, etc). In either case, it is difficult to draw hard conclusions about national ATE program achievement.⁹ In subsequent communication with WMU team members, we realized that they struggled with these issues, along with NSF officers, and planned to address them in the next evaluation study, if a continuation grant is awarded.

It is clear that the evaluators from Western Michigan spent long periods of time conceptualizing and implementing their evaluation of the ATE program. While there may have been delays in developing and implementing the surveys and site visits, the WMU evaluators did deliver on these aspects of their evaluation design. Their strengths as a team lie in organization, planning, and management. Their talents and interests are nicely suited to structured quantitative evaluation designs that address program implementation and goal achievement. In the evaluation of the ATE program, they assembled a dedicated and talented group of educators, industry and business specialists, and evaluators to be part of the effort. Their advisory panel and evaluator group (i.e., advisory committees) were composed of leading professionals in the field of evaluation who not only supplied keen advice, but also participated in several phases of the evaluation itself.¹⁰ To date, the WMU evaluators have met their obligations to both the original NSF grant and their advisory committees. When NSF officials were asked what they thought of the Western Michigan team, ATE program officer Elizabeth Teles stated, “They [WMU] *really* tried to understand the ATE program. They came to meetings and visited sites. They were sensitive to project needs and our needs.” Sitting next to Teles, Gerhard Salinger, the co-lead program officer for the ATE program, concurred.

⁹ Evaluations have no pre-ordinate mandate to focus on program outcomes and achievements. It is fully within the scope of evaluators to negotiate for a focus on quality of processes and description of activities as a method of collecting representational data on the quality of a program. Having said this, The Evaluation Center did make promises to supply outcome/achievement data, but a search for independent empirical evidence of impact did not characterize this evaluation.

¹⁰ While the inclusion of advisory panel and evaluator group members on the site visit teams (and as Issue Paper authors) originally raised metaevaluation questions, we have since come to understand and appreciate the logic of this choice. We have no evidence to support the worry that panel and group members were co-opted by the WMU evaluators. Both WMU and NSF program officials supported this active participation of these individuals.

3. Evaluation of the ATE Program

The evaluators at Western Michigan did a masterful job of managing the entire project. Site visitors, project PIs, and NSF program personnel often stated that The Evaluation Center staff continually updated them on progress of the evaluation and supplied necessary documents and directions for carrying out their responsibilities for data collection and report writing. Project Director Susie Kelly of the Northwest Center for Sustainable Resources said, "They were very explicit. Their objectives were clear from the start. They gave us plenty of time and notification." This was no small achievement considering the large number of site visitors, advisory panel members, and actual ATE projects spread across the country. We metaevaluators too were recipients of the generous assistance provided by the Western Michigan evaluation team. Evidence of their attention to detail and program management can be seen at their web site devoted to the ATE evaluation [<http://ate.wmich.edu>]. There, a description of the national ATE program and their ATE evaluation exists along with survey data and biographical sketches of WMU staff and their advisory committees.¹¹

The Evaluation Center has produced several products as a result of their 3-year evaluation of the national ATE program. Chief among them are:

- Status Report 1 ("The Nature of the ATE Program")
- Status Report 2 ("Findings from a Survey of ATE Projects and Centers")
- Survey 2001 Report: ("The Status of ATE Projects and Centers")
- 13 Site Visit Reports (not available to NSF officials or the general public)
- 9 Issue Papers (all nine in draft form)
- Site Visit Guidebook (anticipated in Fall, 2001)
- Monitoring Report for 2000 and 2001 (for NSF use)

Each of these writings to date was comprehensive and significant in scope. None of them was assembled with haste. There was clear attention to detail and accuracy in each report. Several sites had complaints about the "representation of their work," but most felt the site reports captured the activities and achievements of their projects in a fair and objective manner.¹² Additionally, each report attempted to build on previous findings so a fuller picture of the national program might emerge.

The polished versions of the Issue Papers are intended to be this year's final statement on the ATE programs. Currently, in first draft form, these Papers do not deliver a synthesis of present and past data. Survey data were not fully incorporated into the first drafts. The literature reviews, while insightful, were not sufficiently linked to the site visit reports and ATE issues. Likewise, fieldwork from the 13 sites was not integrated into the Issue Papers as completely as expected by us metaevaluators and the NSF program officers. Without comprehensive revision of these Papers, the NSF staff as well as current and future ATE projects, may fail to appreciate and learn from the Western Michigan efforts.¹³

¹¹ At the time of this writing, information about this metaevaluation is absent from the WMU web site.

¹² More detailed discussion and analysis of the site visit reports can be found in section 5 of this report.

¹³ See section 8 of this report for a more thorough look at the Issue Papers.

While the evaluation team was vigorous in its attention to information exchanges, we were not satisfied with two choices the team made in designing and carrying out their evaluation. One important role for evaluators is to question program manager decisions and policies. There seemed to be too little attention paid to ATE program and policy decisions made by NSF. The evaluation team moved forward to evaluate program activities and achievements. Why no evaluation of the ATE program goals as expected in CIPP¹⁴ evaluations? Nor of selection criteria for awarding grants? Nor as to the choice of what areas count as “advanced technical fields”? A critique of ATE policies and procedures got minimal priority. While the entire design should not be engaged with these higher order administrative questions, time spent investigating them would have been worthwhile for both NSF and the ATE program.

The second choice troubling us was the decision to commit deeply to fully-structured protocols for the site visits. They were referred to as Guidelines but their length and detail largely determined the topics and style to be used. Even though site visitors were asked to rely on personal experience and expertise, we think the protocols diminished the opportunity for the site visitors to provide the rounded, contextualized, interpretive commentary that later was needed for writing Issue Papers. The protocols were to be a universal framework for data collection at all sites, structuring the data, making the protocols more comprehensive and easily assimilative across the sites visited and to facilitate aggregation and comparisons. Probably nothing seemed more obvious at the time. But what this did, in our eyes, was to dedicate the resources of the site visits to the matter of project compliance and accomplishment. The broad history of site visits is fixed upon the purpose of compliance with contracts and regulations. And the history of evaluation is often fixed upon the determination of goal fulfillment. And we have seen those purposes prominent in the protocols and the site visit reports. Another purpose of evaluation is thus squeezed: the interpretation of the driver phenomena in a context of stakeholder values, with implication for policy-making and practice. It may not have been foreseeable at the time the protocols were being developed that this interpretive, constructive, policy side of evaluation would be greatly needed, but it became so.

Overall, did NSF “get its money’s worth” with the WMU evaluation grant? We did not look at budgets or expenditures of the WMU evaluation, but we do believe that NSF received a worthwhile and valuable evaluation of the ATE program. The Evaluation Center performed well. They implemented their plan, made adjustments when needed, and by-and-large, delivered on their promises. While richer description and more exploration of problem areas and issues would likely provide readers with greater vicarious experiences (and thereby allow them to understand the unique challenges faced by the ATE projects), the WMU staff of evaluators and site visitors did a commendable assessment of the quality of the ATE program around the eight drivers.

¹⁴ Context, Input, Process and Product Evaluation, as developed by Stufflebeam. See Daniel L. Stufflebeam, George F. Madaus, and Thomas Kellaghan, (2000). *Evaluation Models: Viewpoints on educational and human services evaluation*. Second edition. Boston: Kluwer.

4. Site Visits

The Western Michigan evaluation team visited 13 ATE sites (ten projects and three centers) as part of their overall evaluation effort. These visits were dedicated to finding out what was actually happening at sites across the country, to “illuminate what the ATE program is doing.”¹⁵ It was an opportunity to dig deep, ask program issue questions, fill in the contextual background and gather personal perspectives. Annual survey data had been gathered and interpretations of quality were beginning to emerge, but the evaluation needed first-hand descriptions of project activities and validation of survey results. As evaluators, they knew triangulation of results was vital.

The selection of sites to be visited was an aspect we questioned.¹⁶ The 13 sites were selected by WMU from a list of twenty supplied by ATE project officers using WMU criteria¹⁷. The ATE people knew the sites longer and better than the WMU people and provided 20 that could inform the evaluators effectively. But the sample of 13 could have been more illuminative in terms of geographical diversity, success rate, and grant size. Mainly we would like to have seen more from sites which were struggling with ATE challenges and constraints.

After speaking with ATE project external evaluators and principal investigators, we were not convinced that the struggling sites and small grant sites were well represented¹⁸. Some of the visited projects had problems, of course, but by and large, the group was doing quite well. In a national program of this scope and nature, we expected a larger “failure rate.” At the Issue Paper meeting in Washington, DC, July 2, 2001, several members of the evaluator group commented that for experimental technology education programs we might expect 50% or so to be “duds.” These experimental sites were supposed to be pushing the envelope and looking into uncharted waters, so one might reasonably expect trouble. If no ATE projects were struggling, one might claim that ATE was too conservative in their grant-awarding process and not providing enough support to the more innovative, but possibly riskier, proposals. While no sample of 13 can be expected to generalize well to a larger population, we think a more diverse sample would have better portrayed the ATE program.

Once sites were selected, WMU turned attention to staffing the visiting teams. They developed a matrix of the 13 project topical areas and potential site visitors’ qualifications. In detail they matched expertise, interest, and experience of the site visitors to the domains of the projects. They especially tried to balance each team with both academic and industry personnel

¹⁵ From a WMU memo to site visitors entitled “Goals of the ATE Evaluation Site Visit.” We saw merit in enriching the database with site visits but we believe that the annual survey could contribute more too. For example, situational items could be used to gather perceptions of relevant business, engineering, career development and ethical issues. The same questions need not be asked every site and every year.

¹⁶ All three centers chosen were from the Western part of the US. The ten projects selected did not include one in the Midwest, Midsouth and Mountain states. Maps from the “ATE Program: 1999 Awards and Activities” NSF 00-112, p.28-29) show a much greater diversity of site locations than were ultimately selected for the 13 visits.

¹⁷ With one exception, all sites invited to participate in the evaluation agreed to do so. The one had an unexpected PI turnover and opted out. Another project from the site list replaced it.

¹⁸ Of the ten projects visited, the median grant award was \$573K, high of \$1.7M and low of \$232K. NSF used \$500,000 as the dividing line between small and large project grants. With this cut, WMU sampled both large and small projects sufficiently, but we would have liked to see proportionately more small projects (those closer to 200-300 thousand) in the case study sample.

so their combined knowledge would fit well with the particular sites.¹⁹ According to the criteria used, the teams were well formed and had the ability to understand and reflect upon project activities and goals. We heard minor concerns from projects that had visitors with no industry or “practical” experience, but more typical were comments from PIs along the lines of “We really valued Al’s experience in industry” or “Lester understood the community college scene” or more generally “The visitors gave us new perspectives...some educational, others business related...it was a good combination of expertise.” Within the given resource and time constraints, WMU did a nice job of matching site members to project needs. To increase team expertise, without undue cost, it might have been valuable to recruit local content experts to assist with the team visits.

From all indications, the site visit teams were thoroughly prepared for their visits. Evaluation managers in Kalamazoo provided them with documentation about each site, previous evaluation data if available, protocols for interviewing and observing, and guidelines and directions for carrying out the visit.²⁰ Scheduling was managed flawlessly by WMU and both PIs and site visitors commended the WMU staff on “paying attention to all the details.” Each site developed an agenda for the visits: with the goals and preferences supplied by WMU. Negotiations took place between WMU and projects and final agendas were forwarded to site team members and sites.²¹ One of the center directors said, “It was a pretty quick turnaround, things happened quickly.” As was characteristic of WMU’s work style, there were no surprises for site visitors or the project staff. Indeed, the site visitors were not just prepared, but gracious as well, “They were just wonderful visitors” commented a project PI. She continued, “They called in advance to plan things out and get to know us. It was nice. I was fearful and a little anxious at first, but it was very pleasant. I would welcome them back anytime. They were just fabulous with our faculty and students.”

Visits lasted approximately one day for projects and two days for centers. To gain a basic understanding of compliance and attainment, project activities and priorities, this was an adequate time on site. To dig deeply and probe for issues, this time was insufficient. When team members and PIs were asked whether the visits were “too long/too short/about right”—almost everyone agreed that a little longer would have been ideal, but the duration was probably about right given the circumstances. Sites were already burdened with other evaluative tasks and, for some, these visits were “above and beyond” expectations. Gloria Rogers, a site visitor, felt they did about as well as could be expected in a day. Rogers specifically commented, “There is a point of diminishing returns. There are constraints. I think we got about 80% in the time we were allowed. Another day and maybe we would have gotten that other 20%. I’m not sure that would be worth it.”

In phone interviews with project administrators, similar comments were heard. Maricopa Advanced Technology Education Center Director Michael Lesiecki used almost exactly the same

¹⁹ Teams usually consisted of two members for a project visit and three for a center visit.

²⁰ If there was a concern, it was for information overload. Two site team members reported feeling overwhelmed with pre-visit information. They felt “swamped” before and during the visits. The 14 “Critical Questions” given to site visitors helped somewhat in allaying this feeling.

²¹ This collaborative working arrangement was the *modus operandi* for WMU. Nothing was forced on team members or sites. Feathers were not ruffled. Perhaps to a fault, the Evaluation Center staff and their site visitors accommodated almost all project requests. They were respectful guests.

words Gloria Rogers used. An exception to this satisfaction with site duration came from center personnel. Two of the three centers had concerns about the length of the stay—they claimed it was not sufficient for capturing the extent and quality of their activities. As one principal investigator stated, “We have very complex and spread-out activities. I don’t think they spent enough time with us to really see our work. What they reported on did not detail our center’s achievements. They did not fully represent us.” This PI went on to explain that the visitors “worked hard” but without more people or more time she said she did not believe it was possible for her Center’s work to be documented.

We as metaevaluators realized the visitors did not have on-site time to “digest” what they were uncovering. Wayne Welch complained of “not having time to let things percolate...to think.” We too worried about how much any site team could learn about a project or center in a day or two in the field.²² Could team members dig deeply? Did they take the opportunity after reflection back in their offices to call or email back for clarification? Not many did. Their draft reports were due in two weeks. Protocols, deadlines, and agendas continually pushed them to move forward. Common structure and goals were necessary for cross-site analysis and continuity across visiting teams. We believe the visits did deliver useful and insightful data about individual projects and the national ATE program as a whole, but we still ponder the missed opportunities and potentially simplified site reports.²³

²² It should be noted that a “day’s effort” typically meant 7 AM to 7 PM or longer. Site visitors worked hard from morning through evening trying to comprehend the sweep of project efforts.

²³ It might have been useful to have attached, with each site report, a cover letter (for WMU use only) indicating potential trouble spots, issues, concerns, that authors suspected but did not investigate nor include in the site report. This would protect sites but allow WMU to contemplate program challenges.

5. Site Visit Reports

The 13 site visit reports were intended for the WMU staff, their collaborators, and the sites themselves. It was never the intention they be forwarded to NSF, the ATE officers or the general public. This promise of confidentiality was kept. These reports were written to speak of compliance and meritorious accomplishment with focus on the drivers. This emphasis worked toward a necessary cohesion, but also constrained author contributions to policy questions and as to topical choices and freedom of expression. Anecdotes, pictures, and stories were not the purview of these reports, although encouraged in the written guidelines to visitors. Description of project activities in relation to the relevant drivers was the mandate. The editorial policy was described by Nanette Keiser, ATE evaluation project manager at WMU:

There were two types of edit processes. The site visitors, at a meeting in October 2000, approved these two written processes. First, the team leader assembled and edited the "draft" site visit report from parts written by fellow team members, along with parts s/he wrote (usually the project description). This was forwarded on to WMU staff members within 2 weeks of the site visit. WMU staff then forwarded the report on to the site for comments and highlighting of any misperceptions and inaccuracies. When changes were received from the sites, WMU staff made a determination if each site's suggested changes were substantive in nature. If they were not, WMU staff made the changes, added a cover page, and did grammatical revisions as needed. If the changes were substantive, the report was forwarded to the team leader, and if s/he chose to do so, team members were re-engaged to discuss and finalize these changes. Again, the team leader made the final determination regarding these changes.

Confirmation by member checking did occur by passing the draft reports through WMU to the sites. Some site report authors did not have ample opportunity to make post-check changes, apparently because of limited turnaround time.

In interviews, project staffs often complimented WMU on their report, only at times quibbling with minor interpretational issues. For projects, the facts were not in contention. The resulting reports were accurate, dry and informative.²⁴ Principal investigator Kathleen Alfano of CREATE in the Santa Barbara area said, "They did well with the report. It gave us a sense of validation. Now we know we are doing well compared to others." Centers had larger, persistent, concerns. Two of the three center directors said their site reports did not reflect their work. One center PI said, "The report was too critical and just not on target for some things. They [the site visitors] weren't on site long enough to get a full appreciation for what we are doing across the country. We never got a chance to shine."²⁵ "We sent our objections to WMU, but not much changed in the final report." Is this simply "sour grapes" or a legitimate critique? Another center director commented positively on the WMU member checking process, "They sent a draft for us to respond to. I really liked that. We sent additional information that they incorporated into the final report. We corrected a number of minor misconceptions. The report turned out fine." All projects and centers had the opportunity to respond to first draft site reports. At all

²⁴ More "engaging" site reports (intended more as promotional tools than evaluative reports) can be read in two American Association of Community Colleges books: *Students at the learning edge: Advanced technological education programs at community colleges* by Ashlock and Wright and *The learning edge: Advanced technological education programs at community colleges* by Mahoney and Barnett.

²⁵ This same PI claimed that the site visitors were too focused on the drivers, "Sometimes those just didn't fit our work very well, but they kept returning to the drivers."

Sites but two centers it was felt the reports adequately reflected their goals, activities, and achievements.

Why did centers and projects react differently to the reports? Here is one hypothesis. In addition to being larger, more highly funded, more diffuse, and more complex than the projects, centers were designed to be "centers of excellence." These "centers of excellence" are regional or national clearinghouses for their content area. They are experts and resource people for their disciplines. They help solve problems. They offer professional development and standards. They recruit industry and academic partners to produce and deliver needed materials and services. They are not typically the ones seeking assistance, but rather providing assistance. They play an information services role.²⁶ And so do the site visitors. It is possible that in some way the visitors were seen as invading their turf. The site visitors worked from academic models and theories in this ATE business-oriented setting. One center principal investigator said, "It was as if they had a different model or set of criteria for success. This was problematic and it showed in the final site report." The projects, on the other hand, without this expertise, struggling in certain ways, were disposed to appreciate the observations of the WMU team. Speaking of her site visitors, Oklahoma project director Anita Watkins²⁷ commented, "They offered help and advice. They pointed us toward some valuable information and places to go for more grants and assistance. They even offered to give us continued support if needed. They were very open and helpful." We concluded that there was a sufficient difference here to call for more differentiated protocols for projects and centers in the future.

If the site visit reports were to inform a collection of policy-based Issue Papers and to "capture the richness of the projects, the excitement and the human side of the program,"²⁸ then they were less than a success. But the reports were intended to confirm the good works of projects and to validate and illuminate the national survey data, as such they were a success. We would have more highly commended Western Michigan had their site visit reports achieved all four goals. But maybe they achieved what they could have. We were not in the trenches day-in and day-out like the WMU team. They had other pressures, other pulls, and too little time. The reports were helpful for Issue Paper authors and managers at WMU who needed to interpret survey responses from across the nation, but not helpful enough for policy making. Evaluators cannot perfectly capture quality and utility, even if contractually obligated to do so. In this case, WMU made no promises to submit "case study like" reports with thick description and interpretive implications for policy. Yet as the fiscal year closed, WMU increasingly placed the Issue Papers high in priority. Gradually the evaluation emphasis switched from providing a status report to formatively evaluating the drivers. That process happened too slowly to shape the site visit reports.

²⁶ See Judith W. Little and Milbrey W. McLaughlin (1993). *Teachers' work: Individuals, colleagues, and contexts*. NY: Teachers College Press.

²⁷ In Okmulgee, Watkins directed the project entitled, Preparing High Performance Technicians in Distinctive Manufacturing.

²⁸ This notion of rich, descriptive data surfaced at an advisory panel meeting in October, 2000. There was no final agreement that this was a major goal of the site visits. But the idea is intriguing and worthy of mention.

6. Technology Education

The ATE Technology Education program was created to support community colleges across the country providing technology training to students already employed or hoping to be employed in local industry or agency programs. Examples of new technology fields served: Environmental sciences, Water quality maintenance and restoration, Toxic waste, Microelectronics, Semiconductors, Computer science, Electronic networking, Global positioning, Health sciences, Biology, and Aerospace. It was up to the individual college to recognize the need and opportunity for developing new programs in subject matter fields needing special courses. Of course, it often was an employer, public official, or field representative who put the idea in the heads of campus people.

One of the broad questions WMU site visitors raised had to do with the readiness of the college to provide this training, the competence in the technology held by the college faculty members providing the training, stock of state-of-the-art instruments, and acquaintance with state-of-the-art technology. The social dynamics of the site visit situation were such that in order to get cooperation from their hosts and good answers to their questions, the site visitors could not press aggressively on very many topics. Each of these questions risked offending the interviewee. But visiting team members Norman Gold, Ann Igoe, Albert Schwabenbauer, and Gloria Rogers each spoke to us at length indicating their attention to state-of-the-art questions. Igoe said that projects "with active high-tech partners were doing good things and were *high-end* in terms of technology, but some others were not *high-end*, even though they thought they were." Senior associate (for the evaluation) Frances Lawrenz said, "It is possible that too little attention was paid to technology, maybe it should have been a driver."

The rationale for training paralleled the rationale of many technical programs on a community college campus: that the employers knew what trainees need, sometimes providing training specifications and even instructional materials. It is supposed that there should be an interaction among trainers noting what was being covered elsewhere or should be covered according to new industry and technology product announcements, but the community service motive of the colleges probably modulated that role. So the rationale for the WMU evaluation was to learn from the employers and other partners whether or not the topical content of the training was up-to-date.²⁹

The site visitors were alerted to indications that there might be some mismatch between training and industry needs but the protocols did not ask for probing. In our metaevaluation questioning of project directors and site visitors, this issue was not identified as a problem. Actually, our questioning pursued the matter differently, asking if the site visit team was staffed with sufficient technological expertise to recognize whether or not the training content might be out-of-date or in some other way faulty.

It was clear that staffing of the WMU site visit teams was oriented to the education side of technology education, the side represented by the drivers. We were told at the outset that it was known that some of the projects were inexperienced or poorly informed about the technical sides of professional development, course development, student services, etc. And we believe the

²⁹ Again, we stress the potential usefulness of having local content experts in the site visit teams.

WMU staff recognized that educational services evaluation was the strength of their proposal. So they assigned people to their site visiting teams giving high priority to education and lower priority to technology. There clearly were some experts on particular technologies in some teams, and matched well with the sites visited, but not all the teams were so staffed.

Wayne Welch and Arlen Gullickson were the site team for the project, Advancing Careers in Technology and Science, headquartered at Collin County Community College north of Dallas. The project was primarily a human services project, facilitating technical careers starting in high school and into life-long learning. Although the affiliated colleges offered training in Cisco Systems (internet hardware and software), computer assisted drafting, and telecommunications, this project was not responsible for it, and thus was not directly concerned with technology content of courses. Asked about the match of the site team to the project, Welch said,

There were just two of us, both educational evaluators, and we were there just a day, not conducive to deep thinking. It was a day originally booked full of presentations, so we had to negotiate to ask the questions for which we needed answers. So we could have used a third person and it probably would have helped if that person had been familiar with technology education as offered by these colleges. Some issues probably could have been probed better. But I think we got what we went for, an assessment of the effectiveness of this career facilitation program, as input, of course, for evaluating the national program.

The WMU staff had been urged to use its evaluation inquiries to make the vitality of the program apparent, to capture "the richness of the projects, the excitement and the human side of the program." The richness would include some interaction with cutting-edge industries struggling to avoid the economic downturn. This drama probably was at least alluded to during some of the site visits but it did not get much mention in the site visit reports, possibly because the drivers more or less invited a more prosaic write-up.

In summary, the WMU evaluation followed its plan to concentrate on the procedural side of the training. Responsibility to assure that training content was solid and up-to-date was not emphasized, but no signs emerged to suggest it should be. Still, the extent to which campus faculty members were checking out industry and employer perceptions of the quality of the training received less attention from the WMU team than it deserved.

7. ATE Drivers

The collection of "drivers" originated with ATE but expanded in discussions between the WMU evaluation staff and NSF people as a conceptual structure attuned to ATE aims and questions about the work of the funded projects and centers. The questions were said to regard: Collaboration; Standards development; Course, curriculum and/or materials development; Professional development; Program improvement; Recruitment; Student services; and Sustainability/Transportability/Dissemination.

There was a clear connection with ATE program objectives that were, according to Status Report 1:

1. Develop model instructional projects in advanced technology fields.
2. Provide professional development to faculty and instructors in advanced technology fields.
3. Establish innovative partnership agreements.
4. Acquire and implement state-of-the-art instrumentation.
5. Develop and disseminate instructional materials.

Metaevaluation colleague Wee-Haur Pek pointed out that several drivers (professional development, collaboration, materials development, and sustainability/transportability/dissemination) were directly drawn from Objectives 2, 3, and 5. The other four drivers (standards development, program improvement, recruitment and student services) seemed all to be steps within Objective 1. The WMU team had difficulty fitting Objective 4 into the protocols possibly because "only 5% of the project sites specifically mentioned instrumentation as a project objective." WMU recommended that "additional attention and support be given to this objective."³⁰

The list of titles makes it apparent that the ATE work was seen as training, perhaps hinting that the work most in need of WMU evaluation would concentrate on the problems and accomplishments of training. With training in mind, the drivers were set forth as the conceptual structure for the site visits. Questions on other topics would be raised and the projects and centers were expected to draw attention to what they were accomplishing. The evaluation visit was to tease out information on problems and accomplishments, but the important problems and accomplishments would be those associated with the eight drivers.

The drivers were seen to be of generalizable value for project quality control and support was given to the prospects of providing a Site Visit Guidebook for ATE projects and possibly other educational projects funded by NSF. Development of the drivers is a good example of the commendable client orientation of the evaluation team and of the collaborative aura between WMU and NSF.

These drivers are respected constructs and divisions of labor within an Instructional Technology tradition. In recent years that field has been dominated by the special questions of distance learning and computer-based instruction but the driver concepts remain important. Thus the drivers have an immediacy of potential use for the ATE contractors and a research base and

³⁰ We note that the team attempted to address this objective within selected drivers such as curriculum development and program development.

literature as well. It is our view that much instructional technology research is well below the quality of studies NSF has funded in curriculum development and educational research over the years, and currently is handicapped by “best practices” thinking and hortatory reviews.³¹ Still it draws deeply upon professional experience and deserves at least some of the respect practitioners give it. And of course, even if the research were poor, the drivers as areas of concern are not therefore less legitimate.

The precaution we were led to is that center agendas, and site visits, and the evaluation as a whole, might be drawn by the practice and literature of this field into admonitions (of the kind: "The trainer needs to make eye contact with the trainees." and "You have to walk before you run."), not recognizing that there are many circumstances in which such admonitions may be counterproductive. So we looked for this mindset in the site visit work, and found some, but considered it not a problem.

The idea of having the drivers as conceptual structure for organizing the evaluation is good. We feel it is an improvement on a project goals orientation because, we believe, it more quickly gets at the difficulties the projects may be coping with. Our own preference for structuring the evaluation around issues (drawing observations and reporting more deeply about local context and interpretation) would have made distillation of findings across 100+ sites (and even across the 13 sites visited) more difficult. The site visit teams found the driver framework useful for organizing their visits (and later the Issue Papers). Drivers could have dragged the work down. We asked the team members if the drivers drew attention away from other matters that needed lots of it, and they responded that, while good focus came from the drivers, they had no difficulty moving to other matters as needed. Among other topics hypothetically neglected (according to us) because of the heavy orientation to drivers were: local technology history, involvement of the central office, commingled funding, politics, community relations, staff turnover, students lacking study skills, and library resources. We commented on the fact that there was less description of the human story, the dramatics in the site reports, than some readers wanted, and wondered if the drivers had kept them from gathering such "oral history." None of the site visitors asked this question indicated that the drivers had kept them from providing a more dramatic account of project life, noting the brief time they had to get acquainted with it.

Most project and center staff members supported the drivers as being (as they understood it) what the evaluation was supposed to be about. One project external evaluator, however, said that the drivers were “blindings,” sharing our worry that the driver structure might have disadvantages.³² He/she was clearly in the minority. Referring to the site visitors and their protocols, John Duryea, project manager at Tech-4 in Orlando, said, "They were not restrictive at all. We were able to tell our story. They made sure to come back to certain questions at the end of the day, but otherwise they were very accommodating of our needs." The drivers were not being trivialized.

³¹ "Best practices" are the things that successful projects do. So it is also true that they are the things that projects with the most resources do. Many projects do not have those resources. A technology should not be based on what requires extra resources. It is for this reason that research and evaluation needs to record what marginal and struggling projects are doing, for from them can come insights about successful coping.

³² Here, as a metaevaluation maneuver, we raised the possibility of something bad happening, to see if it was refuted, and it was, particularly by John Duryea.

8. Issue Papers³³

Our metaevaluation assignment included a consideration of the use of the ATE “Issue Papers” as a summary statement of this year's evaluation of the program, plus a reading of the first drafts of those Papers and a hearing of discussion by the authors and critical friends. In the words Wayne Welch used in his Issue Paper on Advisory Committees,

An issue is defined here as a topic or situation that requires further consideration to make a decision or reach a conclusion. Five of the issues addressed by the ATE evaluation come from the core program objectives mandated by Congress. These issues include dissemination, materials development, collaboration, professional development, and program improvement.

Four other issues evolved from discussions and concerns expressed by NSF program officers and the ATE evaluation staff. They are recruitment and retention, sustainability, evaluation and advisory committees.

These issues are explored in a set of nine Papers, called Issue Papers. The issue Papers were selected as a way to interpret and report the findings from the WMU studies. The Papers are based upon information contained in two status reports and the reports of site visits to a sample of 13 projects and centers. Additional context was obtained from the program guidelines, NSF staff, literature reviews and the involvement of the authors in various aspects of the ATE program. This included, for example, serving as a project evaluator or as a member of a site visit team.

The Papers were written by different authors, each with their unique background and experience. As such, they vary in tone, content, length, and how much judgment and recommendations they contain. However, all are designed to provide information to help make decisions or reach conclusions about the issues in the ATE context.

The Issue Papers were expected to be useful to those who formulate policy but also useful to those at sites where policy changes will be most felt.

The Issue Paper topics were assigned by the WMU evaluators to provide a synthesis of happenings at the 113 projects and centers as sampled by the annual survey and the site visits in 2000. Although authors were not told to restrain themselves from using previous data and experience, or to avoid gathering their own, the expectation was that there would be adequate findings in the surveys and site visit reports for all that needed be said at this time about the drivers. It was our understanding that these Issue Papers would deal with generalizations and illustrations as needed to back up a Site Visit Guidebook, a final product of the evaluation.

The contract for the evaluation spoke of a Guidebook but did not call for Issue Papers. The initial data gathering design did not orient much to advisement of policy makers and practitioners on such organizational matters as professional development, course development and sustainability. The aim of the evaluation team was to prepare a survey and a set of face-to-face interview and observation protocols that recorded the activities of the projects, with emphasis on attainment, colorings of merit, and maintenance of goal pursuit. This is the traditional expectation of an all-project evaluative survey and a set of evaluative site visits.

³³ Issue Papers were originally called White Papers in the WMU evaluation.

Because the purpose of the site visits was to use a sample of visitation reports to convey the nature of the program as a whole, the evaluators struggled with a difficult sampling process. This is not a process unfamiliar to any evaluators, but the aggregation process is one that has not been routinized, validated, nor even much debated. How to put the stories of 13 projects together to make a story of a program having other projects. The process is made difficult when the number of sites is large, too large not only for all to be visited but even too large for anyone to have acquaintance with a large portion. The inclination of the WMU team was, as most professional colleagues would urge, to structure the site visit input so that it could aggregate in a statistical sense, even when the data were categorical. But the site visitors needed to be drawn to thinking about what has implication for the program as a whole, even if it is not of high importance at the project level. The protocols raised questions about the drivers that would most easily aggregate, most easily tell what is common about the projects, but raised few questions about what is fluid, unique, personal, and dynamic at the sites visited. The site visitors drew from different disciplines and experience and were inclined to look at both what was common and unique. They took the commitment to drivers seriously, agreeing on-site as well as off, that the improvement of the training side of these projects was of high priority. They found other matters of interest but the protocols were persistent in returning them to the common questions. So the site visit reports were endorsements of the projects' devotion to the drivers but not richly informative of how people coped with problems in order to accomplish what they did. Because of this long preparation for sustaining and enriching the survey reports, the site visits were not ideally suited to the preparation of Issue Papers.

There was perhaps no way the evaluation team could have anticipated the late-term need for some kind of policy testing in the natural setting. That is to say, what happens at individual projects unintentionally and often unconsciously provides a set of covariations not unlike a highly controlled policy study. For example, with the driver called Collaboration, there is a question as to how much responsibility the faculty should take in deciding whether or not the local partners and employers are defining workplace competency in ways that best serve the trainees. Policy research could be set up to look into it, perhaps it already has been. But some things that faculty members at to-be-visited sites already are doing vary as to questioning training specifications externally generated. Especially if the detail and tone of the protocol are supportive, this variation could be recognized and inquired about by a site visitor. One might not find any of the ten most pressing policy questions to be addressed in this fashion, but that disposition to helping prepare Issue Papers and a Guidebook is recognized as valuable, now that the Issue Papers are undergoing revision.

The evaluation team could not have recognized this two years ago, not even a year ago. There were no Issue Papers. So it would be a mistake to say that the survey and site visit protocols should have been designed with this end in mind. In retrospect, it appears that a stance by the designers to hold down the priority on project compliance and attainment might have made it easier to pursue issue interpretation and policy development. A better way to prepare a set of Issue Papers probably would be to send each author to a site or two for a few short visits. Concentrating on the activity encountered, the problems and personnel at hand, preparation for probing the annual survey data and for writing his/her Issue Paper might have been improved. We find it interesting to make these speculations, but we do not presume that the original evaluation design was faulty because it did not have a crystal ball.

9. Summary

The site visit part of the evaluation of the ATE program coordinated by The Evaluation Center at Western Michigan University fully met the standards for evaluation of educational programs.³⁴ The protocols for data collection, the preparation of site visitors, the conduct of the visits, and the site visit reports complied with requirement and attended to the needs of stakeholders. The evaluation team fulfilled their contractual obligations. They oriented to the developing nature of the ATE program. For the most important issues, the site visit teams were appropriately staffed. The evaluators were considerate of the circumstances of the projects and centers. They provided information in accordance with mutual understandings with the ATE managers. It was recognized that minor improvements in design and operation had been possible, more apparent after the fact than before, but the process of evaluation was fully compliant with obligations.

Good information was obtained from the 13 sites visited. The information gathered followed the intended protocols reasonably closely. By plan, the information centered on a set of "drivers," organizational constructs important for training, such as: collaboration with partners, professional development, and project sustainability. These were examined at each site. As intended, an effort was made to extrapolate to the program as a whole, and the results of that difficult generalization will not be known until later this year.

Almost all site visitors, project principal investigators, and ATE program personnel stated that The Evaluation Center staff acted with competence and consideration. They managed the project very well, particularly in matters of communication. Management of the project was no small achievement considering the large number of site visitors, advisory committee members, and actual ATE projects spread across the country. The advisory committees and the lead evaluators conceptualized the study at the outset in a traditional consideration of project compliance and accomplishment, later moving more into an orientation to study the drivers for policy purposes. Both of these undertakings were well thought out, but time and other resources did not allow them to move very far with the second purpose.

From survey data, The Evaluation Center produced annual status reports, then added site visit reports as input to topical papers called "Issue Papers" dealing with broad conceptualizations of the drivers as pertinent to ATE policy and implications for practice. Those Issue Papers are scheduled for publication at midwinter. The input writings have been comprehensive and significant in scope, with clear attention to detail and accuracy. Several sites complained about the "representation of their project work," but most site people interviewed said the site reports captured the activities and achievements of their projects in a fair and objective manner.³⁵

While the evaluation team was vigorous in its attention to information exchanges, our metaevaluation drew attention to two shortcomings. One was the little attention paid to ATE program history and early policy decisions. The second was the constraint put upon fieldwork

³⁴ The Joint Standards include standards for utility, feasibility, propriety and accuracy. Joint Committee on Standards for Educational Evaluation (1994) *The Program Evaluation Standards*.

³⁵ More detailed discussion and analysis of the site visit reports can be found in section 5 of our report.

by the volume and targeted nature of the protocols. That constraint appeared to limit inquiry into the uniqueness of individual sites and interpretations for deeper understanding of the drivers.

Among other problems examined were: sampling of sites, objections to site visits at centers, lack of time for digesting site data, scarcity of contextual data, too little differentiation between projects and centers, and low priority to policy questions.

Overall, the WMU evaluation team did a highly commendable job of setting up a site visit operation and integrating the reports with the annual survey and increasing interest in the drivers. The Evaluation Center performed well within their design. They implemented their plan, made adjustments when needed, and delivered on their promises. They could have done even better had they gathered richer descriptions and more explored project problems and issues. But all in all, the Western Michigan University evaluators and site visitors conducted a commendable assessment of the quality of the ATE program around the eight drivers.

10. References

- Ashlock, T., & Wright, S. (2001). *Students at the learning edge: Advanced technological education programs at community colleges*. Washington, DC: Community College Press.
- The Evaluation Center. (2000). *The nature of the ATE program* (Status Report I). Kalamazoo, MI: Author.
- Gullickson, A., Lawrenz, F., & Keiser, N. (2000). *Findings from a survey of ATE projects and centers* (Status Report II). Kalamazoo, MI: Western Michigan University, The Evaluation Center.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards*, (2nd ed.). Thousand Oaks, CA: Sage.
- Little, J. W., & McLaughlin, M. W. (1993). *Teachers' work: Individuals, colleagues, and contexts*. New York City: Teachers College Press.
- Mahoney, J. R., & Barnett, L. (Eds.). (2000). *The learning edge: Advanced technological education programs at community colleges*. Washington, DC: Community College Press.
- National Science Foundation. (2001). *Advanced technological education program solicitation*. (Document NSF 01-52). Washington, DC: Author.
- National Science Foundation. (2000). *Advanced technological education program: 1999 awards and activities*. (Document NSF 00-112). Washington, DC: Author.
- Smith, N. L. (1990). Cautions on the use of investigative case studies in meta-evaluation. *Evaluation and Program Planning*, 13, 373-378.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan T. (2000). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (2nd ed.). Boston: Kluwer Academic Publishers.