One-time case study. This design trains participants without any pre-test measure, but a post-training measure is taken during the program or shortly afterward. When this design is used, there usually has been little thought given to evaluation design.

Still, this is the most convenient evaluation design, because it requires only one measure. It's also the least rigorous, because it makes no comparisons. The best that can be said for this design is that it's more informative than taking no measure. This design is useful when constraints permit no preliminary data collection, if the primary evaluation client simply wants data collected to confirm or invalidate perceptions about a training program, or when participants have no prior background in the subject matter.

For example, when IBM offered a program on beginning Japanese to their international sales representatives, no pre-test was given because of their lack of exposure to the language. In such a case, it's appropriate to take performance measures only after participants' initial exposure to training content.

Data-collection tools

Data-collection tools can be categorized as either quantitative or qualitative instruments.

Quantitative instruments:
■ performance records and tests
■ standardized questionnaires and survey instruments
■ personnel assessment instruments.

The candor and accuracy of questionnaire and survey responses can be strengthened by assuring respondents of anonymity

Quantitative data have the following characteristics:
■ relatively easy to measure and assign dollar values to
■ objectively based
■ use a common measure of performance
■ credible to management

Qualitative instruments:
■ interviews
■ observations
■ focus group meetings
■ case studies

These instruments may be supplemented by other tools. For example, an evaluator may use a checklist of behaviors to guide an observation or may videotape a focus group meeting for additional review.

Qualitative data have the following characteristics:
■ difficulty in standardizing
■ subjectivity
■ behaviorally oriented
■ less credible to management.

At first, quantitative data appear superior to qualitative data. Yet quantitative data are more influenced by outside factors than qualitative data. Also, the appropriateness of quantitative or qualitative data depends on an evaluation's purposes, evaluation clients' questions, and overall evaluation design.

Quantitative data collection is more suitable when the following circumstances hold:
■ Evaluation is to determine whether a training program should be continued or expanded.
■ Evaluation's purpose is to identify a training program's economic impact on the organization.
■ A rigorous evaluation design is used.
■ Standardized data about a training program are needed.
■ The specific training is crucial to strategic businesses goals or safety.
■ Formal evaluation reports are required.

Qualitative data is more suitable under these circumstances:
■ The focus of the evaluation is to improve the program, to discover unanticipated consequences of the evaluation itself or to determine how a training program's success varies at different sites or among categories of participants.
■ Quantitative information needs to be augmented to provide depth and detail about a program's success.
■ Quantitative data are unavailable (for example, employee agreements may prohibit the collection of certain quantitative data).

To select the best design and data-collection methods for use in the real world of politics, personalities, and methodological imperfections, evaluators must match appropriate data-collection methods with evaluation purposes. So, in practice, quantitative and qualitative data are often gathered together.

The collection of information from multiple data sources or through several methods is called triangulation. Although more expensive and time-consuming, triangulation increases the probability of confirming training's responsibility for changes in employee performance and organizational measures.

Evaluation Practices

ASTD's research revealed that the actual practice of evaluation doesn't often follow the strict recommendations of evaluation literature. This is largely explained by the fact that many training practitioners haven't found the literature's advice applicable or useful for their organizations.

But, as well-known author and management consultant Thomas J. Peters has said, "What gets measured gets done. . . . Even imperfect measures provide an accurate strategic indication of progress, or lack thereof." So practitioners have employed various practical evaluations.

Here's an overview of current evaluation practices among organizational leaders in training, telling how and why they subscribe to their various practices. The evaluation techniques and practices explored don't meet traditional academic notions of rigor, but do provide valuable information, are reproducible, and can be quickly and easily conducted. Most of the training managers that par-

ticipated in ASTD's research effort believe that there's value in a concerted effort to increase the practice of employee training evaluation along these lines.

All the organizations represented in this study evaluate some aspect of their training programs. In terms of the four-level Kirkpatrick model (see page S-15), 75 to 100 percent of them evaluated training programs at the participant reaction level. Virtually all of them also evaluated participants' knowledge gains in some of their training programs. Twenty-five percent of their training programs were evaluated at this, the learning level.

Behavior change on the job was the least measured: among companies surveyed, only about 10 percent evaluated training at this level. Employee training was only evaluated at the organizational results level about 25 percent of the time, despite new pressures on training practitioners to assess the economic worth of HRD activities.

Sixty-six percent of the training managers reported that HRD professionals are under increasing pressure to show that programs are producing favorable bottom-line results. These managers had a strong track record in training evaluation or had high management acceptance of training as a way to meet real operational needs. So, in their experience, increased pressure did not mean that upper management doubted that training could be beneficial.

The reason usually given for closer scrutiny by management was that employee training is being recognized as a significantly large expenditure. But greater management attention coupled with movement toward cost reduction can be particularly injurious to expenditures (such as those for employee training) that have hard-to-isolate or long-term payoffs.

Although most training programs are evaluated at the reaction and learning levels, these levels aren't always consistent with the reasons for evaluation. Research suggests that evaluation conducted for the proper reasons helps determine training's impact on job performance and economics within an organization.

Most organizations evaluate training programs to meet the following demands:

■ **Training department demands.** For quality assurance, trainers gather information to direct their efforts to improve training effectiveness. Trainers also want to demonstrate training programs' worth to top or operating management. And, training managers want to build databases for future planning and analysis.

■ **Employee demands.** After a decade of downsizing and flattening of organizational pyramids, employees are seeking training that's timely and useful in meeting their new job responsibilities.

■ **Management demands.** Managers are scrutinizing training as a tool for gaining a competitive advantage. Many managers believe that having the best trained work force increases competitiveness.

In an economically competitive environment, however, it's necessary to justify training expenditures to ensure adequate returns on training investments. Many organizations are only willing to pay for training that seems relevant and efficient. But interviews with training managers reveal that most training decisions aren't the result of data derived from assessment of training's worth. Rather, decisions about training still tend to be based on management's perceptions about training programs' worth.

Consequently, most training investment decisions are a matter of tradition—and continue training programs already in place. For example, supervisory programs usually are conducted year after year but undergo occasional adjustments to meet changing company needs.

Some organizations, such as Aetna and Johnson & Johnson, allocate as much as 85 percent of their training budgets for the continuation of existing programs. That necessitates program monitoring to make sure that the programs continue to address line managers' training needs. Frequent communication between training department staff and line managers promotes trainers' understanding of business needs.

Of couse, some trainers avoid evaluating training programs for fear that hard results may not justify the expense of progams. That sentiment is particularly likely in an evaluator who had direct responsibility for part of a program. To avoid having evaluators with a vested interest in evaluation results, some organizations prefer to use outside evaluators of in-house training and in-house evalutors of contracted training. But outside evaluators' lack of in-depth familiarity with the organization carries its own problems. At least one company, Digital Equipment Corporation, has established an internal department for training quality assurance to resolve the inside/outside evaluator dilemma.

The driving force behind most employee training investment decisions is the line management structure. Line managers usually recognize employees' needs for new job knowledge and skills, because managers are closest to the problem. Generally, training investment decisions are made when line management reports a performance problem likely to require a training solution. The training department performs a front-end analysis to determine whether the difficulty is a problem that has a training solution. In most organizations, line or operating management initiates 75 to 95 percent of employee training.

But top management usually is responsible for initiating programs that are company-wide, related to overall company policy, start-ups of broad new activities (such as comprehensive quality assurance programs), or of special interest to executives. And, it's often top management that approves or disapproves a specific dollar amount for training investments.

Many decision makers assume that once behaviors are corrected, organizational economic indicators will improve automatically. Some organizations require no tangible evidence from training to demonstrate its worth other than its connection to strategic business goals. That's the straightforward reason some training managers give for not accounting for training through measurement and evaluation: upper management doesn't require it. But without accounting for training, it's hard to say whether training contributes enough organizational value or might be worthy of greater resources.

Who initiates training investments?

The extent of line management's involvement in initiating training was reflected in ASTD's investigation of training practices in pace-setting organizations. The

figure shows the percentages of training investment decisions made by line management (alone or in conjunction with the training department), top management, or the training department.

As the percentages shown in the box indicate, line management increasingly controls the investments an organization makes in training. Almost all of the organizations listed noted that a market-driven approach—repeated assessment of what training the organization needs rather than unquestioned continuance of existing training in products or programs—is the key to training department success. Therefore, training departments must learn which programs line management needs to function effectively.

Specific questions that training managers should ask line managers include:

■ What are departmental objectives for the fiscal year?
■ What problems have impinged on employees' assigned responsibilities, and do these responsibilities have training implications?
■ What are the career goals of key people in the department?

Communication between training department staff and line managers to focus on organizational problems, issues, plans, and strategies will identify the knowledge and skill gaps that impede performance. For example, Chase Manhattan Bank encouraged constant communication between line management and the training department by eliminating their training course catalogue. Chase believed that having a course catalogue discouraged the training department from rigorously assessing the training needs of each business unit.

Who Initiates Training Investments?

Company	Line Management	Top Management	Training Department
Polaroid	75%	—	25%
Johnson & Johnson	75%	15%	10%
Motorola	75%	10%	15%
Upjohn	95%	5%	—
DEC	95%	5%	
Arthur Andersen	75%	—	25%
Aetna	90%	—	10%
AT&T	80%	5%	15%
Chase Manhattan	95%	5%	—
Vulcan Materials	50%	50%	—
New England Telephone	90%	10%	
Xerox Customer & Marketing Education	75%	—	25%

Source: American Society for Training and Development, 1988.

Current practices

Research shows that most companies evaluate all their training programs in some fashion. All the companies that ASTD investigated build evaluation design into overall program design when program objectives are established. Sophisticated evaluation methods aren't applied to training programs if a program is a continuation of previously successful efforts, when less rigorous techniques are adequate for answering the majority of questions decision makers have about a program, or when practical considerations prohibit using such techniques.

Even in organizations relatively advanced in training evaluation, sophisticated statistical methods or controls are rarely employed. This may be for the good reasons above or, as Brandenburg and Smith have pointed out, because of trainers' lack of evaluation expertise (including uncertainty about when to measure for change) and training departments' concern that evaluation is tantamount to criticism.

Whatever methods they use, evaluators are increasingly using multiple data sources—combinations of quantitative and qualitative data.

The organizations that ASTD investigated hold conflicting views about the need to demonstrate training's connection to the bottom line. These views affect how evaluations are conducted within the organizations and how information about training programs is communicated to management.

One of the most popular purposes for evaluating training is to demonstrate training's worth to top or operating management. The organizations investigated showed the following approaches to training and evaluation:
■ Top management is greatly supportive of investments in training. Management perceptions of training's worth is the critical factor.
■ Line management views training as a strategic lever for achieving key business objectives. Evaluation focuses on whether key business objectives improve after training. If so, training receives at least a share of the credit.
■ An evaluation must demonstrate how a training program has made specific contributions to business objectives and, to the extent possible, must discount alternative explanations for improvements.

Organizational profiles: economic results

Vulcan Materials. Because of increasing industrial com-

Evaluation Practices

petitiveness, Vulcan Materials is evaluating training at the economic-results level. One-third of Vulcan's training investment decisions are based on cost-benefit analysis conducted before training implementation. Training department proposals include analysis of projected costs and benefits; evaluation design is built into program development. In order of preference, the priority of evaluation measures is output, quality, cost, timeliness, behavior (including that associated with attitudes), and observations.

As noted earlier, few organizations use statistical methods and controls for training evaluation, but management now places greater emphasis on demonstrating training's contribution to organizational economic goals. Accordingly, Vulcan's former human resources manager developed a simple, practical method for estimating training's economic results.

Before a training program's implementation, middle managers are asked to estimate the savings they expect to result for their departments. Middle managers also rate their confidence (on a 0 to 100 percent scale) that the training program itself (as opposed to other factors such as managers' on-the-job reinforcement of practices) will be responsible for the savings. Projected savings are multiplied by the "confidence" percentage to yield a forecast of total cost savings from training.

Once a training program is complete, evaluators examine actual cost savings or revenue increases that the organization has incurred. The training department then asks line managers to estimate (again, through a percentage) how responsible they now believe the training was for these improvements. The actual savings/revenue amount is multiplied by this percentage to provide an estimate of the total cost savings the training program has provided to the organization; that figure is then compared to the forecast.

An example of that practice is shown in a course to improve first-line supervisors' supervisory skills. En-hancement of supervisory skills was expected to reduce production worker turnover. Middle managers of the production workers' supervisors were asked to project the savings they expected to make as a result of reduced turnover and to indicate their confidence rating of training's responsibility for this savings.

The supervisor training program was estimated to cost $10,000. The overall confidence level of middle managers that the program would be responsible for reducing turnover among production workers was 50 percent. Middle managers estimated that the benefits from reduced turnover would equal $200,000. That expected total benefit of reduced turnover was multiplied by the 50 percent confidence level to estimate the training program's contribution to reduced turnover at $100,000—10 dollars of benefit for each dollar spent on training.

After the training department made a proposal to top management outlining the expected savings from reduced turnover, the training program was implemented. A study six months later revealed that actual total savings from reduced turnover amounted to $100,000. Middle managers again expressed a 50 percent confidence level that training was responsible for the savings. The actual total benefits of reduced turnover was multiplied by 50 percent to determine training's actual responsibility for improvements in turnover.

The result showed that the training program's contribution to reduced turnover was estimated at $50,000 or five dollars of benefits for each dollar spent on training. This success story of reduced turnover was publicized with some credit attributed to the training department.

Arthur Andersen regularly evaluates training at the economic results level, and evaluation is invariably designed into the original program. Anticipation of economic downturns when training may need more justification causes increasing pressure to show cost-benefit

Savings Forecast and Actual Savings

Savings forecast

$200,000 (expected benfits) x 50 (managers' confidence level of training responsibility) = $100,000

$100,000 (expected savings) – $10,000 (training costs) = $90,000 savings forecast

Post-training evaluation

$100,000 (expected benefits) x 50 (manager's confidence level of training responsibility) = $50,000

Actual savings

$50,000 – $10,000 = $40,000

Source: interview, Phillips, 1988

for training. Top management is generally receptive and supportive of evaluation, and eager to assess training's contribution to achieve strategic business goals.

That management perspective is reflected in evaluation policies. Arthur Andersen is unusual in that its evaluators use a wide range of statistical methods, including experimental designs as well as more common qualitative and quantitative evaluation designs.

For example, using the control group method and a pre-test and post-test evaluation design, Arthur Andersen compared accelerated learning techniques to traditional instructional design and delivery methods to determine which was more economical.

Thirty-eight participants were selected based on similar work and personal characteristics. Participants were randomly assigned to either an experimental (accelerated learning approach) or control (traditional approach) group. The traditional learning approach was evaluated as being superior to the accelerated learning approach in terms of participants' knowledge gains and reactions at the end of training. But a study four months later revealed that behavior improvements on the job were higher for the accelerated learning group than for the control group. That led to further research by Arthur Andersen on the economic value of accelerated learning techniques.

Johnson & Johnson, an organization with an excellent training reputation, evaluates all training programs. Although the primary evaluation emphasis is on behavioral change on the job, evaluators also pay attention to advancing operational initiatives.

For example, a training program that addresses a key operational initiative shares credit if improvements occur after training. Management accepts this claim of credit without requiring the support of statistical controls.

More than 80 percent of training decisions are based on qualitative data obtained from worldwide employee surveys. Actual decision making about where to invest in employee development usually isn't based on hard data or rigorous financial analysis to determine potential return on investment.

Recently, Johnson & Johnson's corporate headquarters mandated a strong training policy that exhibits top management's greatly increased support for training. Management is looking more intensely at employee training as a tool for gaining competitive advantage. But that hasn't lead to increased management demands for an economic-results level of evaluation. Instead, in response to the strong increase in management support and funding for training, there is self-generated pressure from the training department to demonstrate training's effectiveness.

Motorola is another organization where management's perspective of training influences evaluation of training programs at the results level. Top management is strongly supportive of training, to the point of having set mandatory minimums for training investments in all divisions. The chairman of Motorola has publicly singled out training as a critical means of supporting strategic business goals.

At Motorola, training plays a key role in achieving strategic goals and operating initiatives. Motorola's training and education center is charged with operating an extensive in-house training and education effort to raise employee skill and knowledge levels. Motorola's managers are committed to training as a means of improving productivity, performance, and profitability through the development and expansion of workforce skills.

Motorola's managers are also committed to the principle of participative management. They believe that only by developing and expanding employee skills will effective employee contributions—and consequent productivity improvements—be maximized.

Because sophisticated evaluation methodology is costly, it isn't used to evaluate training at the results level. In fact, Motorola's trainers believe that if they were to be required to evaluate training at the results level it would mean they had lost the demonstrated confidence of management. Almost all Motorola training programs are evaluated using quantitative and qualitative measures in order to improve training effectiveness and, at times, to demonstrate training's effectiveness.

There's no direct effort to demonstrate training's contribution to Motorola's economic objectives. But indirectly, by designing training programs to address five of Motorola's key operational initiatives, the training department can make a legitimate claim of contributing to those objectives' achievements. The operational initiatives are having zero production defects, reducing total cycle time, integrating production and manufacturing, becoming a customer-driven company, and developing a participative management culture.

Although training programs can share in the credit for improvements in measures of these initiatives that occur after training, the absence of statistical controls means training can't eliminate rival explanations for improvement. Still, Motorola's management has great confidence in training's contribution to economic goals.

That is partly because the training population is predominately production personnel. Result measures are usually more precise for them because their work is highly task-specific. For managers and staff personnel, units of measure are much less obvious, and there are many more non-training influences on work output.

Polaroid. Since the 1980s, Polaroid has demonstrated that training affects the bottom line. Usually, the link between training and an improved bottom line hasn't been a matter of hard data analysis.

The units of measurement for evaluating training are the same units that line managers use as performance standards (for example, production units per hour per employee). Projections about training's impact on the bottom line are made without the support of statistical controls. Sometimes, benefits are based on projections of past, similar experiences.

A few years ago management showed strong endorsement of training's worth at Polaroid: training expenditures were doubled although the company was going through an overall downsizing of 30 percent. Training was perceived as a vital economic force although hard data were not available to substantiate that belief.

Evaluating training at the behavioral level

In some organizations, evaluation looks at training's on-the-job application. ASTD's research indicates that the most popular reasons for evaluation are to gather information that helps decision makers improve the training process and to facilitate participants' job performance. So evaluation is done to ascertain whether participants meet program learning objectives and transfer learning to the job setting.

The organizations that ASTD investigated regarded training and line management collaboration in evaluation design as critical to accurate identification of employee behaviors necessary to contribute to organizational goals. The organizations interviewed use a variety of practical qualitative and quantitative measures of individual performance to demonstrate positive transfer of what was learned in training to the employees' jobs. Virtually all the organization's representatives stressed the importance of waiting a reasonable period of time before assessment of on-the-job behavioral changes, typically six months.

Organizational profiles: behavioral results

Johnson & Johnson acknowledges training's direct support of strategic business goals. One reflection of that is the recent addition of a state-of-the art management training center at Johnson & Johnson's corporate headquarters. To ensure quality training, Johnson & Johnson's training director has directed that new emphasis be placed on measurement to augment the company's existing evaluation strategy.

Currently, Johnson & Johnson uses a six-months' post-training self-report form to evaluate training participants' on-the-job changes and management potential and practices. Subordinates of managerial and supervisory trainees are surveyed annually about their perceptions of mangerial performance changes.

Johnson & Johnson recently developed a first-line supervisors' competency model to identify critical supervisory work skills. Training now is being used as the strategic lever to help 350 first-line supervisors achieve these competencies. This program's evaluation has three parts:

■ Participants define their perceptions of change.
■ Participants' supervisors fill out questionnaires.
■ Five to eight subordinates of participants respond to a survey.

The Travelers Companies. Travelers stresses evaluation of training at the behavioral level with few attempts at translating evaluation results into dollars earned or profit. Travelers believes that it's more relevant to evaluate a training program's effect on management behavior, supervisory tactics, and strategic planning efforts.

Travelers regards evaluation as critical because of training's perceived importance in facilitating management competencies.

Travelers has identified management competencies for supervisors, managers, and directors and has developed a continuum of 27 training programs for each managerial level. The critical managerial competencies were identified through surveys of 700 Travelers' managerial personnel nationwide. Fifty Travelers managers identified as extremely high-potential performers were also surveyed to help isolate the traits and skills most important for Travelers' managers. Analysis of those traits and skills shaped training and evaluation plans.

Trainers at Travelers would prefer to use more statistical evaluation methods, but current staff capacity and lack of management demand for precision discourage the use of more rigorous evaluation methodology.

At present, Travelers uses a variety of qualitative and quantitative methods. Quality assurance checks are performed to assess program validity and relevance during training implementation. Qualitative data indicate whether management training programs need revision. Information about participant reactions is gathered immediately after a training program and also through evaluator observations and post-course group discussions with participants. The useful evaluation information that management has been receiving may eventually pave the way for more rigorous evaluations.

Travelers trains instructors in evaluation techniques (such as observation and group discussion), so they can act as self-evaluators. It's not economical for Travelers to use an evaluator/observer in every training program, especially those that have been operating for some time. For such programs, evaluation is primarily a maintenance function. But, a professional evaluator is usually called in to assess new training programs.

Evaluation data are used to discern whether participants view training as valuable and transferable to the job. As a rule, if more than 25 percent of a training program's participants express dissatisfaction with an aspect of the course as it relates to the job, revisions will be made. When revisions are to be made in a course, evaluation data are reviewed by a committee of line managers, instructional designers, an evaluator, and an instructor.

New England Telephone conducts evaluations primarily to determine whether training meets the expectations of line managers once workers are back on the job. New England Telephone doesn't use statistical methods but prefers to build a case for their training programs' value through reaction sheets and focus group meetings.

New England Telephone uses three reaction forms.

The first, a questionnaire, is administered to participants immediately after a course. It asks participants to evaluate the instructor and facilities and to assess the transferability of the training curriculum to the job.

The evaluator administers a second reaction form to participants three months after training. This form asks participants to reconsider training's effectiveness for their overall jobs and to assess why it was or wasn't valuable.

The training department sends a third questionnaire to line managers in order to assess a program's effectiveness in serving their mission.

Focus group meetings are held several months after training and may last an entire working day. These meetings reinforce the role of training in facilitating corporate objectives by tying training closer to strategic business

goals. By examining discrepancies between training expectations and actual accomplishments, they also enhance credibility.

Focus group members include an evaluator, middle managers, instructors, participants, and instructional designers. The evaluator moderates the meeting and is responsible for condensing participant reaction sheet data and qualitative data gathered from interviews with participants and the instructor. Discussions focus on necessary changes. Major questions that the focus groups address:

■ Does the course content require revision to meet existing or changing business requirements?
■ Could the program be redesigned to be more effective?

When developing a training program, New England Telephone holds focus group meetings for the client and training department. The client is closely involved in establishing training program learning objectives. The after-training focus group discusses whether these learning objectives were adequately realized in the program.

The training department claims that by understanding and serving the strategic business goals of the various business groups in the organization, training can share in the responsibility for economic improvements. Consistent with this approach, New England Telephone's evaluation unit doesn't serve as a judge but as an internal consultant charged with improving the relevance and quality of training.

Xerox Corporation Customer and Marketing Education.
Customer and Marketing Education's primary evaluation goal is to determine whether the skills, knowledge, and attitudes presented in training have been transferred to particpants' job performance. Line management and the training department communicate before training program goals and objectives are established. After a training program, an evaluator determines whether the learning objectives set by the client have had a positive effect on job performance.

Questions that evaluations seek to answer:
■ Do participants use what they learned in the training program?
■ Which training objectives did and did not transfer to participants' jobs?
■ If transfer to jobs didn't occur, what factors are responsible?

Evaluation begins in the pilot stage of program development, when a program is in draft form and hasn't been delivered on a regular basis. This first evaluation step helps identify and correct confusing or irrelevant material before a program is fully implemented.

Pilot-test evaluation data are gathered from several sources: tests of the curriculum, classroom observations, reaction forms from participants, and debriefings with participants, instructional designers, and line managers.

The second evaluation step takes place during training. Training programs are monitored to determine whether participants achieve program learning objectives and to obtain participants' reactions. Data are collected over a period of time—from tests, participant reaction sheets, and after-course discussions with participants.

This approach permits the evaluator to examine trends and patterns. For instance, the evaluator looks to see whether specific types of participants are reacting similarly to a training program. In this step, evaluation seeks to answer whether participants achieve program objectives and whether the program meets participants' needs.

The final evaluation step takes place after training and centers on whether programs learning objectives have transferred to the job setting. This evaluation steps seeks answers to these questions:
■ Are people using what they learned in training on the job?
■ Which learning objectives transferred to the job?
■ What factors were responsible for training not transferring to the job?

For this step, data-collection methods include the following:
■ observations of former participants in their job settings
■ roundtable discussions with participants after training
■ interviews with participants
■ questionnaires
■ telephone interviews
■ examinations of job records
■ interviews with former participants' managers.

Unless a client requests quantitative measures, Xerox finds that using a variety of qualitative methods is more appropriate for evaluation than a strictly quantitative or statistical approach.

Evaluators observe former participants in their work settings in order to verify data gathered from interviews, round table discussions, or questionnaires. These observations also help determine whether any barriers in the working environment inhibit the transfer of learning to the job setting. Barriers might include conflicting organizational policies and practices or managerial unwillingness to allow participants to use the new skills. Barriers also may be uncovered through group discussions (among an evaluator, managers, instructors, and instructional designers) that focus on learning transfer.

AT&T uses a combination certification-evaluation program that measures whether trainees can do a job following training. Very few AT&T training investment decisions are based on hard productivity data, but decisions are invariably based on analysis of competency deficiencies.

AT&T enables an individual business unit to certify employee performance on the job through a needs analysis. The analysis identifies the knowledge, skills, and abilities that workers must have to perform adequately.

Evaluation design is always a part of program design. And AT&T evaluates to ensure that design standards are valid for producing effective training. AT&T regularly evaluates all training with the primary aim of assuring that the training is up to date and serves its intended purpose. Rather than using criterion-based testing, AT&T uses line managers and expert judgment to determine whether behavior has improved on the job. AT&T also uses a five-point self-report form to measure participants' reactions to a course, their perceptions about whether training has enabled them to perform their jobs better, and their assessment of training's contribution to their career development. ■