Performance-Learning Satisfaction Evaluation System


Richard A. Swanson

April 1996

University of Minnesota
DEPARTMENT OF VOCATIONAL & TECHNICAL EDUCATION • ST. PAUL, MN 612-624-7481 Fax 612-624-4720
Introduction

The gap between evaluation sound theory and practice and typical practice is a serious problem for human resource development-performance improvement (PI-HRD) efforts in industry, business, and government organizations (Swanson, 1990). PI-HRD includes personnel training and development, organization development, and more general performance improvement efforts. A literature review of the summative evaluation noted that evaluation, when conducted, is not conducted effectively (Parker, 1986).

A dilemma facing most practicing PI-HRD professionals is that top managers and/or clients typically neither ask for nor require formal evaluations. And, when top managers do evaluate, it is usually in response to a crisis and invariably comes too late.

A typical work-place scenario consists of the busy PI-HRD practitioner doing what the company wants, feeling successful, and not being regularly required to prove the added value that results from PI-HRD contributions. With a full agenda of important development and delivery tasks, the busy professional finds it difficult to evaluate. However, almost all important organizational processes and functions are regularly evaluated in terms of their effectiveness, efficiency, and bottom-line contributions to the enterprise. In addition, it has been clearly established that effectiveness evaluation data, particularly bottom-line performance results, are the key to gaining support for the PI-HRD function from top management (Kusy, 1988). It is clearly irrational not to evaluate PI-HRD programs and processes.

The purpose of this document is to present a practical and valid evaluation system that can be applied to any PI-HRD program in business and industry. The system, titled the Performance-Learning-Satisfaction Evaluation System (PLS Evaluation System) was originally developed and tested by the author for Manville Corporation, Control Data Corporation, and Kellogg Company. Since then, the system has been applied by practitioners and researchers in a wide variety of settings. Most recently CIGNA HealthCare has implemented and advanced the application of the PLS Evaluation System.

Both Parker's review of literature and Kusy's study of management support of effectiveness evaluation established the need for this study. Additional criticisms of the present state of evaluation theory and practice provide further impetus (Alliger & Janak, 1989; Holton, 1995; Newstrom, 1995). Sadly, the most vocal defenders of inadequate theory and practice demonstrate little understanding of sound evaluation practice or theory (Kirkpatrick, 1995 & 1996). In addition, evaluation is one of the primary phases of the systematic PI-HRD process. The first four phases of the process include analysis, design, development and implementation (or some variation of these elements).

Note: The author wishes to thank Dr. Timothy McClernon, VP. of Sales Performance Consulting, CIGNA HealthCare and CIGNA Corporation for their contributions to the advancement of the PLS Evaluation System.
PLS System Overview

The PLS Evaluation System addresses the evaluation phase PI-HRD and consists of five major elements:

1) PLS Evaluation Model
2) PLS Evaluation Plan
3) PLS Evaluation Tools
4) PLS Evaluation Schedule
5) PLS Evaluation Report.

The PLS Evaluation System is grounded in and utilizes sound economic, systems, psychological theory (Swanson, 1995). The evaluation system logically connects the performance goals specified in the up-front performance analysis with the performance outcomes. Also acknowledged are the related or mediating outcomes of learning and satisfaction.

Without a connection to up-front analysis, it is unlikely that performance results at the organizational, process, and individual levels will be found. The fundamental flaw of providing an intervention and then looking for results (or transfer) has been at the crux of PI-HRD poor practice (Swanson, 1994). Even so, the results of placing a rigorous evaluation process on top of a poorly conceived PI-HRD intervention can cause an organization to reassess its practices. I conducted one such evaluation that resulted in top management’s sobering realization that the most expensive and most popular executive development program they had been supporting for ten years had not produced any discernible positive impact on the organization, its processes, or individuals.

The dilemma facing the PI-HRD profession is that the professional advice they are getting leaves them relatively impotent in terms of evaluation. The advice from evaluation theorists is overwhelming and unworkable. They create scenarios that are complex, situational, and cautious to a fault. The advice from evaluation practitioner advocates are simple, situational, and flawed beyond repair.

What is needed and what is proposed with the PLS Evaluation System is a loose-tight evaluation system. It is loose enough to accommodate almost all situations. It is tight enough to consistently honor core questions, techniques, and reporting results as part of every evaluation. In addition to the PLS Evaluation System conceptual components, job aids, tools, and references, there is an auxiliary data processing system. This low-cost system consists of a computer, software, a data input scanner, and a printer. This system allows for the production of program specific scannable data entry sheets, scanning of completed sheets into the computer memory, and the manipulation of the data into a standard format report.

The PLS Evaluation System can stand the tests of sound theory and practice. This does not mean that the process is painless. There are costs to good evaluation. And, as with all business decisions, the benefits should exceed the costs. PI-HRD professionals who have been operating with inadequate evaluation will most certainly need to absorb a high portion of the initial costs (in terms of development their own expertise and installation of the system) before they yield the benefits of being taken seriously as a contributor to the organization.
PLS Evaluation Model

The evaluation model embraces the domains of performance, learning, and satisfaction (PLS). Each of the three domains are divided into two options for a total of six evaluation options labeled A-F.

The three domains of performance, learning, and satisfaction are not levels or a hierarchy (see Holton, 1995; Newstrom, 1995). They are domains of independent worth from an evaluation perspective. While there can be a relationship between the domains, it cannot and should not be assumed that there is a direct positive relationship. Contrary to the practitioners myth, the research shows that participants most satisfied with a program are not necessarily those who learned the most (Alliger & Janak, 1989; Dixon, 1990). High or low satisfaction can be found among low, medium, and high achievers. And, because participants have gained knowledge and expertise does not mean that they will use it in the workplace (Gielen, 1995).

Performance can be thought of as (a) business results at the organizational, process or individual levels and/or (b) financial results or benefits in terms of money or monetary ratios. Business results that are monetized become financial results. For example, the business result of increasing market share by 5% could be converted into numbers of sales times the financial value of each sale. Another example would be to increase the hourly production rate of injection molded items (that meet quality standard) from 30 to 35. In each of these examples the business results can be monetized and then expressed in terms of financial results. The general financial results goal of the PLS Evaluation System is that the benefits exceed the costs by a 2:1 ratio. If the 2:1 return on investment goal is achieved, it can be said that the program achieved 100% of its goal. A 2:1 return on investment, especially in a year or less, is ambitious. Organizations may want to choose and adhere to a lesser goal.

Learning can be thought of as (c) knowledge demonstrated through mastery of information and concepts and/or (d) expertise in demonstrating the skillful workplace behaviors in a simulated or actual work situation. Knowledge is most efficiently measured through paper and pencil tests. This is because paper and pencil tests allow for a larger sample of learning from each participant in a minimum amount of time. While knowledge is a component of expertise, demonstration and assessment of skillful workplace behaviors pushes the training closer to reality. As with the performance and satisfaction domains, rational goals need to be established for the learning areas of knowledge and expertise. The 100% goals for knowledge and expertise must be connected to the behavioral outcomes established for the intervention. These outcomes generally focus on (1) not meeting, (2) meeting, or (3) exceeding standards. The general learning results goal of the PLS Evaluation System is that participants, on the average, meet learning and expertise standards. Thus, if meeting standard is 2 on a 1-3 scale, an average score of 2 would be attainment of 100% of the goal.

Satisfaction can be thought of as the (e) perceptions of the interventions by the participants directly involved and/or (f) perceptions of the sponsor (or supervisor) as to the participants and/or interventions. The overriding concern is that perceptions be acceptable, not negative and not jubilant. Since most development and improvement efforts require change, and change is not comfortable, it is illogical to pursue extremely high ratings. And, extremely low ratings end up being culturally counter productive. The standard 4 Likert scale (2 levels each of negative and positive with an acceptable mid-point score of 2.5) establish 2.5 average as the "acceptable" satisfaction goal. Thus, if the satisfaction rating average is 2.5, it can be said that the program achieved 100% of its goal. Only after
there is clear evidence that performance and learning goals are being reached should there be focused concern for raising the 2.5 satisfaction goal.

The PLS Evaluation Model, key references, and definitions are displayed in Figure 1. Additional references are presented at the conclusion of this manuscript.

---

**Figure 1 About Here**

---

**PLS Evaluation Plan**

In the PLS Evaluation System, planning decisions are made about which evaluation domains and options within each domain will be used to evaluate a specific performance improvement program or intervention. The decisions are recorded on the PLS Evaluation Plan.

**Domains & Options.** The rows of the PLS Evaluation Plan follow the three evaluation domains and the six evaluation options presented in the PLS Evaluation Model (Figure 1). Figure 2 is a sample completed PLS Evaluation Plan for “Communication” training. The vertical columns require the planner to make a series of decisions in relation to the evaluation domains and options.

---

**Figure 2 About Here**

---

**Choices.** The first set of decisions has to do with the choices of inclusion among the three domains and six options. Which will be included? The general assumption is that all three domains are important and that every effort should be made to include all three. And, the further assumption is that serious consideration should be made to pursue to all six of the A-F options (two options each within each of the three domains).

**Points on the Data Collection Timeline.** Once the evaluation option choices have been made, data pertaining to the program or intervention will be collected within a timeline. This timeline can generally be thought of as being before, during, or after the intervention. Thus, it may be possible to evaluate participant knowledge before, during, or after a program. And, as you can see by the PLS Evaluation Plan sheet, there are three intervals of time in the before stage (1-3), three in the during stage (4-6), and three in the after stage (7-9). For example, time interval 4 is during the formal program and at the very beginning, 5 is in the middle if the program, and 6 is at the end of the formal program.

The idea that the evaluation timeline identifies specific time intervals to collect specific data can be quickly planned and illustrated with this planning sheet. For example, the evaluation planning sheet helps illustrate the idea that participants have before program knowledge (time interval #3) that differs from program ending knowledge (time interval #6).

**Other Comparison Options.** It is not always appropriate or desirable to make within group comparisons (e.g. a person’s expertise at one point compared to that same person’s expertise at another point in time). Other Comparison Options include columns 10, 11, and 12.
Column 10 refers to another cycle of the same program (e.g., the June section of Communication Development compared to the July section of Communication Development). Column 11 refers to a standard—an attainment level set that has verified meaning. For example, participant satisfaction rating should be acceptable, but not excessively positive. The research demonstrates clearly that gaining new knowledge and expertise is hard—not necessarily an easy experience. Thus, the standard is to have an acceptable average rating of 2.5 or higher on a 4.0 scale, not a 4.0. Column 12 is the norm. A large data base provides averages that can be used as a point of comparison.

**Data Analysis Plan.** Just what data will be collected, compared, and presented as part of the final evaluation report for a specific intervention? The *Data Analysis Plan* column allows evaluation planners to capture the data to be compared to answer the evaluation question for a specific evaluation option (A-F).

An example here is: E6 (participant satisfaction at the end of program) compared to E11 (standard of 2.5 or higher on a 1-4 scale). Another example is: D2-D3 (average expertise before attending program) compared to D7-D8 (average expertise after participating in the program).

**Evaluation Plan Notes.** These notes are organized in six rows to match the six evaluation options (A-F). Comments and functional decisions effecting the evaluation effort are noted.

### PLS Evaluation Tools

Evaluation tools can be constructed and used for each of the six evaluation options (A-F). Of these, performance scores are indicators of the business effects that result from a PI-HRD program. Learning scores are indicators of the amount of knowledge and expertise acquired by the participants during the development program, and the satisfaction scores are indicators of how participants and their sponsors perceive the program. Although there are many possible ways to configure evaluation tools, the *PLS Evaluation System* focuses on a on six reasonable options (A-F) and recognizes that not all six are always used.

**Performance.** In the *PLS Evaluation System*, the tools for measuring the performance results back in the organization are *business results* and *financial results*. Business results contrast the productivity of either the organization, business process, or individual employees (1) before and after the PI-HRD program or (2) against a comparison group, standard, or norm. Figure 3 is the job aid for evaluating business results. It is always desirable to utilize the organization’s existing direct measures of performance such as sales, items produced, customer’s served. In that organizations are dynamic, judgments may be required in terms of attributing the portion of performance attributable to the intervention being evaluated. For example, one large company-wide intervention I studied carried a marketing component and a system-wide service improvement component. In the end, the VP of Marketing and the VP of PI-HRD met with the President to agree on the distribution of the business results and the financial gains to their respective contributions. In this case, and in every case I have participated in, estimates of the PI-HRD contribution by line managers have been understates. As might be expected, line personnel tend to want to take as much credit as they can. Even so, this state of affairs still leaves high quality PI-HRD interventions in the realm of excellent business investments.

*Financial results* is used to determine the economic value of the PI-HRD program, the benefit of which is determined by subtracting the cost of the program from the performance value resulting from the program. The financial cost accounting methods and norms
utilized by the organization should be consistently applied to all calculations. Thus, if employee salary during the time they are participating in a development program is not normally costed back to the intervention, it is not considered as part of the cost basis. In many calculations this salary cost variable will already show itself in lost productivity from being absent from the workplace. On the performance value side, it is generally easier than you think to determine the value of a unit of performance. For example, in one organization the assistant comptroller was able to tell me the average net income resulting from a sale. In another organization, the VP of Manufacturing informed me of the financial worth of the hourly output of fully functioning machining centers. This critical financial information is almost always a few tiers away from the place where the actual performance takes place. Financial benefits are a matter of subtracting costs from the resulting performance values. Interpretation of this data has been discussed in earlier sections of this article. Figure 4 is the PLS Evaluation System job aid for conducting financial benefit analysis of PI-HRD programs.

**Learning.** Learning is measured by knowledge tests, tests of expertise, or both. Knowledge tests measure the cognitive information learned by participants. Two types of paper and pencil knowledge test items, multiple choice and matching, are encouraged because they can be scored objectively and are not as susceptible to guessing as are other forms of test items and not as subjective and time consuming as verbal testing or written essays. In constructing knowledge tests, care must be taken to ensure that the tests produce valid and reliable results. A test is valid when it measures what it is supposed to measure and it is reliable when it produces consistent results. The job aid for constructing knowledge tests (Figure 5) includes sample test items, validity and reliability criteria, and helpful test construction references.

Tests of expertise measure what the learners can do by examining either the products that the learners produce or the processes exhibited by the participants as a result of their learning. An in-program test of expertise must also be valid and reliable. Interpretation of this data has been discussed in earlier sections of this article. The job aid for constructing in-program performance tests (Figure 5) provides examples, criteria for validity and reliability, and helpful references on tests of expertise.

**Satisfaction.** The PLS Evaluation System requires that "satisfaction" be measured for every intervention. Participant satisfaction is measured by having each participant complete the Program Evaluation Form at the completion of the program (Figure 7). This is a standardized form that should not be altered in terms of the front page questions. The participant satisfaction score is calculated by tallying all the participant responses to the front page questions. Ordinal values are then assigned to descriptors as follows: Very good (4), good (3), fair (2) and poor (1). The overall participant satisfaction score is obtained by averaging the scores and determining the mean satisfaction score which will fall within the 1-4 range. Sub-scores on the individual questions can also be computed this way. Interpretation of this data has been discussed in earlier sections of this article. The second side of the satisfaction form can be used for program specific questions of interest or rating the attainment of the specific program objectives.

The comments written by the participants on the participant satisfaction form are not included in the participant satisfaction score. They provide open-ended feedback for the facilitators, presenters, and program planners for program improvement.

Sponsor satisfaction is measured by using the Sponsor Evaluation Form (Figure 8), which is completed by each participant's supervisor or sponsor. After the responses are gathered, the average sponsor satisfaction score for the PI-HRD program is computed in the same manner that the average participant satisfaction score and sub-scores are determined. As
before, the written comments provide the facilitator, presenter, or program planner with immediate, open-ended feedback.

Using standard participant and sponsor satisfaction forms for all training courses allows for the comparisons of PI-HRD programs with each other and across time, making it easy to identify and document recurring problems and/or successes.

Figures 3 - 8 About Here

**PLS Evaluation Schedule**

The *PLS Evaluation Schedule* is simply a matter of operationalizing the *PLS Evaluation Plan* in terms of activities over time. In addition, each activity is coded as to when it is to be completed, who is responsible for the activity, who determines if completed activity meets quality standards, and who else is involved. Many organizations utilize standard project management tools that can be inserted for this realm of activity. Figure 9 is a simple one-page alternative.

Figure 9 About Here

**PLS Evaluation Report**

The *PLS Evaluation Report* is an “Executive Summary” of the effectiveness of a program (or set of identical interventions). The intention is that every program offering be evaluated in terms of its’ effectiveness and that the results be reported to the appropriate stakeholders in the organization. *PLS Evaluation Reports* have a standard format, standard sections, and standard means of reporting data. And, reports are almost always short and generally 2-4 pages in length.

There are eight standard sections listed below followed by a description and then illustrated in Figure 10:

- Organization and Program Identification Heading
- Program Purpose
- Program Description
- Evaluation Summary
- Approval
- Distribution List
- Evaluation Results (Performance, Learning, Satisfaction)
- Improvement Proposal

In that the *PLS Evaluation Report* is an executive summary, there is available evaluation data exceeding what is contained in the report. This additional data is retained and used by the department for tracking and improving specific elements of the program and for responding to specific evaluation inquiries.

The “Organization and Program Identification Heading” provides the critical identification data as to the program name, dates of the training, location, and number of participants. This individual program offering level of information is how all data should be stored in the
data base. Like sessions, such as 24 offering of an identical training program, can be combined into a cumulative report. That report should have an equally descriptive, yet unique, header.

“Program Purpose” is a simple and direct statement, usually in 50 words or less, as to the connection of the program to the business goals and/or improving a core business process. In comparison, the “Program Description” is a 50-75 word description of the program itself including the title, length, descriptive features in terms of content and/or method, and the immediate outcome.

The “Evaluation Summary” is a simple presentation of the selected evaluation tools (A-F) and the percentage of attainment of the established goal within each. For example, Financial Results could be 200% of the goal while Expertise could be 98% of goal.

“Approval” is a space for the signature of the appropriate person responsible for the program. The “Distribution List” is the list of those persons who should be receiving evaluation all reports related to this program. There is likely a short list of people who should receive evaluation reports for all programs and an additional list for each specific program. Together they make up the distribution list for any single report.

“Evaluation Results” is the largest single section of the report. The evaluation data is collected and reported in the effectiveness domains of Performance (A. Business Results & B. Financial Results), Learning (C. Knowledge & D. Expertise), and Satisfaction (E. Participant & F. Sponsor). All data is converted into a standard “percentage of goal” scale. For example, the participant satisfaction goal is to achieve an acceptable or higher average participant satisfaction rating. Thus, a goal of 2.5 on a 1-4 scale. Therefore, 2.5 represents 100% of the goal. If the average was 2.0 (below the goal), the percent of goal attained would be reported as 75%. A satisfaction average of 3.0 would be 125% of the goal. The purpose of this simplified percentage of goal reporting method is to ensure the accurate interpretation of the evaluation data by the most number of people as they view this report and all subsequent reports.

The “Improvement Proposal” is the last section of the PLS Evaluation Report. This 25-100 word statement focuses on the evaluation results (1) in context of the purpose of the program and the description of the program, and (2) what needs to be done improve performance.

The PLS Evaluation System is designed in a manner that allows the utilization of computer hardware and software to create instruments, scan data into the computer and to call out the data required for the PLS Evaluation Report. A sample report is presented in Figure 5.

Figure 10 About Here

Conclusion

The purpose of this document was to present a practical and valid evaluation system that can be applied to any PI-HRD program in business, industry, and government-- for profit or non-profit organizations. The system, titled the Performance-Learning-Satisfaction Evaluation System (PLS Evaluation System) has been applied by practitioners and researchers in a wide variety of settings.
Evaluation References


Figure 1

PLS EVALUATION MODEL
"Performance-Learning-Satisfaction"
A three-domain evaluation taxonomy for Performance Improvement
and Human Resource Development

PERFORMANCE

• Business results...organizational, process, or individual units
• Financial results...benefits in terms of money or monetary ratios

LEARNING

• Knowledge.....mastery of the information and concepts
• Expertise.... demonstration of workplace expertise

SATISFACTION

• Participants.....those people directly involved in the intervention
• Sponsors.....of the participants and/or intervention

© Richard A. Swanson 1995

**Evaluation**
Evaluation is a systematic collection of evidence to determine if desired changes are taking place.

**Reliability**
Reliability is a measure of consistency.
(e.g. Does the test yield consistent results with repeated use?)

**Validity**
Validity is a measure of accuracy.
(e.g. Does the test measure what it is supposed to measure?)

**Summative Evaluation**
Determining the effectiveness of an intervention.
(e.g. Has the goal been reached?
Does the program have worth or merit?)

**Formative Evaluation**
Determining the need for changes to guide the process and improve the odds of reaching a goal.
(formative evaluation is part of the intervention/change process)
Figure 3. PERFORMANCE - Business Results Job Aid

THREE TYPES OF TESTS OF BUSINESS RESULTS:

1. **Organization Performance (sample A & B):**

   A. The performance improvement staff decided to evaluate the effectiveness of the new product sales training program by considering whether training made an impact on yearly sales. To conduct this evaluation, they utilized a design which examined sales volume per month across each sales district before training and after training. The length of the before and after training phases was contingent on the sales district of which the sales representative was a member. The staggered line represents the actual training program which occurred over a two-day period. The following figure illustrates this method:

   ![Graph showing sales before and after training]

2. **Process Performance (sample A & B):**

   a. Doing the right activities: National Healthcare has a defined sales process and wishes to determine if the major process are being followed. Sales personnel monitored in terms of the level of adherence to the process activities.

3. **Individual Job Holder Performance (sample A & B):**

   a. Quantity: The blow molder operator is responsible for a single complex machine that produces a high volume of plastic bottles for liquid products. The quantity of acceptable quality bottles produced per shift is recorded.

   b. Time: The blow molder operator is expected to be able to shut down a machine, re-tool it, and start it up in an established time period.

**ESTABLISHING CONTENT VALIDITY:** (does the test measure what it is supposed to measure?)

1. Determine if the organization regularly collects data on the performance of the organization, process, work group, or individual workers in the area under investigation.
2. Make sure that unit of performance selected is the same or a good approximation of the performance need specified in the original needs assessment.

**ESTABLISHING RELIABILITY:** (does the test yield consistent results?)

1. If using organizational records, inquire about the reliability of the data collection methods.
2. Use controls such as comparison of group performance during earlier time periods before and after the program.

### THREE TYPES OF FINANCIAL DATA:

1. **Cost Analysis Worksheet**

<table>
<thead>
<tr>
<th>Forecaster</th>
<th>Date</th>
</tr>
</thead>
</table>

   1. Analysis: Performance/Needs
      | Staff | External consultant costs | Materials |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      Sub-total $

   2. Analysis: Expertise/Work behavior
      | Staff | External consultant costs | Materials |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      Sub-total $

   3. Design
      | Staff | External consultant costs | Materials | External support costs |
      |       |                           |           |                       |
      |       |                           |           |                       |
      |       |                           |           |                       |
      |       |                           |           |                       |
      Sub-total $

   4. Development
      | Staff | External consultant costs | Materials |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      |       |                           |           |
      Sub-total $

   5. Implementation
      | Participant | Facilities | Tuition/fees | Staff | Materials |
      |             |           |             |       |           |
      |             |           |             |       |           |
      |             |           |             |       |           |
      |             |           |             |       |           |
      Sub-total $

   6. Evaluation
      | Staff | External consultant costs |
      |       |                           |
      |       |                           |
      |       |                           |
      |       |                           |
      Sub-total $

   7. Total costs (sum of all sub-totals) Total $

2. **Performance Value Analysis Worksheet**

   **A. Data Required for Calculations**

   (a) What unit of measure was used to describe the performance? $
      (units)$

   (b) What was the level of performance per worker at the conclusion of development?

   (c) What was the level of performance per worker before intervention?

   (d) What dollar value was assigned to each unit?

   (e) How long was the development time to reach the goal (b)?

   (f) How long was the evaluation period?

   (At a minimum this is the longest time of all options that were considered.)

   (g) How many workers participated in the development?

   **B. Calculations to Determine Net Performance Value**

   (h) Did participants produce during development? $\text{No} = 0 \quad \text{Yes} = b + c$

   (i) The number of units produced per worker during development $= e x h$

   (j) The number of units produced by a worker during the evaluation period
       $= (f - e) x b + i$

   (k) The dollar value of a worker's performance during the evaluation period $= d x j$

   (l) The net performance value gained per worker $= k - (c x d x f)$

   (m) Do you want to calculate the total net performance value of all participants? $\text{Yes} = 1 \times g \quad \text{No} = 1$

3. **Financial Benefit Analysis**

   Performance Value
   $\cdot$ Cost  $
   $\cdot$ Benefit $

---

**ESTABLISHING CONTENT VALIDITY:** (…does the test measure what it is supposed to measure?)

1. Make sure the cost categories are the same as those regularly used in the organization. Have someone in accounting and HRD department verify categories.

2. Make sure the unit of performance and its worth is reasonable and acceptable to the decision makers in the organization.

**ESTABLISHING RELIABILITY:** (…does the test yield consistent results?)

1. Double check the individual numbers and their manipulation in the formula.

2. Have a second analyst prepare a financial benefit analysis.

Figure 5. LEARNING - Knowledge Test Job Aid

TWO TYPES OF PAPER AND PENCIL TEST ITEMS:

1. **Multiple choice test items (samples):**

   To speed up nut turning on tasks where space is limited or where bolts with long threads prevent the use of sockets, use the _________ wrench.
   
   (a) crescent  
   (b) combination  
   (c) ratchet  
   (d) allen

   Indicate the best answer by circling the number. The most important property of an objective test is:
   (1) Ease of marking  
   (2) Accuracy of scoring  
   (3) Its reliability  
   (4) Its validity  
   (5) Complete sampling of the syllabus

2. **Matching test items(samples):**

   For each item, write a number to indicate that the statement applies to:
   (1) Norm-referenced assessment  
   (2) Criterion-referenced assessment  
   (3) Both norm- and criterion-referenced assessment  
   (4) Neither norm- nor criterion-referenced assessment

   __________ Assessment is mastery-based  
   __________ Some people must fail; otherwise assessment is too easy  
   __________ Assessment is useful for making predictions

   Directions: Column A contains a list of advantages of varied shopping outlets. Choose from column B the outlet which best fits each advantage in column A and insert the identifying letter in the space provided. Responses in column B may be used more than once.

   Column A  
   (1) "One-stop" shopping  
   (2) Offers 24-hour service  
   (3) Armchair shopping  
   (4) All prices may be lower  
   (5) Product demonstrated at home  
   (6) Open counter display

   Column B  
   (a) Mail order  
   (b) Door-to-door  
   (c) Vending machine  
   (d) Department store  
   (e) Specialty  
   (f) Used clothing

**ESTABLISHING CONTENT VALIDITY:** (...does the test measure what it is supposed to measure?)

1. Make sure that the test matches the content taught and its relative emphasis.

2. Use a matrix with content breakdown on one axis. Use low- and high-level thinking on the other axis. Weight the distribution of items according to the learner's time-on-task or importance.

<table>
<thead>
<tr>
<th>Program Content</th>
<th>Thinking Level</th>
<th>Low</th>
<th>High</th>
<th>TOTALS Test Items No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit No. 1</td>
<td></td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>Unit No. 2</td>
<td></td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Unit No. 3</td>
<td></td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Unit No. 4</td>
<td></td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>35</td>
</tr>
<tr>
<td><strong>TOTALS</strong></td>
<td></td>
<td>20</td>
<td>20</td>
<td><strong>40</strong></td>
<td>100</td>
</tr>
</tbody>
</table>

**ESTABLISHING RELIABILITY:** (...does the test yield consistent results?)

1. Use at least 25 test items for any one test.

2. Use as many items as possible, being careful that the test time does not become unreasonable.

3. Eliminate items that all test takers get wrong or right.

**SUGGESTED REFERENCES:** Gronlund, N.E. (1988); Parker, B. (1986); Phillips (1994)
Figure 6. LEARNING - Expertise Test Job Aid

TYPES OF TESTS OF EXPERTISE (process and product)

1. Process Check Sheet (samples):
   - Course: Basic Tools and Hardware --> Expertise Check list --> Terminal Learning Objective:
     - Task: Fasten the hardware to the sub-assembly workpiece
     - Conditions: Given a torque screwdriver, Phillips head tip, no. 10-32 x 1/2 Philips pan head screw, no. 10 flat washer, no. 10 hex nut, subassembly workpiece
     - Standard: Meet the torque specification of 58 inch/pounds

<table>
<thead>
<tr>
<th>Critical Item</th>
<th>Procedural Steps</th>
<th>Satisfactory</th>
<th>Unsatisfactory</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Determine torque specification</td>
<td>1</td>
<td>0</td>
<td>specification will be 58 inch/pounds</td>
<td></td>
</tr>
<tr>
<td>2. Locate the adjustment knob</td>
<td>1</td>
<td>0</td>
<td>at the butt end of torque screwdriver handle</td>
<td></td>
</tr>
<tr>
<td>3. Locate torque indicator line</td>
<td>1</td>
<td>0</td>
<td>transparent plastic collar at midpoint of torque screwdriver</td>
<td></td>
</tr>
<tr>
<td>4. Turn adjustment knob</td>
<td>1</td>
<td>0</td>
<td>clockwise/counter clockwise to 50 inch/pounds</td>
<td></td>
</tr>
<tr>
<td>5. Line up torque indicator line</td>
<td>1</td>
<td>0</td>
<td>at the 58 inch/pounds setting</td>
<td></td>
</tr>
<tr>
<td>6. Lock the adjustment knob</td>
<td>1</td>
<td>0</td>
<td>adjustment knob locked in place</td>
<td></td>
</tr>
<tr>
<td>7. Insert Phillips head tip into position</td>
<td>1</td>
<td>0</td>
<td>at the end of torque screwdriver opposite adjustment knob</td>
<td></td>
</tr>
<tr>
<td>8. Position screw</td>
<td>1</td>
<td>0</td>
<td>into no. 10 predrilled hole in subassembly workpiece</td>
<td></td>
</tr>
<tr>
<td>9. Position flat washer</td>
<td>1</td>
<td>0</td>
<td>over shank of prepositioned screw in step 8</td>
<td></td>
</tr>
<tr>
<td>10. Position hex nut</td>
<td>1</td>
<td>0</td>
<td>over shank of prepositioned screw in step 8</td>
<td></td>
</tr>
<tr>
<td>11. Fasten hardware in place</td>
<td>1</td>
<td>0</td>
<td>finger tightened</td>
<td></td>
</tr>
<tr>
<td>12. Position torque screwdriver</td>
<td>1</td>
<td>0</td>
<td>tip of torque screwdriver securely seated in screw head</td>
<td></td>
</tr>
<tr>
<td>13. Tighten screw</td>
<td>1</td>
<td>0</td>
<td>until locking sound occurs</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total score = 13

Minimum mastery level score = 13

2. Product Specifications (samples):

<table>
<thead>
<tr>
<th>Amperage Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 10 8 6 4 2</td>
</tr>
</tbody>
</table>

Directions: Rate the instrument according to the following criteria by placing an "X" in the appropriate blank.

- **Criterion** | **Description** | **Yes-No**
- Quality | Does it measure quality of the performance? | ____ |
  - Skill | ____ |
  - Attitude | ____ |
- Efficiency | Does it measure efficiency of the operation? | ____ |
- Ease of use | Does the language, design, length and degree of detail promote ease of use? | ____ |
- Achievement of goals | Does it achieve the goals of measuring student progress, diagnosing, certifying and evaluating instructions? | ____ |
- Adaptability | Does it serve, with little revision, for self-evaluation, peer evaluation and supervisor evaluation? | ____ |
- Validity | Does it measure what it was designed to measure? | ____ |
- Reliability | Does it provide trustworthy or essential measures? | ____ |

Recommendations for change: ______________________________________________________________________________

CONTENT VALIDITY: (does the test measure what it is supposed to measure?)
1. Make sure that the process check sheet contains all the critical steps specified by the work behavior analysis.
2. Make sure that all the product specifications, quality and quantity, are included in the evaluation criteria.

RELIABILITY: (does the test yield consistent results?)
1. Have trainee exhibit the process at least twice and produce at least two products.
2. If No. 1 is not possible, have trainee talk through the process while doing it or describe the specifications to ensure correct rating.

Figure 7. SATISFACTION - Participant Satisfaction Form

Program Title ___________________________ Date ________________
Location __________________________________________
Presenter(s) __________________________________________

Please answer the following questions to help us improve future programs.

<table>
<thead>
<tr>
<th>Question</th>
<th>Very</th>
<th>Good</th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My ability in the content area of this program.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. My level of knowledge/expertise in this area prior to the program.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. My motivation to learn this new material.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. My supervisor's encouragement in learning this material.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Quality of instructor presentations.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Quality of the information presented.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Amount of time to practice new material.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Quality of feedback on my performance during the program.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Quality of physical environment.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. The expectation that I will use this material on the job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. The use of my time in participating in this program.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The most valuable part of this program: ______________________________________

The least valuable part of this course for you? _______________________________

If you rated any item "poor", please provide some additional explanation ________

Additional comments: ________________________________________________________

Participant Name (optional) ________________________________
Figure 8. SATISFACTION - Sponsor Satisfaction Form

Program Title

Dates of Program

Participant(s):

Sponsor/Supervisor Completing this Form:

Date: 

Please answer the following questions to help us improve future programs.

Now that your direct report has completed the development program and is back on the job, what is your impression of the effectiveness of the program?

<table>
<thead>
<tr>
<th></th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Participants have performed better on the job following the development program.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Attending the development program was a good use of the participant's time.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additional comments:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
EVALUATION REPORT
(Company) Healthcare Sales Performance Consulting

Program: “Communicating the (Company) Advantage”
Dates: January - September, 199X, 24 Groups
Participants: 180 Healthcare Sales/Account managers

Program Purpose:
When a (Company) employee meets with a potential or existing customer, they create a professional impression, build a credible, trustworthy relationship, and communicate the value of (Company) with a direct emphasis on our medical management capabilities.

Program Description:
“Communicating the (Company) Advantage” is an intensive two-day program that focuses on listening, questioning, and presentation skills. It is designed to improve sales associates’ ability to communicate the “medical management capability” in a way that improves sales results. During the two-day program, sales managers co-facilitate by providing technical expertise, and serving as a communications coach. Sales managers also are required to go through the program and to coach and evaluate sales associates back on the job.

Evaluation Summary:
In summary, the 199X, “Communicating the (Company) Advantage” programs were very effective. The program exceeded its goals in all of the areas evaluated thus far, ten months following the first program. See page two for a further breakdown of these results and contact Sales Performance Consulting if you wish additional information.

<table>
<thead>
<tr>
<th>Performance</th>
<th>Goal Attainment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Results</td>
<td>409%</td>
</tr>
<tr>
<td>Learning</td>
<td></td>
</tr>
<tr>
<td>Expertise</td>
<td>117%</td>
</tr>
<tr>
<td>Satisfaction</td>
<td></td>
</tr>
<tr>
<td>Participant Satisf</td>
<td>149%</td>
</tr>
<tr>
<td>Sponsor Satisf</td>
<td>147%</td>
</tr>
</tbody>
</table>

Approval: V.P. of Sales Performance Consulting Date: 10/26/9X

Distribution List:
- President
- Sr. V.P. of Sales & Mkg.
- V.P. of National Accounts
- Sr. V.P. of HR
- V.P. of HRD
- Area Operating Officer

“Communicating the (Company) Advantage” page 1 of 5
EVALUATION RESULTS: 1995 Communicating the (Company) Advantage Program

The evaluation of this program is reported according to the effectiveness domains of performance, learning, and satisfaction (PLS).

PERFORMANCE

The overarching goal of this program was to increase sales through communication and presentation skills. Listed below is a summary of sales attributed directly to this performance improvement program.

<table>
<thead>
<tr>
<th>Business Results by Type of Sale</th>
<th>Financial Results in Terms of Total Premium Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Markets “Y” Business</td>
<td>$5,500,000</td>
</tr>
<tr>
<td>1 Dental</td>
<td>290,000</td>
</tr>
<tr>
<td>2 PPO Firms</td>
<td>2,000,000</td>
</tr>
<tr>
<td>18 Markets “X” Business</td>
<td>57,620,000</td>
</tr>
<tr>
<td>1 National Account</td>
<td>13,000,000</td>
</tr>
<tr>
<td><strong>24 Sales</strong></td>
<td><strong>Totals</strong> $78,410,000</td>
</tr>
</tbody>
</table>

The business results reported at this point are taken from 48 responses as to sales attributed to the program. These responses come mostly from sales personnel that attended the program; however, some managers responded on behalf of their whole office. It is important to note that sales data has not been submitted by all offices.

The approximate profitability of the communication program is as follows: The average value of a single reported sale is $3,267,083; the average net profitability on a single sale is 3.0%, or $98,012/sale; the total profit for this group of sales $2,352,300.

The direct program cost per group is approximately $12,000, or $288,000 for all twenty-four groups. The approximate return on investment (ROI) of this program is equal to the total profit divided by the total cost of the programs. The financial results goal is a 2:1 return on investment (ROI) within 12 months. The financial results obtained were 8.17:1 ROI, or 409% of goal, for this 9 month period. The attached exhibit provides more detail on sales attributed thus far to this communication program.

If we assume the goal for the “business results” is to increase average sales by 5% per program participant per year, business results will be based on an increase in the number of target market firms sold per rep per year. The business results will not be available until 12 months following the program when all sales have been accounted for.

A. Business Results

24 sales reported thus far, totals not available at this point in time

B. Financial Results

"Communicating the (Company) Advantage" page 2 of 5

© 1996 Swanson
EVALUATION RESULTS: 1995 Communicating the (Company) Advantage Program

While an approximation of ROI is a valuable measurement, the comments of participants add an additional dimension to the financial results. One sales manager's comments are listed below:

- "I am certain it [COMM I Program] has made a material difference in our results here in Boston. We have been awarded at least 5 substantial pieces of business since the initial training took place where the decision was made on the strength of our oral presentation. The five wins were as follows: Company A, $20 mil; Company B, $11 mil; Company C, $3 mil; Company D, $3 mil; Company E, existing client conversion to managed care.
- In the case of Companies A, B, and E our competitor was Company X. In each case it came down to the oral presentation, as the metrics/financials for us and Company X were virtually the same. Preparation for the presentations was done using a lot of the principles and the process outlined in the communications training."
  - (Name), Sales Manager, Boston

LEARNING

Each participant was tested on “communication expertise.” Managers used standardized rating forms to evaluate the expertise of participants on a multi-dimension, 3 point per dimension scale. The goal was a “meets clients’ expectations” rating of 2 for each dimension. The average expertise for all groups immediately preceding training was 1.71, or 86% of goal. The average expertise for all groups at the end of the training was 2.12, or 106% of goal. The average expertise for all groups 60 days following training and coaching was 2.33, or 117% of goal. The overall improvement in communication expertise was 31% over the period. No paper and pencil “knowledge test” was used based on the assumption that the knowledge was required to succeed on the communication test of expertise.

C. Knowledge

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D. Expertise

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Training</td>
<td>0%</td>
<td>20%</td>
<td>40%</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
<td>120%</td>
<td>140%</td>
<td>160%</td>
</tr>
<tr>
<td>End of Training</td>
<td>0%</td>
<td>20%</td>
<td>40%</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
<td>120%</td>
<td>140%</td>
<td>160%</td>
</tr>
<tr>
<td>60 Days After Training</td>
<td>0%</td>
<td>20%</td>
<td>40%</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
<td>120%</td>
<td>140%</td>
<td>160%</td>
</tr>
</tbody>
</table>

"Communicating the (Company) Advantage" page 3 of 5

© 1996 Swanson
EVALUATION RESULTS: 1995 Communicating the (Company) Advantage Program

SATISFACTION

Participants and participant managers rated their satisfaction with the communication program. The responses of 167 participants and their supervisors are reflected here. The goal for both measures was to have a positive average satisfaction rating of at least 2.5 on a 1-4 scale. The “participant satisfaction” results were 3.72, or 149% of goal. The “sponsor satisfaction” results were 3.68, or 147% of goal.

Participants in all areas expressed overwhelming enthusiasm and satisfaction with the program. Listed below are some of their comments:

- “Every (Company) employee should be subject to at least some aspects of this seminar.”
  - (Name), New York City Sales

- “As a Client Manager, I do not have the same opportunity to use my communication skills in selling “new” cases. Instead, I use it in selling renewals, in selling additional ‘Y’ business, in solving customer issues, and working with my internal customer’s to resolve our clients’ problems.”
  - (Name), Chicago Sales

- “I appreciate my company investing in me through this program.”
  - (Name), New York City Sales

- “This is a great program! There is nothing more important than communicating the vision!”
  - (Name), Jacksonville Sales

- “It was definitely time well spent. I view it as an investment into our success.”
  - (Name), Chicago Sales

- “I am absolutely convinced that as a result of the CCA and follow-up training with Stump Speeches that my people are more impactful—that is they can communicate our story more convincingly, more consistently than ever before. I have no doubt that those skills have impacted our ability to sell in the marketplace. On absolutely every case we sold the skills training leveraged our ability to sell the case. Similarly, the training leveraged our ability to avoid losing our cases.”
  - (Name), Sales Manager, Florida

Suggestions for Improvement
The key suggestion for improvement is to follow up more vigorously on the utilization of the coaching form by the sales managers.
EVALUATION RESULTS: 1995 *Communicating the (Company) Advantage* Program

Exhibit A.

SALES ATTRIBUTED TO COMMUNICATION SKILLS I PROGRAM

<table>
<thead>
<tr>
<th>Number of Sales</th>
<th>Name of Sales Associate</th>
<th>Name of Account(s)</th>
<th>Premium Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 PPO Firms</td>
<td>* NYC</td>
<td>Account A**</td>
<td>$1,000,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account B</td>
<td>$1,000,000</td>
</tr>
<tr>
<td>4 Target Markets</td>
<td>Virginia</td>
<td>Account C</td>
<td>$900,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account D</td>
<td>$400,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account E</td>
<td>$170,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account F</td>
<td>$400,000</td>
</tr>
<tr>
<td>4 Target Markets</td>
<td>Boston</td>
<td>Account G</td>
<td>$3,000,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account H</td>
<td>$3,000,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account I</td>
<td>$20,000,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account J</td>
<td>$11,000,000</td>
</tr>
<tr>
<td>1 Dental</td>
<td>Philadelphia</td>
<td>Account K</td>
<td>$290,000</td>
</tr>
<tr>
<td>2 Target Markets</td>
<td>Carolina</td>
<td>Account L</td>
<td>$3,500,000</td>
</tr>
<tr>
<td>&quot;Y&quot; Business</td>
<td></td>
<td>Account M</td>
<td>$2,000,000</td>
</tr>
<tr>
<td>2 Target Markets</td>
<td>Great Lakes</td>
<td>Account N</td>
<td>$500,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account O</td>
<td>$350,000</td>
</tr>
<tr>
<td>1 Target Market</td>
<td>Chicago</td>
<td>Account P</td>
<td>$2,500,000</td>
</tr>
<tr>
<td>4 Target Markets</td>
<td>Miami</td>
<td>Account Q</td>
<td>$8,400,000 (combined)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account R</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account S</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account T</td>
<td></td>
</tr>
<tr>
<td>1 National Account</td>
<td>Carolina</td>
<td>Account U</td>
<td>$13,000,000</td>
</tr>
<tr>
<td>2 Target Markets</td>
<td>Tennessee</td>
<td>Account V</td>
<td>$3,000,000 (combined)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account W</td>
<td></td>
</tr>
<tr>
<td>1 Target Market</td>
<td>Birmingham</td>
<td>Account X</td>
<td>$4,000,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total Value</td>
<td>$78,410,000</td>
</tr>
</tbody>
</table>

* The salesperson reporting the specific sale(s) are listed.
** The actual company names (accounts) are listed.