ESTABLISHING TRAINING CRITERIA ON AN ECONOMIC BASIS

STEVEN L. JOHNSON
Calspan Corporation

INTRODUCTION

The Problem

The training community has toiled for many years in an attempt to establish exactly what is meant by the phrase "training effectiveness." Although no one can precisely define it, the concept of effective training is of interest to psychologists and managers alike. The question is continuously being posed by managers: "is our training program cost-effective?" Subsequently, training psychologists begin expounding on the multitude of factors that comprise the known principles of learning (reinforcement, habit patterns, immediate feedback, etc.). At the end of the discussion, the managers conclude that they know (viscerally) that training is necessary but they recognize that the psychologists cannot convince each other, let alone "laymen." Thus, the schism between the "shrink" and the "bean counter." The goal of the present paper is to develop a quantitative and, more importantly, a communicable framework for establishing training effectiveness criteria that can be used and understood by both the training specialist and the manager.

Training Decisions

In the process of developing a training program, an obvious question relates to which behaviors (tasks) will be trained and which will not. That is, for any nontrivial operator's job, it is impossible to exhaust the possible contingencies that might arise. In fact, humans are often included precisely to utilize their conceptual generalization capabilities. The only point here is that one decision that must be made is whether or not to train on a particular task. To date, there has been no algorithm established to assist in this decision.

A second decision which is an extension of the same issue is, if a task is to be trained, what level of proficiency should be required? That is, what should be the training duration (which the manager translates into dollars). It is well known that the rate of learning exhibits diminishing returns as a function of training time.

A third question, that has become prominent in military training, is related to the complexity (cost) of training devices. As with the relation between learning rate and time, there are diminishing returns of learning rate as the complexity of the training device surpasses a particular level. Figure 1, illustrates the relationship of learning rate and device complexity (cost). The "bean counter" is obviously interested in establishing where the additional cost of training (time and/or devices) is not warranted. This paper will present an approach to establishing the required information necessary to make the above decisions and a quantifiable format of this information.

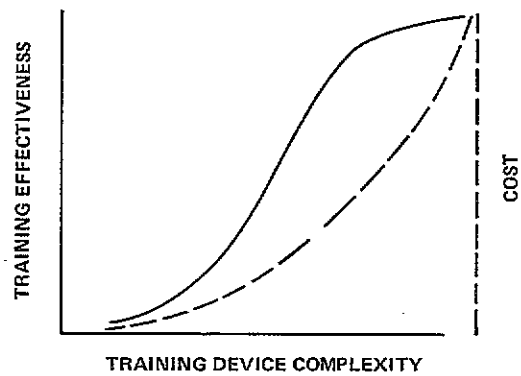

Figure 1. Cost and Training Effectiveness as a Function of Training Device Complexity

The Quality Control Aspect of Training Criteria

Training programs can be viewed as means of increasing the reliability (i.e. quality) of one of the components in the man-machine system. Therefore, the "training criteria" (time, performance levels, etc.) are analogous to the quality levels utilized in industrial quality control schemes. Therefore, the factors that are involved in establishing quality levels in a Bayesion economic analysis model (costs, payoffs, and probabilities) are appropriate to establishing the cost-effectiveness of training criteria on the basis of a Loss Function. The goal of this approach is to determine the out-going level of the trainee on the basis of minimizing total costs that relate to that quality level.

METHODOLOGY

The general approach discussed in this paper involves viewing training as a means of increasing the reliability (quality) of the human operator. The training criterion (outgoing quality) is established within the context of decision theory parameters (costs, payoff, and probabilities).

Increasing Reliability Through Training

Although it is seldom discussed in this context, training is simply a means of increasing the reliability of one of the components in a man-machine system. In the same sense that increasing the quality of any other component (in terms of material or workmanship) increases reliability, the quality of the human component is increased through training. There are both advantages provided by the increased reliability and costs incurred by the increase.

Training criteria are the means by which the quality of the human product is assured. In the same sense as the producers' and consumers' risk are variables in a quality control scheme, the amount of risk (cost) involved with undertraining or overtraining a human is also a variable.

The "natural" quality of the process is an important aspect to consider in quality control schemes. Similarly, the "natural" reliability of the human for various tasks is important in establishing the training criteria. There is an extensive amount of research literature on the topic of human reliability. This information is often "descriptive" of the human (e.g., Askren and Regulinski, 1969; and Beck, Hayman, and Markisohn, 1967); but it also includes information that is "prescriptive" in nature (Feherman and Siegel, 1973; Siegel, Wolf, and Lautman, 1974; and Sontz and Lamb, 1975). These studies have resulted in an extensive methodology for assessing human reliability; however, there is little direct information as to increases in human reliability through training (Goldman and Slattery, 1964). This relationship must be inferred from the reliability data in combination with information about learning.

Learning Rate

The literature on reliability and human performance provides extensive data and a methodology for estimating human performance parameters. That is, the range of the potential "outgoing qualities" for various types of tasks can be, and in some cases has been, established (within rather large bounds). The question remains, however, as to the rate of quality increase (learning) as a function of time. In fact, there is a significant amount of information pertaining to learning rates for both discrete-serial and continuous perceptual-motor tasks (e.g., Fleishman, and Parker, 1962; and Neumann and Ammors, 1957). For various types of tasks, the learning rate versus time function (the learning curve) can be specified with relative confidence. In fact, the stability of the learning curve has resulted in severe difficulties in demonstrating effects due to different "training techniques." That is, for a given task, trainees learn "in spite of the training technique." For the purposes at hand, this is a significant advantage. Figure 2 illustrates the general shape of the learning curve.

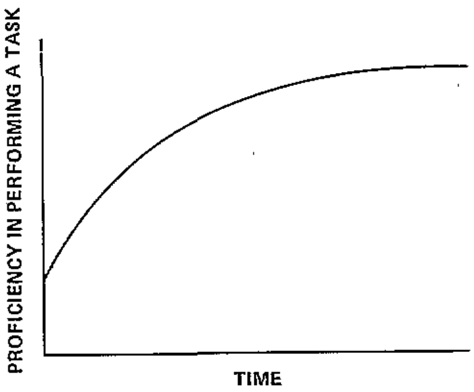

Figure 2. Typical Learning Curve

Information Requirements

As discussed in the previous sections, there is, to a greater or lesser degree, an extensive literature base pertaining to human reliability and learning rate as a function of time. The present section addresses the other information that is required to establish an outgoing quality level of a training program on the basis of economics.

The first (and in the extreme, the most obvious) parameter is the probability that the operator will ever perform the particular task. For example, emergency procedures to handle abnormal situations often occur with a relatively low probability. Intuitively (and pragmatically) there is a low probability of occurrence beyond which training on that task would not be "justified." Similarly, on the other end of the continuum, for a task that occurs routinely, training could probably be "justified." The point is that (as with the relationship between the probability of a defect and the requirements for a quality control scheme) the existence or extent of training on a task is affected by the probability of that task occurring.

A second set of parameters that is required for an economically based decision is that of costs. There are two general categories of cost. One category involves the costs of training. These costs include personnel costs (trainees, instructors, managers), training equipment procurement costs (from textbooks to complex system simulators), and operation and maintenance costs (including those associated with facilities). Another category of costs that is important in establishing training criteria involves the cost of a failure to perform the operational task. Included in this category are such things as costs due to waste when an operator allows a process to go out of control, equipment damage, and loss of life or personal injury.

Obviously, some of these costs are more easily arrived at than others. For example, the "cost of human life or injury" is potentially very difficult to ascertain. However, in terms of law suit settlements, these numbers are being established. In addition, tacticians have been establishing equally difficult costs in war-game simulations that are used to evaluate a weapon system's effectiveness. The point to be made here is that, although assessing (or at least asserting) costs is often difficult, it is possible. A second point is that "establishing" these costs is not within the technical area of the training specialist and cannot, therefore, be his responsibility. This is analogous to the costs in the economic decision process of establishing outgoing quality levels in any quality assurance situation. It cannot be the quality control specialist's responsibility to determine the costs of "unacceptable" quality.

The last set of parameters that impacts upon training criteria decisions involves the payoffs afforded through training as a function of time. That is, what are the probabilities that the human will "handle" the situation given that the situation does occur. Similarly, how does this probability vary as a function of training time. These questions relate to the previously discussed human reliability estimates and the learning curve. Whereas, the probability of a situation occurring and the costs involved are the responsibility of managers (with the counsel of experts in reliability and economics), this aspect of the information necessary for a decision is within the technical expertise of the training (human performance) specialist. Now that the categories of information have been established, the following section of this paper will address the computation involved in the decision "process."

COMPUTATIONS

The purpose of this section is to present a "first-cut" algorithm that is potentially useful in making training criteria decisions. First, it is necessary to establish the notation that will be used and define the meaning of the parameters.

- P_{occ} - probability that the situation will occur.
- P_o - probability of performing the task without training, given that the situation occurred.
- P_t - probability of performing the task after training time t, given that the situation occurred.
- $K_{\Delta t}$ - cost of a unit of training time.
- K_{occ} - cost of a failure to perform the task.

The increase in the probability of the human to perform a task after training time t is:

$$(P_t - P_o)\,(P_{occ})$$

The pay-off value contributed by the training is:

$$(P_t - P_o)\,(P_{occ})\,(K_{occ})$$

The cost of training for a period of time t is:

$$(K_{\Delta t})\,(t)$$

Therefore, "the break-even point" in terms of the training criterion is when the pay-off value of the training is equal to the cost of the training. That is, when:

$$\frac{(P_t - P_o)\,(P_{occ})\,(K_{occ})}{(K_{\Delta t})\,(t)} = 1 .$$

When the left side of this equation is greater than 1, additional training is cost-effective; when it is less than 1, it is not.

Manipulating this equation results in:

$$\frac{(P_t - P_o)}{t} = \frac{(K_{\Delta t})}{(P_{occ})\,(K_{occ})} .$$

As discussed in the previous section, providing the data for the right-hand side of the equation is the responsibility of the combination of manager, reliability expert, and the economist. The information pertaining to the right-hand side is the responsibility of the training specialist.

Recall, from the previous section that the increment in rate of learning is not a *linear function of training time. Therefore,* the optimization process of the decision involves increasing training time (t) to the point where the equation is satisfied. Figure 3, graphically illustrates this procedure.

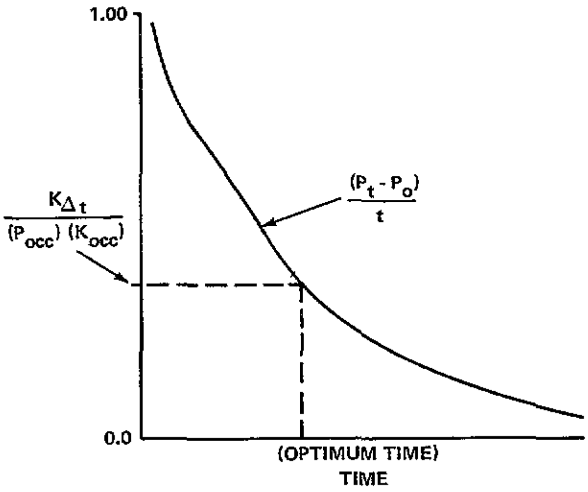

Figure 3. Function Relating Training Decision Parameters

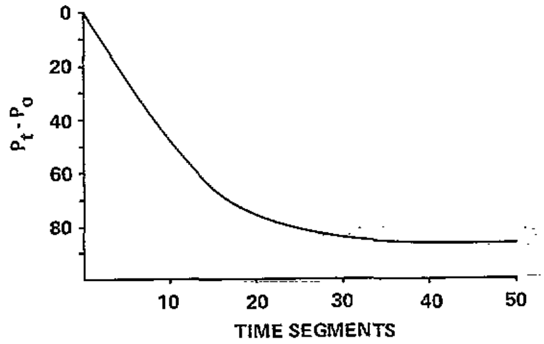

Figure 4. Learning Curve Adapted from Fleishman and Parker (1962)

Table 1

$P_t - P_o$	t	$\frac{P_t - P_o}{t}$
0	0	.00
.25	5	.050
.49	10	.049
.66	15	.044
.76	20	.038
.84	30	.028

HYPOTHETICAL EXAMPLES

For the purpose of illustrating the method, a simple example will be presented. Assume that a training program is being initiated in a widget factory. There is a particular situation that can lead to a malfunction that approximately one out of twenty operators experience (e.g.,average tenure in the job of 5 years, probability of situation occurring is 0.01 during any particular year). If the situation occurs, the operator must perform a series of steps in a procedure within a short period of time. If he fails, the widget-making machine explodes with a replacement cost of \$16,000. The total cost of one unit of training time is \$35.00. In addition, assume that the learning rate is approximated by the curve found in the study by Fleishman and Parker (1962). The question is: "How long should the operator be in training?" The parameter values for this example are:

$K_{\Delta t} = 35$; $K_{occ} = 16{,}000$; and $P_{occ} = .05$.

Figure 4 illustrates the function, $P_t - P_o = f(t)$ and Table 1 gives discrete values of $P_i - P_o$, t, and $(P_t - P_o)/t$.

The calculation results in:

$$\frac{K_{\Delta t}}{(K_{occ})\ (P_{occ})} = .0438.$$

Therefore, it can be determined from the table that approximately 15 time units of training is optimal.

A second example, in a different context, is an overcoat manufacturing factory in which there is a training program for seamstresses. In this case, the P_{occ} is the probability that the customer will decide upon buying the coat on the basis of the sewing (as opposed to the material or design). The K_{occ} is the lost revenue if the customer does not by the coat. In this case, however, the P_t is the probability that the customer will accept the coat, given that the sewing is one of his decision points. Through training, the P_t value increases as in any other learning situation. Given the required cost data, probabilities, and learning curve, the extent of training that is cost-effective can be determined.

The purpose of the present paper, and the examples given, is to demonstrate the feasibility of approaching training requirements in terms of decision theory parameters. The next section discusses the next steps in refining this approach.

NEXT STEPS

An integral part of the approach described in this paper is the use of the learning curve characteristics in determining training criteria. It was stated earlier that we have an extensive amount of information pertaining to the shape of the learning curve for many tasks. However, what is needed for this approach (as well as many other aspects of human performance) is a descriptive "taxonomy of tasks" that vary in their learning characteristics. This taxonomy could be used by training learning specialists to estimate the function $P = f(t)$ for the tasks that make up their particular operational job. There have been selected efforts in the area of task taxonomy (e.g. Miller, 1971). In fact, Meister and Mills (1971) and Mills, Bachert, and Hatfield (1975) have categorized tasks in terms of human reliability. Swain (1963, 1974) has also demonstrated this possibility. Although this work does not provide learning curve information, it does serve to set the boundaries in terms of human performance reliability.

The actual calculations required by this methodology could be significantly simplified by applying an appropriate transformation to the curve that would result in a linear relationship. For example, a logarithmic transformation is the obvious starting point for this type of effort.

In addition to increasing the confidence in the learning data and simplifying the computation requirements, a major "next step" is to determine the sensitivity of the various parameters. This type of analysis will allow one to allocate his time appropriately in terms of refining the information to which the cost-effectiveness is most sensitive.

IMPLICATIONS OF THE METHODOLOGY

One of the more interesting implications of the proposed methodology is that it illustrates the absurdity of the "perfect mastery" approach on a cost-effectiveness basis. That is, on the basis of costs and payoffs, it is the rate of learning as a function of time that is the important metric. The appropriate criterion for terminating training for a particular individual is when his increased proficiency gain per unit time (i.e., dollars) is less than the payoff contributed to operational success. This approach is compatible with the techniques of Computer Managed Instruction (or adaptive training in perceptual-motor skill training) in that the rate of proficiency increase is accessible. The point is that an individual's present proficiency is not the only criterion variable; it is also important to evaluate his rate of advancement and subsequently estimate his potential for further advancement.

Another important implication of this methodology is the categorization of the types of information necessary to make valid training system design decisions and the assignment of responsibilities for the various information. Decisions relative to training criteria are the joint responsibility of operational personnel, managers, and training specialists.

The problems of establishing training criteria are very complex; however, there must be a framework from which to start if progress is to be made. The present paper discusses the applicability of techniques presently developed in the areas of quality assurance, training technology, and decision theory to the problem of establishing cost-effective training criteria. The types of data required, and the appropriate responsibilities for those data are discussed. There is no apparent reason why the cost-effectiveness of training cannot be established through the same methods as have been used for other systems for many years and with significant methodological advances.

REFERENCES

Askren, W.B., and Regulinski, T.L. Quantifying human performance for reliability analysis of systems. Human Factors, 1969, 11 (4), 393-396.

Beek, C., Hayman, L., and Markisohn, G. Human reliability research. NPRDC - TR - 430, 1967.

Feherman, P.J., and Siegel, A.I. Prediction of human reliability. Part II: Validation of a set of human reliability predictive techniques. Wayne, PA: Applied Psychological Services, 1973.

Fleischman, E.A., and Parker, J.F., Jr. Factors in the retention and relearning of perceptual-motor skill. J. of Exper. Psychol., 1962, 64 (3), 215-226.

Goldman, A.S., and Slattery, T.B. Maintainability: A Major Element of System Effectiveness. New York: Wiley, 1964.

Meisler, D., and Mills, R.G. Development of a human performance reliability data system: Phase I. USAF, Aerospace Medical Research Laboratory Report, AMRL-TR-71-87, 1971.

Miller, R.B. Development of a taxonomy of human performance: A user-oriented approach. American Institute for Research, AIR-R71-8, 1971.

Mills, R.G., Bachert, R.F., and Hatfield, S.A. Quantification and prediction of human performance: Sequencial task performance reliability and time. USAF, Aerospace Medical Research Laboratory Report, AMRL-TR-74-48, 1975.

Neumann, E., and Ammons, R.B. Acquisition and long-term retention of a simple series perceptual-motor skill. J. Exper. Psychol., 1957, 53 (3), 159-161.

Siegel, A.I., Wolf, J.J., and Lantman, M.R. A model for predicting integrated man-machine systems reliability. Wayne, PA.: Applied Psychological Services, 1974.

Sontz, C., and Lamb, J.C. Predicting system reliability from human data. Proceedings of the 1975 Annual Reliability and Maintainability Symposium, IEEE No. 75, CHO 918-3 RQC, 105-109.

Swain, A.D. Method for performing a Human Factors reliability analysis. Sandia Laboratory Report, SCR-685, 1963.

Swain, A.D. Human Factors associated with prescribed action links. Sandia Laboratories Report, SAND-74-0051, 1974.

ABOUT THE AUTHOR

MR. STEVEN JOHNSON is a senior psychologist and project manager in the Human Factors Section of Calspan Corporation where he is developing a training system for the Air Force's B-1 strategic bomber aircrew. He did research in visual detection, modeling, training techniques, and training system development. He was task leader for training conducted for the Navy's E-2C early warning aircraft. Prior experience at Aviation Research Laboratory included work on controls, displays, and design of cockpit layouts. Academic and research background in experimental psychology included visual and auditory perception, information processing, and sensory physiology. He received the M.S. degree at University of Illinois Aviation Research Laboratory and is doing research for the Ph.D. dissertation in Human Factors at State University of New York at Buffalo.