20

Mining Human Performance Data

David A. Nembhard
University of Wisconsin-Madison

Introduction and Overview	463
Mining for Organizational Learning	464
Methods	464
Individual Learning	467
Data on Individual Learning	468
Methods	468
Individual Forgetting	474
Distributions and Patterns of Individual Performance	474
Other Areas	476
Privacy Issues for Human Performance Data	477
References	477

INTRODUCTION AND OVERVIEW

The mining of human performance data has potential for wide-reaching benefits in manufacturing and service organizations, for researchers and human behavior specialists. Until recently the vast majority of research in these areas was based on relatively small controlled experiments and resultant data. With the rapid growth of data acquisition systems in the workplace, there are now massive stores of data that have potential for use in addressing many existing research questions and in informing decision-making processes in these organizations themselves. The primary purpose of many of these data acquisition systems is often mundane, such as barcode tracking of raw materials and products, or payroll accounting systems that keep track of employees' working hours. However, the sheer volume of data collected often creates an opportunity to mine these records as secondary data. Much as the data from point-of-sale systems are transforming marketing research, the records from bar-code scanners and networked systems offer vast new repositories of information on individual and group performance. As will be illustrated later in this chapter, a common key component of human performance data mining is the availability of temporal databases related to human work or activities.

The sections that follow describe several useful approaches for mining data related to human productivity. These approaches include measurements of organizational improvement and learning, individual learning, forgetting, human computer interaction, and cognitive science data.

MINING FOR ORGANIZATIONAL LEARNING

Scholars have studied learning behavior in organizations for more than a century. Important contributions have been made by behavioral psychologists studying the individual learning process, industrial engineers studying manufacturing costs, systems theorists studying organizational dynamics, and artificial intelligence scholars modeling the human thought process. More recently behaviorists have added important insights into corporate culture and mechanisms for promoting and managing change.

The use of learning curves for measuring and improving efficiency and cost has been demonstrated in numerous manufacturing and service organizations. Effective organizations use their experience to "tune" operations to do the same work at a lower unit cost. Observers of aircraft manufacture in the 1930s noted that "as the quantity of units manufactured doubles, the number of direct labor hours it takes to produce an individual unit decreases at a uniform rate" (Yelle, 1979). Unfortunately, the observed cost reductions result from a complex mix of individual learning, technology changes, economies of scale, and changes in management systems (Dutton & Thomas, 1985). Thus, the internal structure and causal mechanisms that affect organization-level learning curves are still not well understood (Argote, 1993).

Various researchers and practitioners have used systems analysis, case studies, and observation to find ways for companies to increase the rate of organizational learning. This "action research" agenda tries new approaches and observes their effects on organizational processes and outcomes (Argyris, 1989). These efforts have created a strong interest in helping employees to "learn how to learn" (Argyris, 1982; Senge, 1992). An increase in learning rates may provide benefits in terms of cost reduction, productive output, and efficiency. The measurement of organizational learning is also useful for cost estimation and production planning.

A second stream of work has studied the economic impact of midlevel organizational learning. They show how diversity in learning behavior can exert a complex influence on overall productivity (Adler, 1990; Kantor & Zangwill, 1991; Meredith & Camm, 1989). Both types of intermediate level research have focused more on cognition and problem solving than on motor learning and skill acquisition. They reflect an implicit assumption that learning in professional and managerial ranks will have a greater impact on organizational and economic performance. In particular, scholars want to help organizations foster "double loop learning," in which an organization's decision-makers learn how to learn more effectively (Argyris, 1982).

Methods

Organizational learning (productivity improvement) has been studied from a variety of perspectives. Most commonly, the overall productivity of the organization has been examined over a time scale to gauge the rate and overall effect of increased productivity with increased organizational experience. Methods for achieving this have often centered on linear and nonlinear regression. Example 1 illustrates a straightforward approach using linear regression.

Example 1: Regression Based Organizational Learning Model. This example describes a linear regression based model of organizational learning adapted from Epple, Argote, and Devadas (1991), wherein the conventional learning curve is written in the form given in Equation 1, where y denotes output, x is cumulative output (as a surrogate for acquired experiential knowledge), T is the time worked, with C and r parameters related to the organization's characteristics. The larger the parameter r, the more rapidly productivity increases due to learning. Parameter C in this formulation is a time constant.

$$\frac{y}{T} = Cx^r \tag{1}$$

Because organizational data generally is stored based on a discrete time increment (e.g., hours, days, weeks), Equation 1 may be rewritten in a form suitable for estimating an organization's parameters given such discrete data, where t is a discrete time increment. Thus, Equation 2 allows for the prediction of an organization's output at time t, y_t, given cumulative experience up to the previous time increment, x_{t-1}.

$$\frac{y_t}{T_t} = Cx_{t-1}^r \tag{2}$$

To linearize this expression so that standard linear regression methods may be used, logarithms are taken of Equation 2, letting $c = Ln(C)$ and yielding a form more convenient for estimation as shown in Equation 3.

$$Ln(y_t/T_t) = c + r\,Ln(x_{t-1}) \tag{3}$$

To illustrate the application of Equation 3, consider the data for t, x_t, and y_t given in Table 20.1, where for simplicity each period is of equal length (i.e., $T_t = 1$ for all t). By taking logarithms as indicated in Equation 3 and using the values in Table 20.1, the dependent and independent variable terms may be written in matrix format as

$$Y = \begin{bmatrix} 2.48 \\ 2.71 \\ 3.00 \\ 3.04 \\ 3.22 \\ 3.26 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 1 & 2.30 \\ 1 & 3.09 \\ 1 & 3.66 \\ 1 & 4.08 \\ 1 & 4.38 \\ 1 & 4.65 \end{bmatrix},$$

respectively. Using least squares estimation, the estimates of the model parameters, c, and r in Equation 3 may be determined using Equation 4 in which X' is the transpose of matrix X

TABLE 20.1
Fitting a Learning Curve Using Linear Regression: A Numerical Example

Time, t	0	1	2	3	4	5	6
Output, y_t	10	12	17	20	21	25	26
Cumulative output, x_{t-1}	—	10	22	39	59	80	105
In (x_{t-1})		2.30	3.09	3.66	4.08	4.38	4.65
In (y_t/T_t)		2.48	2.71	3.00	3.04	3.22	3.26

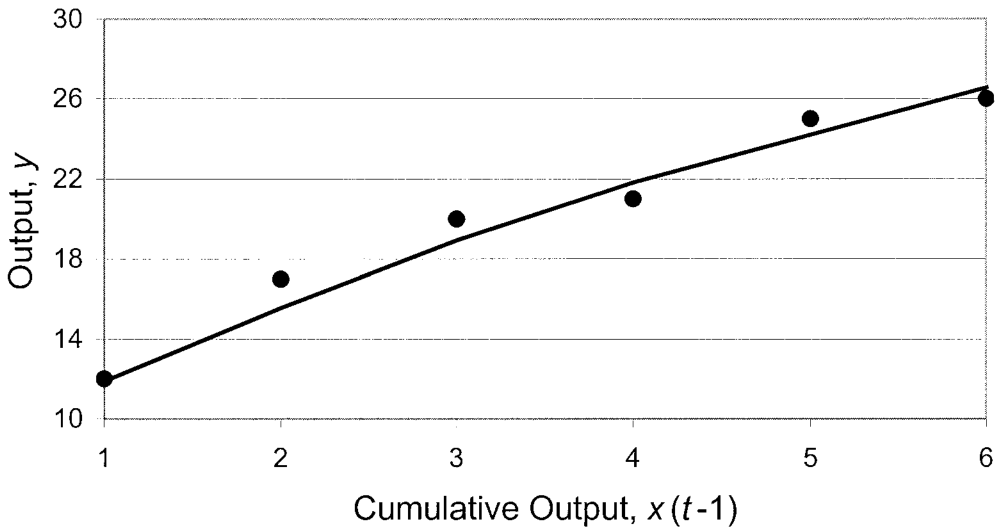

FIG. 20.1. Fitted linear model for Equation 3.

(see Neter, Kutner, Nachtsheim, & Wasserman, 1996 for a detailed discussion of the derivation and use of Equation 4).

$$\begin{pmatrix} c \\ r \end{pmatrix} = (X'X)^{-1}(X'Y) \tag{4}$$

The least squares parameter estimates for the data given in Table 20.1 are then $\binom{c}{r} = \binom{1.688}{0.342}$ with the corresponding plot of the fitted line and original data as shown in Fig. 20.1.

Equation 3 may be extended as shown in Equation 5, with an additional parameter, α, to represent potential diminishing returns on labor. Equation 3 is a special case of Equation 5 in which $\alpha = 1$.

$$Ln\,(y_t) = c + \alpha Ln(T_t) + r Ln(x_{t-1}) \tag{5}$$

The form in Equation 5 is suitable for estimating and predicting organizational learning characteristics given for each time period t, the productivity during t, the cumulative organizational experience from all previous periods, x_{t-1}, and the present time T_t. Additional information may be mined from an organization's production history to answer a broader range of questions such as whether the number of hours in a work shift affects the diminishing returns on increased labor hours. Adding an additional term, $\beta Ln(n_t)$, to the right hand side of Equation 5 yields such a model, where n_t represents the number of shifts per week, and h_t, the number of hours per shift with organizational parameters αgnd β, respectively. Note that by definition $T_t = h_t n_t$, making Equation 5 a special case of Equation 6, where $\alpha = \beta$.

$$Ln\,(y_t) = c + \alpha Ln(h_t) + \beta Ln(n_t) + r Ln(x_{t-1}) \tag{6}$$

The implicit assumption in forming the models shown in Equations 5 and 6 is that knowledge acquired through learning by doing does not depreciate (i.e., there is no forgetting effect). Models of organization knowledge depreciation vary widely. However, in the context of a linear model this may be represented by replacing the cumulative experience term x with a cumulative knowledge term that accounts for the depreciation of knowledge over time with depreciation parameter λ, as given in Equation 7.

$$Ln(y_t) = c + \alpha Ln(h_t) + \beta Ln(n_t) + r Ln\left(\sum_{s=1}^{t-1} \lambda^{t-s-1} x_s\right) \tag{7}$$

Depending on the information content in the data and the decision support questions being raised, a greater or lesser number of terms may be included in such models and fitted in an analogous manner to that shown for Equation 3. In general it will be difficult to calibrate the choice of model a priori, because the models have not yet been fitted to the data. However, a common approach to addressing this issue is to fit the data to each of the candidate models to help gauge both the appropriateness of the models and information content in the data.

A complete review of the organizational learning literature is broader than space permits. However, the interested reader can find recent comprehensive overviews of the organizational learning literature by Argote (1999), Cohen and Sproull (1996), and Dar-El (2000). Argote (1999) discusses data driven models such as that illustrated in example 1. Another segment of the literature is conceptual in nature such as the compilation by Cohen and Sproull (1996). A third related segment of the literature is prescriptive in nature and is based somewhat on previous results, many of which were data driven. Additional prescriptive outlines on improving organizational performance can be found in Druckman, Singer, and VanCott (1997) and Druckman and Bjork (1991, 1994).

INDIVIDUAL LEARNING

There are gaps in the understanding of human learning that go beyond the interaction of groups within organizations and extend to the individual level. An important limitation to research and decision-making advancements in this area has long been the difficulty in going beyond the laboratory to obtaining access to adequate field data and the subsequent knowledge embedded therein. Recently data warehousing efforts and information systems technologies have helped by providing sources of human performance data from the field. To analyze these data, effective tools are required to allow us to measure and preferably visualize a multiplicity of human behaviors.

At the individual level, human performance and learning is affected by many factors including individual ability; individual variability; financial incentives; organizational norms and constraints; task complexity; training; the nature of the social environment; technology; and other factors. Learning and improvement also occur in teams, work groups, departments, and other collections of people. In these groups interactions, both obvious and subtle, occur among the technologies, workers, and management systems that often lead to measurable performance improvement (Argote, 1993).

Knowledge of individual learning characteristics and distributions of these characteristics may be related to a number of factors at the individual level, such as workforce training, education, and the local industrial base. Building an understanding of how these factors affect learning rates may be central to improving them. As with organizational level learning, the literature on individual learning is extensive. Belkaoui (1986), Dar-El (2000), and Yelle (1979), provide comprehensive overviews of the numerous models and functional forms for measuring individual human learning behavior. Comparitive studies of many of these and other models are given by Badiru (1992) and Nembhard and Uzumeri (2000a). Several of these models also are shown in a later section.

Data on Individual Learning

Data on individual learning generally takes the form of an individual-related spacio-temporal database. At a minimum an individual record contains information such as an individual (human) identifier, a time stamp, and a location. Additionally, information on quality, complexity, teamwork, and other characteristics of the work may be beneficial but are not always available. The unit of analysis may vary, but it may be convenient to use a unit of production (batch), a part, or in the case of a service related task, a customer. The spatial dimension tends not to constrain the ability to mine such data.

Methods

The mining of human productivity data often has been based on individual curve estimation, or regression analyses. However, there may be opportunities to combine some of these with other techniques in the future, including classification, clustering, and online analytical processing (OLAP; see, e.g., Hasan & Hyland; 2001; Li & Wang, 1996). Regression analyses by Argote (1993), Argote Beckman and Epple (1990), and Epple et al. (1991) allow for the examination of learning characteristics among various groups. Curve estimation models by Nembhard (2000), Shafer, Nembhard, and Uzumeri (2001), and Uzumeri and Nembhard (1998) have helped develop knowledge about individual performance in industrial settings.

Models of organizational learning are effective and valuable when the type of activity and the granularity of the data warrant this view. However, greater detail may allow for the use of individual based models of organizational learning (e.g., Nembhard & Uzumeri, 2000a). The distinction between an individual based model of organizational learning and a model of individual learning is somewhat one of name and desired purpose. For example, Fig. 20.2a

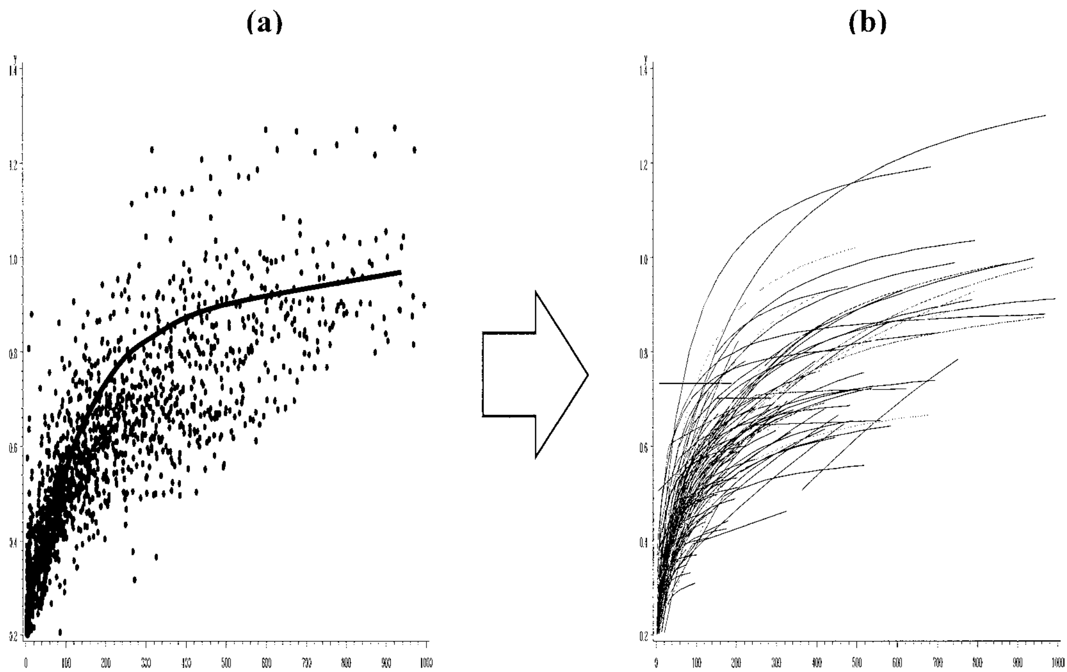

FIG. 20.2. Example of organizational and individual based models.

illustrates the organizational learning curve model given in Equation 3 fitted to organizational data. This view can adequately represent an organization's progress over time. However, in Fig. 20.2b, the same data may be fitted to a set of individual learning models to achieve a similar goal. It is clearly necessary for the data to contain identifiers of individual workers, teams, worker classifications, or other categories for this alternate view to be used. An individual based model then allows for further examination of individual characteristics, capabilities, variations, and potential for improvement. That is, the parameterizations of the individual characteristics can be related to the tasks, technologies, individual experience, age, gender, and education, and so forth. This type of secondary analysis has provided useful insights into human behavioral characteristics (see, e.g., Lance et al., 1998; Nembhard, 2000).

By taking an individual based perspective to organizational learning and forgetting behavior, a mathematical function may be used to describe the observed performance history in much the same manner as curves may be fitted to an organization's history. Standard curve-fitting methods can be applied to produce curves that capture the essential shape of the underlying "signal," as illustrated in Fig. 20.1. Some useful examples of nonlinear curve fitting methods may be found in Dennis and Schnabel (1983). Many of these and similar methods are available in off-the-shelf software such as SAS, SPSS, Minitab, and so forth. The fitted curves provide a filtered summary of the process that is more compact and offers several important benefits including the following.

1. *Reduced data dimensionality*. By fitting the data to an appropriate continuous function that describes the underlying signal in the data, the volume of information required to describe the process is reduced. For example, Fig. 20.3 illustrates 18 discrete data points representing

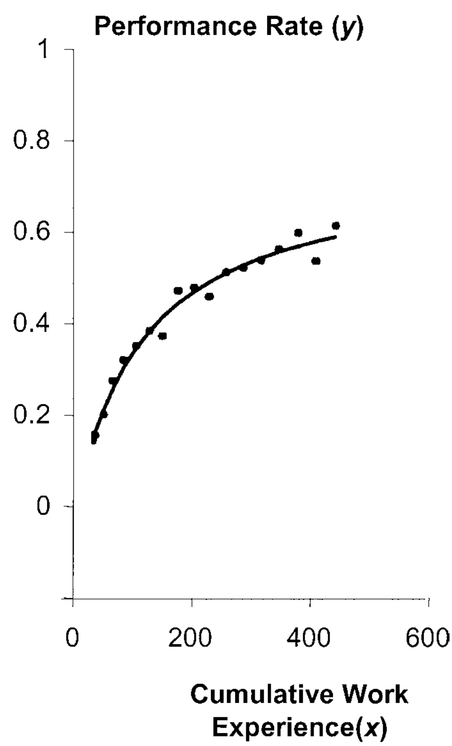

FIG. 20.3. Fitted individual learning curve (using Equation 11).

productivity rates, each with a pair of (x, y) coordinates that are described by a continuous function of only three parameters. Reduced dimensionality may facilitate visualization for decision makers while maintaining the embedded knowledge that aids comprehension and insight.

2. *Noise reduction.* By fitting a continuous function, noise contained in the discrete data is inherently reduced, assuming that the model chosen is appropriate for the human characteristic being modeled. When relevant, the distribution of the noise can be measured and recorded.

3. *Implicit mathematical analysis*. Because the signal pattern is described by the variables in the best-fit equation, an analyst can use the fitted equation to directly perform a number of otherwise complex calculations. For example, the derivative of the fitted function provides an estimate of the signal slope at a given point, which represents the instantaneous rate of improvement. Such information might then be aggregated across individuals and related to specific workplace interventions. For example, we may be able to address questions such as how a change in procedure affects the learning rates across a population of workers.

A general approach for mining human performance data for individual learning and forgetting characteristics and skill measurement is described as follows. Individual productive performance, y_i, can be modeled based on a nonlinear model, f_i, and model parameters, $\underline{\theta}_i$, that best fit the data set for an individual unit i, such that the fitted curve is given by,

$$y_i = f_i(\underline{x}_i, \underline{\theta}_i) + \varepsilon_i, \tag{8}$$

where $\underline{x}_i$ represents the data for curve i, with error term ε_i. That is, in fitting a curve for each individual i, we first select the most preferred function, f_i, for an individual, and then estimate the parameters $\underline{\theta}_i$ from the data using an appropriate nonlinear estimation technique. The underscores on x_i and θ_i indicate vectors the length of which depends on the data and model used, respectively.

When considering multiple individuals, there are two predominant methodological approaches. The first is to select both the best model and the best model parameters for each individual based on a model selection approach such as maximum likelihood estimation (MLE) or least squares estimation (LSE) (see, e.g., Gujarati, 1978). This tends to minimize model selection error because a variety of functional forms, f_i, may be considered, with the best candidates chosen for each individual. However, if comparisons among a large number of individual units, i, is desirable, this approach may be somewhat cumbersome because different models may be most preferred for the various individuals. This realization motivates the second approach, which is to select a single consistent function, g, to be used for all of the individuals in the data being considered. That is, for each individual i fit the model given by

$$y_i = g(\underline{x}_i, \underline{\theta}_i) + \varepsilon_i. \tag{9}$$

By requiring a single model, g, the consistency and relative ability to compare individual units that a single model affords introduces an additional model selection error, δ_i, as given in Equation 10.

$$\delta_i = [f_i(\underline{x}_i, \underline{\theta}_i) - g(\underline{x}_i, \underline{\theta}_i)]. \tag{10}$$

The choice of model, g, can be based on a model scoring approach that maps model-data combinations to numbers the numerical ordering of which corresponds to a preference ordering over the space of models, given the data (Glymour, Madigan, Pregibon, & Smyth, 1996).

Popular rules include the Akaike information criterion (Akaike, 1974), minimum description length (Rissanen, 1978), and Bayes information criterion (Raftery, 1995). A specific scoring system for selecting an individual based learning model is described by Nembhard and Uzumeri (2000a), in which the criteria include model efficiency, stability, and parsimony. Criteria may additionally include other considerations including estimation balance, the interpretability of resultant parameters, and the ability to estimate the model parameters. Several of these potential criteria are described in turn later.

Efficiency. It is important that a model be flexible with respect to the potential shapes of the responses. That is, it should describe any of the regularly observed patterns within the population of individuals. In the case of measuring learning and forgetting, for example, it is beneficial to have the ability to fit both positive and negative improvements and any prior learning based on previous task experience. An efficient model is flexible enough to capture the various shapes of learning among the individuals in a population. For the efficiency measure the average of the mean squared error (MSE) statistics, given by $\hat{\mu}_{MSE}$, across individuals is useful wherein lower $\hat{\mu}_{MSE}$ values indicate a better average fit among all individuals.

Stability. In addition to a good average fit, it is important that every curve in the population of curves fits the model reasonably well. An efficient model may fit a given curve extremely well but fit other individual curves relatively poorly. A stable model is not overly dependent on specific individuals to obtain good model fits (Borowiak, 1989). Stability allows a model to be robust with respect to variations between individuals within the population. Stability may be measured using the variability in the goodness-of-fit, given by $\hat{\sigma}_{MSE}$, for the population of individual curves. Lower $\hat{\sigma}_{MSE}$ values indicate that the model fits individuals in the population more consistently. For instance, given a relatively efficient model, high values of $\hat{\sigma}_{MSE}$ would indicate low stability and, hence, the presence of poor model fits for some individual curves.

Balance. A third scoring criterion is that of estimation balance, in which a balanced model tends to have an equal probability of providing overestimates and underestimates, an overestimate relates to the model *predicting* a higher output than the actual output, on average, and underestimates, a lower output than actual. This may be particularly important if the model is to be used for forecasting, in which imbalance creates a general bias toward overestimation or underestimation. It would generally be desirable to have little or no bias introduced as a result of the model choice. The skewness of the goodness-of-fit may be used as the measure for balance, or alternately, a raw percentage score as used, for example, by Nembhard and Osopthsilp (2001).

Parsimony. In general, it is desirable to limit the number of degrees of freedom used by a model, particularly when the data corresponding to the least data-rich individuals is relatively sparse. This parsimony is additionally useful if a model can be limited to two or three parameters, because they are more readily summarized in two-dimensional or pseudo-three-dimensional graphical representations. We might expect, a priori, that a four-parameter model would tend to fit an individual's data somewhat better than a model with fewer parameters, simply in terms of overall flexibility. However, there is a clear tradeoff between reducing the dimension of the parameter space at the expense of increasing the residual error. The number of model parameters is useful as a measure of such parsimony.

Estimableness. The relative ease or difficulty in fitting a model to each individual's performance history largely depends on the curve fitting algorithm, the starting conditions, step sizes, and termination criteria. In general, it may not be feasible for a single set of starting conditions to result in model convergence for all of the individuals in a data set. For large data sets and/or automated analyses the estimableness of model parameters given a standardized estimation procedure is an important consideration. A potential measure of estimableness is the percentage of individuals for which model parameter convergence can be obtained based on a consistent method and starting conditions. For example, other things being equal one would prefer parameter convergence in 99% of individuals for one model versus only 95% for another model.

Interpretability. Another potentially valuable model selection criterion, interpretability, measures the degree to which a model's parameters have clear contextual interpretations. It is preferable for each parameter to have clear and useful meanings for decision makers. This is somewhat more likely when the models were developed from theoretical bases as opposed to empirically, because theoreticians often start with a set of rational and interpretable parameters and then build models around them based on current knowledge in the discipline. For example, this is true for the model used later, in example 2 (Equation 11).

The literature reveals numerous functional forms, g, for measuring human learning and forgetting. Psychologists have studied and modeled individual learning processes (Mazur & Hastie, 1978), whereas engineers have modeled learning as it relates to manufacturing costs (Yelle, 1979), process times (Adler & Nanda, 1974a, 1974b; Axsater & Elmaghraby, 1981; Pratsini, Camm, & Raturi, 1993; Smunt, 1987; Sule, 1981), setup time learning (Karwan, Mazzola, & Morey, 1988; Pratsini, Camm, & Raturi, 1994), and line balancing (Dar-El & Rubinovitz, 1991). Although the majority of work has centered on the log-linear model (e.g., Badiru, 1992; Buck & Cheng, 1993; Glover, 1966; Hancock, 1967), there have been continuing efforts to identify alternative formulations. Some alternative forms were found to represent observed behavior better than the log-linear model in their respective industrial settings (e.g., Argote et al., 1990; Asher, 1956; Carlson & Rowe, 1976; Carr, 1946; Ebbinghaus, 1885; Elmaghraby, 1990; Globerson, Levin, & Schtub, 1989; Knecht, 1974; Jaber & Bonney, 1996; Levy, 1965; Nembhard & Uzumeri, 2000a; Pegels, 1969; Ramos & Chen, 1994).

Example 2. This section summarizes a numerical example of applying the above approach using real data from 3,874 individuals. The example is adapted from Nembhard and Uzumeri (2000a). Mazur and Hastie (1978) discuss a model of individual learning developed from psychological principles and later tested using experimental data (Mazur & Hastie, 1978) and field data (Uzumeri & Nembhard, 1998). The model is intended to describe learning based on three fitted parameters, k, p, and r, where p represents the prior expertise attributable to a task based on a fit of the model to the data and may be viewed as an estimate of a workers' expertise acquired from past and similar experience. It is in effect shifting the learning curve backward in cumulative work to estimate prior expertise. The fitted parameter k estimates the asymptotic *steady-state* productivity rate per unit time relative to the work standard. This is the rate that can be expected once all learning has been completed. The fitted parameter r is the cumulative production and prior expertise reach $k/2$, starting from the production rate corresponding to zero cumulative work and prior expertise. Thus, r represents the rate of learning *relative* to the individual's steady state productivity rate, k. Note that this definition holds for both positive and negative learning, and that smaller values of r correspond to a more rapid approach to steady state. The response variable, y_x, is a measure of the productivity rate corresponding to x units of cumulative work (i.e., the number of units completed thus far).

$$y_x = k\left(\frac{x+p}{x+p+r}\right) + \varepsilon_x \tag{11}$$

TABLE 20.2
Performance of Learning Curve Models

Model Name	*Form*	*Efficiency,* $\hat{\mu}_{MSE}$ *(change in error)*	*Stability,* $\hat{\sigma}_{MSE}$ *(change in error)*	*Parsimony, Number of Parameters*
Hyperbolic-3	$y = k[(x + p)/(x + p + r)]$	.00577 (0%)	.01183 (0%)	3
Exponential-3	$y = k(1 - e^{(x+p)/r})$	.00598 (+3.6%)	.01382 (+16.8%)[a]	3
DeJong	$y = C_1[M + (1 - M)x^{-b}]$	.00665 (+15.3%)	.01331 (+12.5%)	3
Stanford-B	$y = C_1(x + B)^b$	.00671 (+16.3%)	.01340 (+13.3%)	3
Log-linear	$y = C_1 x^b$	.00711 (+23.2%)	.01434 (+21.2%)	2
S-curve	$y = C_1[M + (1 - M)(x + B)^b]$	.00736 (+27.6%)	.01506 (+27.3%)	4
Hyperbolic-2	$y = k[x/(x + r)]$	.00873 (+51.3%)	.01410 (+19.2%)	2
Pegels	$y = A\ a^{x-1} + b$	.00984 (+70.5%)	.01327 (+12.2%)	3
Exponential-2	$y = k(1 - e^{-x/r})$	.01122 (+94.5%)	.01563 (+32.1%)[a]	2
Levy	$y = [1/\beta - (1/\beta - x^b/C_1)k^{-kx}]^{-1}$	.08096 (+1300%)	.12561 (+962%)[a]	4
Knecht	$y = C_1\ x^b e^{cx}$	.09817 (+1600%)	.16354 (+1280%)	3

Note: Table is adapted from Nembhard and Uzumeri (2000a).
[a]Denotes: convergence criterion not met for some of the individual curves.

Table 20.2 illustrates the use of several model selection criteria (efficiency, stability, parsimony, estimableness) to select a single learning model for multiple individuals. Noting that lower values are preferred for the first three criteria, the hyperbolic-3 model shown in Fig. 20.4 may be preferred, because it performed best on the first two criteria. However, in general such criteria may produce a set of nondominated alternatives, making the eventual choice among models somewhat less clear.

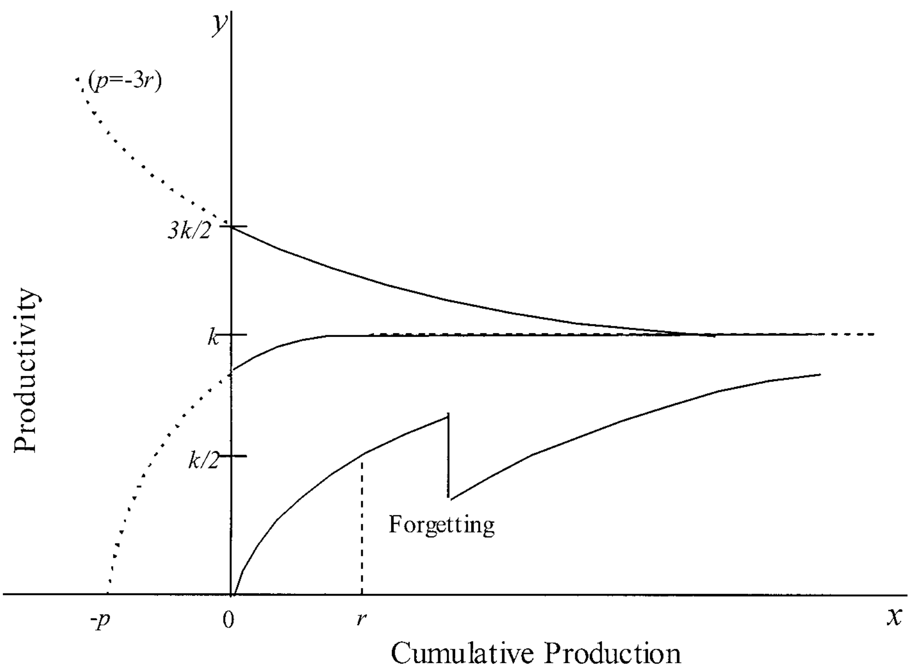

FIG. 20.4. Hyperbolic-3 learning model (in Equation 11).

Individual Forgetting

Research addressing the forgetting or the retention of learned skills suggests that the longer a person studies, the longer the retention (Ebbinghaus, 1885). The forgetting process has been shown to be describable by the traditional log-linear performance function (Anderson, 1985). Based on individuals' ability to recall nonsense syllables, Farr (1987) found that as the meaningfulness of the task increased, retention also increased. The retention of learning at an organizational level also has been found to decline rapidly (Argote et al., 1990). Several research efforts attempted to explain the impact that forgetting has on production scheduling (Fisk & Ballou, 1982; Khoshnevis & Wolfe, 1983; Smunt, 1987; Sule, 1983). Globerson et al. (1989) developed a model of learning and forgetting for a data correction task. The model indicates that forgetting behaves much like the mirror image of the learning process. Bailey (1989) found that forgetting of a procedural task was a linear function of the product of the amount learned and the log of the elapsed time. Nembhard and Osothsilp (2001) presented a survey and comparative study of many of the forgetting models in the literature employing an approach similar to that illustrated in example 2 for the learning process. The study compares forgetting models from Carlson and Rowe (1976), Elmaghraby (1990), Globerson and Levin (1987), Globerson, Levin, and Ellis (1998), Globerson et al. (1989), Jaber and Bonney (1996), and Nembhard and Uzumeri (2000b).

DISTRIBUTIONS AND PATTERNS OF INDIVIDUAL PERFORMANCE

Although modeling individual human performance has many potential direct uses for decision makers and researchers, additional insights may be obtained by studying the *distributions* of improvement behavior across populations of individuals. For example, Fig. 20.5 depicts parameter estimates for k, p, and r, from Equation 11 in three orthogonal views (a): $r \times k$; (b): $r \times p$; (c): $k \times p$. These distributions in general show that workers tend to fall along a continuum. At one extreme, workers learn quickly and then plateau at a relatively low level of performance (bottom left of Fig. 20.5a). At the other extreme, some workers improve slowly and steadily, but climb to high ultimate performance levels (upper right of Fig. 20.5a). Regardless of the mechanisms behind such behavior, these *maps* of individual performance, such as that in Fig. 20.5, can help an organization to build a quantitative description of human performance within its workforce. However, applications extend beyond managerial curiosity. Organizational learning is a group characteristic, and many managerial interventions (e.g., education, training, and technology) are designed to change both individual and group performance. Managers may be more concerned about the mix of learning curve shapes in the workforce than in changing the curves of specific individuals. To deal with individual curves, one would abandon the power of statistics and understand the many subtle forces acting on that specific individual. That is, organizational improvements are more likely to be due to aggregate changes in the mean or the shape of the learning distributions than from changes in a few specific individuals.

Focusing on the group level distributions of learning curve shapes may provide clues to systemic influences that would otherwise be difficult to discern. One straightforward approach for forming such a distribution is to consider the parameter estimates themselves as drawn for a multivariate distribution. Thus, one would estimate multivariate distribution parameters given the population of individual parameter estimates. Figure 20.5 illustrates such a multivariate distribution of three parameters. Using these parameter estimates, one may fit a given

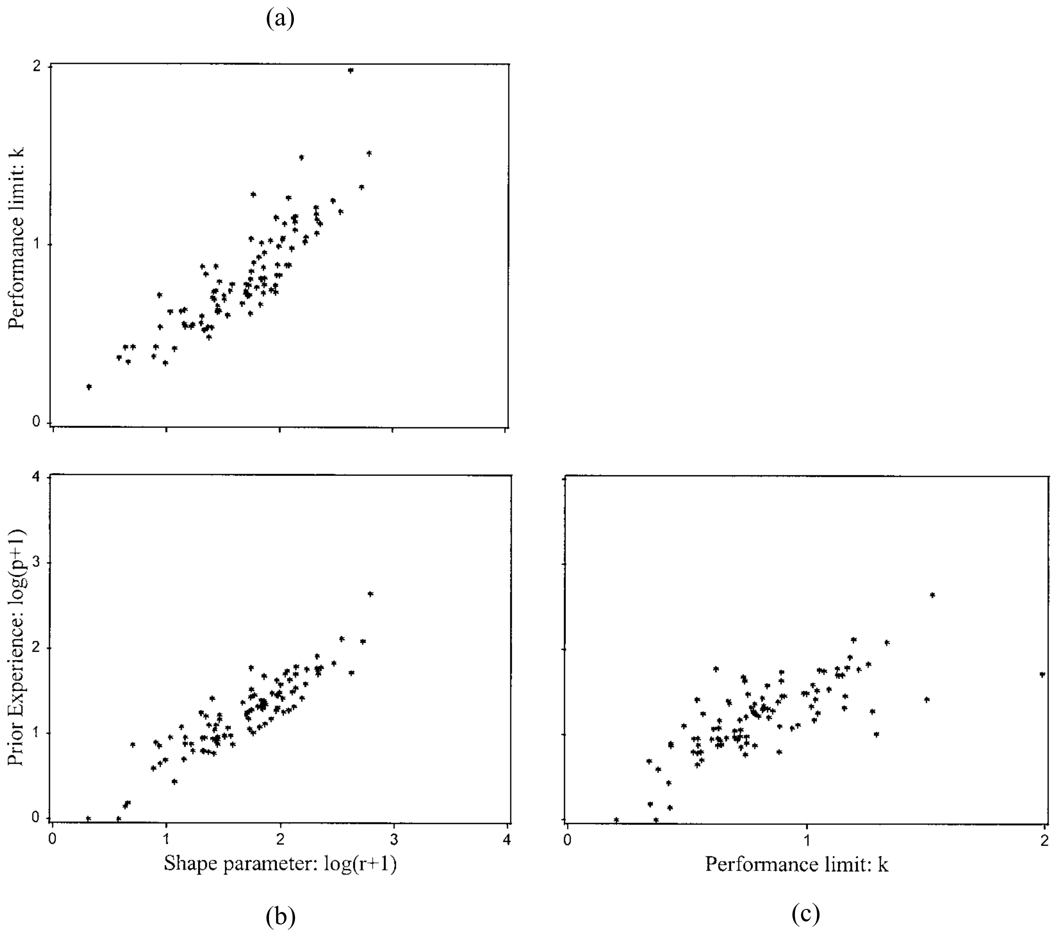

FIG. 20.5. Orthogonal views of the distributions of parameters k, p, r.

distribution such as a multivariate normal (MVN) to further simplify the representation of the original data. For example, the original data for Fig. 20.5 consists of hundreds of observations for each of 110 individuals. The MVN distribution consists of a 3×1 vector of means, and a 3×3 variance–covariance matrix.

The regularity of the learning map in Fig. 20.5 suggests that although parameter estimates are highly variable across individuals, they may be relatively predictable in the aggregate. For instance, we might view a newly hired individual as a random trial selected from the workplace distribution. Managers and researchers may employ statistical modeling and analysis tools to describe and apply the distributions involved to simulations, worker-task assignment, scheduling, line balancing, and other disciplines.

Empirical distributions may be particularly useful when rapid organizational change is taking place. For example, a company opening a plant in a previously unindustrialized area could estimate maps based on early performance data at the new plant or on educational demographics in the region. By comparing these maps to those of mature plants, one might identify systematic differences in the workforces before they are visible directly, thereby providing better production and cost estimates. Maps may also help management to better understand the nature of heterogeneity in their workforces, which has potential for creating new opportunities for taking patterns of worker variability into account in making key decisions. As an example, consider a company that is introducing a new product line. Should the company make the product in a new plant with new workers? Should it make the product in an old plant with experienced workers and shift other products to the new plant? Or should it simply expand its existing plant and use a mix of experienced and inexperienced workers?

Finally, it is important to note that the fitted curves represent a *descriptive* model of learning behavior that does not need to assume a specific causal mechanism. When used in this manner, descriptive curves can be fit to any variable. Instead of production rates, maps could be drawn for reductions in defect rates, for changes in cycle times, or for improvements in worker safety.

To this point, the discussion has assumed modeling of all available data. It should be noted that sampling individual performance records is also a viable approach, particularly when extremely large data sets are available and the data miner intends to estimate a distribution. For sampling, a variety of statistical sampling methods are available. Descriptions of these approaches may be found in Kowalewski and Tye (1990).

At the opposite end of the spectrum there may be circumstances in which data are relatively sparse and aggregation is necessary to obtain useful quantities of information. Although there is no standard established for aggregating human performance data, we may consider various attributes of the individual and the task conditions as candidate variables on which to aggregate. For example, one might aggregate all subjects performing a specific task, or a category of tasks. Alternatively, one might aggregate by years of experience, human skill level, or seniority level, among other variables. The choice of where to aggregate may ultimately be best driven by the intended purpose of the data mining efforts.

OTHER AREAS

The area of human–computer interaction has elements in common with human performance in that it deals with human characteristics and behavior given specific inputs or requirements. Researchers have made use of neural networks, pattern recognition, and visualization approaches for understanding how humans respond and perform computer related tasks (see, e.g., Beale & Finlay, 1992; Brecht & Jones, 1988; Rasmussen, 1983). A large portion of past work relied on closely controlled experiments to expand scientific understanding of underlying human mechanisms and behaviors. However, there are increasing opportunities for mining larger data sets in this area in much the same manner as that for human productivity. For example, given the wealth of data generated daily from Web based activities, information helpful to the understanding of human–computer interaction has significant data mining potential.

There is a relatively long history of research in the cognitive sciences that makes use of human performance data. Much of this work has been in the area of experimental psychology, which seeks to understand the underlying causality and mechanisms by which humans' performance is affected by various internal and external factors. For example, Ebbinghaus (1885) conducted numerous experiments to uncover the underlying mechanisms of human memory. Although these early efforts dealt with modest quantities of data by today's standards, some of these approaches are scaleable to the volumes of data generated by modern organizations.

The cognitive and physiological mechanisms by which individuals acquire knowledge have been carefully examined. Centered in behavioral psychology, artificial intelligence and industrial engineering, this research stream has generated sophisticated models of the processes by which humans acquire and retain knowledge (Anderson, 1982; Hancock, 1967; Mazur & Hastie, 1978).

PRIVACY ISSUES FOR HUMAN PERFORMANCE DATA

Maintaining individual privacy and organizational confidentiality is important and often necessary when dealing with human performance information. Research related studies are generally required to have human subject protocols in place, thereby protecting subjects' privacy. However, as organizations themselves begin to employ the approaches developed by researchers, concern arises from workers and labor organizations seeking to protect workers from potential or perceived abuses of the information.

Employee performance is evaluated in a variety of ways across industries and organizations to make decisions including promotions, layoffs, transfers, and raises. Given that it is not unusual for such evaluations to be made purely based on the judgment of supervisors and peers, one might expect that a more objective measure of performance based on automated data acquisition systems would be openly welcomed by workers. However, although workers in some organizations embrace these new tools, they are not universally accepted. One means of alleviating such fear is to openly communicate what is to be measured and how it will be used, including the protection of privacy when appropriate. For example, if data are collected for the purposes of pay-for-performance, employees are likely to be relatively receptive. Whereas, if their individual learning characteristics are determined, they may not want this information to become public knowledge. A "double blind" is often effective, in which the data miner obtains data with individual identifiers precoded to "anonymize" the individuals. Prior to any analysis, the miner recodes the individual identifier so that those with later access to any analysis will be unable to reassociate the information with specific individuals. Because human performance data mining often yields distributions, the organizations can focus on how the group behaves, which may be more acceptable, and potentially no less valuable, to those being studied, their representatives, and their organizations.

REFERENCES

Adler, P. S. (1990). Shared learning. *Management Science, 36,* 938–957.

Adler, G. L., & Nanda, R. (1974b). The effects of learning on optimal lot size determination: Multiple product case. *AIIE Transactions, 6,* 20–27.

Adler, G. L., & Nanda, R. (1974a). The effects of learning on optimal lot size determination: Single product case. *AIIE Transactions, 6,* 14–20.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369–406.

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.). New York: W. H. Freeman.

Argote, L. (1993). Group and organizational learning curves: Individual, system and environmental components. *British Journal of Social Psychology, 32,* 31–51.

Argote, L. (1999). *Organizational learning: Creating and transferring knowledge*. Boston: Kluwer Academic Publishers.

Argote, L., Beckman, S. L., & Epple, E. (1990). The persistence and transfer of learning in industrial settings. *Management Science, 36,* 140–154.

Argyris, C. (1989). Participatory action research and action science compared. *American Behavioral Scientist, 32,* 612–623.

Argyris, C. (1982). How learning and reasoning processes affect organizational change, In P. S. Goodman (Ed.), *Change in organizations: New perspectives on theory, research and practice*. San Francisco: Jossey-Bass.

Asher, H. (1956). *Cost and quantity relationships in the airframe industry* (Report No. R291). Santa Monica, CA: Rand Corp.

Axsater, S., & Elmaghraby, S. E. (1981). A note on EMQ under learning and forgetting. *AIIE Transactions, 13,* 86–90.

Badiru, A. B. (1992). Computational survey of univariate and multivariate learning curve models. *IEEE Transactions on Engineering Management, 39,* 176–188.

Bailey, C. D. (1989). Forgetting and the learning curve: A laboratory study. *Management Science, 35,* 340–352.

Beale, R., & Finlay, J. (1992). *Neural networks and pattern recognition in human-computer interaction*. New York: Ellis Horwood.

Belkaoui, A. (1989). *The learning curve*. London: Quorum Books.

Borowiak, D. S. (1989). *Model discrimination for nonlinear regression models*. New York: Marcel Dekker.

Brecht, B., & Jones, M. (1988). Student models: The genetic graph approach. *International Journal of Man Machine Studies, 28,* 483–504.

Buck, J. R., & Cheng, S. W. J. (1993). Instructions and feedback effects on speed and accuracy with different learning curve models. *IIE Transactions, 25,* 34–47.

Carlson, J. G., & Rowe, R. G. (1976). How much does forgetting cost? *Industrial Engineering, 8,* 40–47.

Carr, G. W. (1946). Peacetime cost estimating requires new learning curves. *Aviation, 45,* 76–77.

Cohen, M. D., & Sproul, L. S. (Eds.). (1996). *Organizational learning,* Thousand Oaks, CA: Sage.

Dar-El, E. M. (2000). *Human learning: From learning curves to learning organizations*. Boston: Kluwer Academic Publishers.

Dar-El, E. M., & Rubinovitz, J. (1991). Using learning theory in assembly lines for new products. *International Journal of Production Economics, 25,* 103–109.

Dennis, J. E. Jr., & Schnabel, R. B. (1983). *Numerical methods for unconstrained optimization and nonlinear equations*. Englewood Cliffs, NJ: Prentice Hall.

Druckman, D., Singer, J. E., & VanCott, H. (Eds.). (1997). *Enhancing organizational performance*. Washington, DC: National Academy Press.

Druckman, D., & Bjork, R. A. (Eds.). (1994). *Learning, remembering, believing*. Washington, DC: National Academy Press.

Druckman, D., & Bjork, R. A. (Eds.). (1991). In the mind's eye: Enhancing human performance. Washington, DC: National Academy Press.

Dutton, J. M., & Thomas, A. (1985). Relating technological change and learning by doing, research on technological innovation. *Management and Policy, 2,* 187–224.

Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology,* trans. H. A. Ruger & C. E. Bussenius. New York: Dover.

Elmaghraby, S. E. (1990). Economic manufacturing quantities under conditions of learning and forgetting (EMQ/LaF). *Production Planning and Control, 1,* 196–208.

Epple, D., Argote, L., & Devadas, R. (1991). Organizational learning curves: A method for investigation of intra-plant transfer of knowledge acquired through learning by doing. *Organization Science, 2,* 58–70.

Farr, M. J. (1987). *The long-term retention of knowledge and skills: A cognitive and instructional perspective*. New York: Springer-Verlag.

Fisk, J. C., & Ballou, D. P. (1982). Production lot sizing under a learning effect. *IIE Transactions, 14,* 257.

Globerson, S., & Levin, N. (1987). Incorporating forgetting into learning curves. *International Journal of Production Management, 7,* 80–94.

Globerson, S., Levin, N., & Ellis, S. (1998). Rate of forgetting for motor and cognitive tasks. *International Journal of Cognitive Ergonomics, 2,* 181–191.

Globerson, S., Levin, N., & Shtub, A. (1989). The impact of breaks on forgetting when performing a repetitive task. *IIE Transactions, 21,* 376–381.

Glover, J. H. (1966). Manufacturing progress functions: An alternative model and its comparison with existing functions. *International Journal of Production Research, 4,* 279–300.

Glymour, C., Madigan, D., Pregibon, D., & Smyth, P. (1996). Statistical themes and lessons for data mining. *Data Mining and Knowledge Discovery, 1,* 25–42.

Gujarati, D. (1978). *Basic econometrics*. New York: McGraw-Hill.

Hancock, W. M. (1967). The prediction of learning rates for manual operations. *Journal of Industrial Engineering, 18,* 42–47.

Hasan, H., & Hyland, P. N. (2001). Using OLAP and multidimensional data for decision making. *IT Professional, 3,* 44–50.

Jaber, M. Y., & Bonney, M. (1996). Production breaks and the learning curve: The forgetting phenomenon. *Applied Mathematical Modeling, 20,* 162–169.

Kantor, P. B., & Zangwill, W. I. (1991). Theoretical foundation for a learning rate budget. *Management Science, 37,* 315–330.

Karwan, K. R., Mazzola, J. B., & Morey, R. C. (1988). Production lot sizing under set-up and worker learning. *Naval Research Logistics, 35,* 159–179.

Knecht, G. R. (1974). Costing, technological growth and generalized learning curves. *Operations Research Quarterly, 25,* 487–491.

Knoshnevis, B., & Wolfe, P. M. (1983). An aggregate production planning model incorporating dynamic productivity. Part II: Solution methodology and analysis. *IIE Transactions, 15,* 283–291.

Kowalewski, M. J., & Tye, J. B. (1990). *Statistical sampling: Past, present, and future theoretical and practical*. Philadelphia: American Society for Testing and Materials.

Lance, C. E., Parisi, A. G., Bennett, W. R., Teachout, M. S., Harville, D. L., & Welles, M. L. (1998). Moderators of skill retention interval/performance decrement relationships in eight U.S. Air Force enlisted specialties. *Human Performance, 11,* 103–123.

Levy, F. K. (1965). Adaptation in the production process. *Management Science, 11,* 136–154.

Li, C., & Wang, X. S. (1996). A data model for supporting on-line analytical processing. *Proceedings of the Conference on Information and Knowledge Management* (pp. 81–88). New York: Association for Computing Machinery.

Mazur, J. E., & Hastie, R. (1978). Learning and accumulation: A reexamination of the learning curve. *Psychological Bulletin, 85,* 1256–1274.

Meredith, J., & Camm, J. (1989). Modeling synergy and learning under multiple advanced manufacturing strategies. *Decision Sciences, 20,* 258–271.

Nembhard, D. A. (2000). The effects of task complexity and experience on learning and forgetting: A field study. *Human Factors, 42,* 272–286.

Nembhard, D. A., & Uzumeri, M. V. (2000a). An individual-based description of learning within an organization. *IEEE Transactions on Engineering Management, 47,* 370–378.

Nembhard, D. A., & Uzumeri, M. V. (2000b). Experiential learning and forgetting for manual and cognitive tasks. *International Journal of Industrial Ergonomics, 25,* 315–326.

Nembhard, D. A., & Osothsilp, N. (2001). An empirical comparison of forgetting models. *IEEE Transactions on Engineering Management, 48,* 283–291.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Chicago: Irwin.

Pegels, C. C. (1969). On startup or learning curves: An expanded view. *AIIE Transactions, 1,* 216–222.

Pratsini, E., Camm, J. D., & Raturi, A. S. (1993). Effects of process learning on manufacturing schedules. *Computers and Operations Research, 20,* 15–24.

Pratsini, E., Camm, J. D., & Raturi, A. S. (1994). Capacitated lot sizing under set-up Learning. *European Journal of Operational Research, 72,* 545–557.

Raftery, A. E. (1995). Bayesian model selection in social research (with discussion). In P. V. Marsden (Ed.), *Sociological methodology* (pp. 111–196). Oxford, UK: Blackwells.

Ramos, M. A. G., & Chen, J. J. G. (1994). On the integration of learning and forgetting curves for the control of knowledge and skill acquisition for non-repetitive task training and retraining. *International Journal of Industrial Engineering, 1,* 233–242.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics, 13,* 257–266.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica, 14,* 465–471.

Shafer, S. M., Nembhard, D. A., & Uzumeri, M. V. (2001). The effects of worker learning, forgetting, and heterogeneity on assembly line productivity. *Management Science, 47,* 1639–1653.

Senge, P. M. (1992). Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future. *European Journal of Operational Research, 59,* 137–151.

Smunt, T. L. (1987). The impact of worker forgetting on production scheduling. *International Journal of Production Research, 25,* 689–701.

Sule, D. R. (1981). Effect of learning and forgetting on economic lot size scheduling. *International Journal of Production Research, 21,* 771–786.

Sule, D. R. (1983). Effect of learning and forgetting on economic lot size scheduling problems. *International Journal of Production Research, 21,* 771–786.

Uzumeri, M. V., & Nembhard, D. A. (1998). A population of learners: A new way to measure organizational learning. *Journal of Operations Management, 16,* 515–528.

Yelle, L. E. (1979). The learning curve: Historical review and comprehensive survey. *Decision Sciences, 10,* 302–328.