ON THE FORM OF LEARNING CURVES

JOSEPH C. STEVENS AND HARRIS B. SAVIN¹

HARVARD UNIVERSITY

In a wide variety of learning experiments, cumulated measures of performance turn out to be a power function of practice time or number trials, typically with an exponent between 1 and 2. The effects of various conditions on the course of learning can frequently be described in terms of the two constants of the function.

Skinner and his colleagues have made it clear that a cumulative plot of responses against time brings out regularities in behavior that are obscured when time is divided into "trials" and a measure is plotted that is descriptive of the behavior within a "trial." The cumulative curve has seldom been used, however, except for discrete responses such as lever pressing. This is surprising because the cumulative curve is particularly useful when the response is a continuous variable (as in rotary pursuit, for example) and division into trials is arbitrary.

We have replotted data from a variety of learning experiments and have found that the cumulative learning curves can be described very well by a power function:

$$P = kt^n \quad (1)$$

where P is one or another cumulated measure of performance and t is practice time or number of trials. For many purposes, it is useful to consider this equation in its logarithmic form:

$$\log P = n \log t + \log k \quad (2)$$

Thus, in log-log coordinates the power function plots as a straight line with slope n and intercept $\log k$. The size of n is invariant under any multiplicative transformation of the independent and dependent variables and is therefore unaffected by the experimenter's choice of units of measurement.

Evidence for the power function is pre-

sented in Fig. 1, which shows results from eight experiments on learning, plotted in log-log coordinates. The straight lines were fitted by eye, and the constants determined graphically.

The fits are so close that an analytic curve-fitting procedure (e.g., least squares) is unnecessary. These eight functions, together with many similar ones that we have plotted, demonstrate that the power function is a rather general finding. It describes the learning of both human and animal subjects and of the individual as well as the "average" subject. It is interesting that a function of the same form has recently been shown to govern the growth of sensory intensity (S. S. Stevens, 1957).

The only appreciable deviations of the data from the lines occur at the beginning of a few of the curves. These departures are of the kind that would be expected if the learner started with some degree of skill. The curvature at the beginning may indeed prove a useful measure of transfer of training or of retention measured by savings in relearning.

The exponent of the power function is a measure of the *curvature* of the cumulative learning curve, and it almost always has a value between 1 and 2, as it must if the traditional trial-by-trial (noncumulative) learning curve is negatively accelerated. An exponent smaller than 1 would imply a *deterioration* of performance. An exponent larger than 2 would imply a positively accelerated trial-by-trial curve. (Thus, curves traditionally characterized as positively or negatively accelerated might actually differ only in the size of the exponent.) The power function can, of course, be used to predict the form of trial-by-trial curves (in the same way that the law of falling

¹Research supported by Grant G10716 from the National Science Foundation. The authors are grateful for the help and encouragement of S. S. Stevens and R. J. Herrnstein.

<u>Learning Task</u>	<u>Equation</u>	<u>Abscissa</u>	<u>Ordinate</u>	<u>No. Subjects</u>	<u>References</u>
1. Nonsense Syllables	$P = 5.7t^{1.56}$	Trials (1-20)	No. correct anticipations	40	Tsao (1948)
2. Pursuit Rotor	$P = 0.064t^{1.46}$	Minutes (1-31)	Minutes	25	Doré & Hilgard (1937)
3. Ball Tossing	$P = 0.0356t^{1.25}$	No. throws (200-20,000)	No. in target area	1	Braden (1924)
4. U-Maze	$P = 0.027t^{1.19}$	Trials (25-250)	No. correct responses	4 (rats)	Anderson (1937)
5. Coding Letters	$P = 16t^{1.18}$	Minutes (1-20)	No. letters	97	Gentry (1940)
6. Writing Upside Down	$P = 25t^{1.18}$	Minutes (1-70)	No. letters	45	Kientzle (1946)
7. Operant Extinction	$P = 3.42t^{1.17}$	Minutes (3-57)	Decrement in No. responses	1 (rat)	Catania (unpublished data)
8. Typesetting	$P = 46.4t^{1.10}$	Hours (12.7-1227)	No. ems set	50	Kelly & Carr (1924)

Fig. 1. Cumulative learning curves in log-log coordinates. The table beneath the graph gives the equation of each curve, the nature of the variables measured, and the range of values plotted on the abscissa. In order to avoid crowding, a few points are omitted near the top of some of the functions. Otherwise, all available data are plotted.

bodies can be used to predict the behavior of a falling object within any arbitrarily chosen interval of time) by the formula:

$$P + \Delta P = k(t + \Delta t)^n \quad (3)$$

When the interval is allowed to become vanishingly small (*i.e.*, when we take the derivative of the power function), equation (3) assumes the simple form:

$$\frac{dP}{dt} = nkt^{(n-1)} \quad (4)$$

The derivative is of special interest. In pursuit-rotor learning, for example, the derivative may be thought of as the probability of being on the target. Thus, when the derivative reaches 1, performance has reached its maximum (*i.e.*, the subject is on the target all the time), and the power function can no longer hold. In practice, experiments are seldom continued until the limit is reached.

The effects of secondary variables on learning can often be described in terms of the two constants of the power function. Typically, the exponent n and therefore the rate of improvement increases with greater spacing of practice (Doré & Hilgard, 1937; Gentry, 1940; Kientzle, 1946; Kimble, 1949). An inverse relation between n and k is encountered frequently. In pursuit-rotor learning, for example, a subject who performs well determines a function with a relatively high value of k and a relatively low value of n .

The functions of Fig. 1, selected from several dozens like them, illustrate some of the kinds of learning for which the power function consistently provides a satisfactory description of acquisition. We are less sure that the power function applies to other classes of learning. In some cases, it is difficult to decide what variable to measure. An example is the learning of avoidance responses, which do not appear to follow a simple power function. But it has frequently been suggested that two things are learned in avoidance conditioning. The conditioned stimulus must first become "aversive," and only then does learning of the avoidance response commence. The cumulated number of responses may be a power function of trials *after* the conditioned stimulus has become aversive. We can often fit a power function to the data by making arbitrary assumptions about the number of trials before the conditioned stimulus becomes aversive.

But at this stage of the investigation, it is not clear what kind of function best describes avoidance learning.

Although we have plotted numerous curves for rote verbal memorization, all of which are well fitted by power functions (*e.g.*, Tsao, 1948), we are not sure exactly how to interpret them. Recent studies of verbal learning (Estes, 1960; Rock, 1957) suggest that each item is learned on a single trial. If this is true, it is difficult to interpret a plot of the total number of correct anticipations of several different nonsense syllables.

The choice of a suitable variable arises also in the study of a response like bar pressing, which may reach its asymptotic rate after only one rewarded occurrence. We do not know of any meaningful measure of the course of acquisition for this kind of learning. On the other hand, if reward is discontinued after many bar-press responses, the rate of responding decreases gradually. In this case the cumulated-response decrement, and not the cumulated responses, seems to be the interesting variable. If we take the rate just prior to extinction as a base line, the cumulated number of responses *not emitted* grows as a power function of time (Fig. 1).

Although the power function provides an accurate description of many kinds of learning, we do not expect it to fit all data perfectly, even after discounting "random" irregularities. In pursuit-rotor learning, for example, it is well known that performance improves sharply after a brief rest and deteriorates during a single extended trial. Thus, when practice and rest are alternated every few minutes, time on target cannot be exactly a power function of practice time. Although these and other secondary deviations may be important in their own right, they should not be permitted to obscure the basic simplicity of the acquisition curve.

REFERENCES

- Anderson, A. C. The effect of equalizing reward upon the breakdown of a discrimination habit, and its bearing upon reminiscence. *J. comp. Psychol.*, 1937, **23**, 421-437.
- Braden, S. R. An extensive experiment in motor learning and re-learning. *J. educ. Psychol.*, 1924, **15**, 313-315.
- Doré, L. R., and Hilgard, E. R. Spaced practice and the maturation hypothesis. *J. Psychol.*, 1937, **4**, 245-261.

- Estes, W. K. Learning theory and the new mental chemistry. *Psychol. Rev.*, 1960, **67**, 207-223.
- Gentry, J. R. Immediate effects of interpolated rest periods on learning performance. *Teach. Coll. Contr. Educ.*, 1940, No. 799.
- Kelly, C. M., and Carr, H. The curve of learning in typesetting. *J. exp. Psychol.*, 1924, **7**, 447-455.
- Kientzle, M. J. Properties of learning curves under varied distributions of practice. *J. exp. Psychol.*, 1946, **36**, 187-211.
- Kimble, G. A. Performance and reminiscence in motor learning as a function of the degree of distribution of practice. *J. exp. Psychol.*, 1949, **39**, 500-510.
- Robinson, E. S., and Brown, M. A. Effect of serial position upon memory. *Amer. J. Psychol.*, 1926, **37**, 538-552.
- Rock, I. The role of repetition in associative learning. *Amer. J. Psychol.*, 1957, **70**, 186-193.
- Stevens, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153-181.
- Tsao, J. C. Studies in spaced and massed learning: I. Time period and amount of learning. *Quart. J. exp. Psychol.*, 1948, **1**, 29-36.

Received April 24, 1961