Memory & Cognition
2001, 29 (7), 1061-1068

The power law as an emergent property

RICHARD B. ANDERSON
Bowling Green State University, Bowling Green, Ohio

Recent work has shown that the power function, a ubiquitous characteristic of learning, memory, and sensation, can emerge from the arithmetic averaging of exponential curves. In the present study, the forgetting process was simulated via computer to determine whether power curves can result from the averaging of other types of component curves. Each of several simulations contained 100 memory traces that were made to decay at different rates. The resulting component curves were then arithmetically averaged to produce an aggregate curve for each simulation. The simulations varied with respect to the forms of the component curves: exponential, range-limited linear, range-limited logarithmic, or power. The goodness of the aggregate curve's fit to a power function relative to other functions increased as the amount of intercomponent slope variability increased, irrespective of component-curve type. Thus, the power law's ubiquity may reflect the pervasiveness of slope variability across component functions. Moreover, power-curve emergence may constitute a methodological artifact, an explanatory construct, or both, depending on the locus of the effect.

The power function is generally accepted as an appropriate mathematical description of human performance in psychophysics (Stevens, 1971), in skill acquisition (J. R. Anderson, 1982; Logan, 1988; Neves & J. R. Anderson, 1981), and in retention (J. R. Anderson & Schooler, 1991; R. B. Anderson, Tweney, Rivardo, & Duncan, 1997; Rubin, 1982; Wixted & Ebbesen, 1991). Power curves have been observed so frequently, and in such varied contexts, that the term "power *law*" is now commonplace (but see Ru-bin & Wenzel, 1996, for some apparent violations of the power law). Figure 1 contains a representative curve from R. B. Anderson et al. (1997), illustrating the decline in recall performance as a power function of length of the retention interval. The function is represented mathematically as

$$Y = (A)X^{-B},\ -B < 0, \tag{1}$$

where Y is performance level (e.g., recall), X is time (e.g., retention interval), and A and B are parameters defining the precise shape of a given power curve. Parameter B determines the slope[1] of the curve; larger values of B yield steeper slopes. Parameter A is an intercept that specifies the value of Y when X equals 1.0. R. B. Anderson et al. fit the data in Figure 1 to four common functions: linear, exponential, logarithmic, and power. In accordance with the power law, the goodness-of-fit to the power function (using R^2 as the fit criterion) was greater than the fit to the other three functions. Similar findings have been reported by J. R. Anderson and Schooler (1991) and Wixted and Ebbesen (1991).

Repeated confirmation of the power law has prompted theoretical explanations for the phenomenon. Building on Newell and Rosenbloom's (1981) analysis of skill acquisition and Estes's (1956) discussion of curve averaging, R. B. Anderson and Tweney (1997) argued that the power law might characterize forgetting at the aggregate level only, and that the forgetting of single memory traces, or parts of memory traces, might proceed exponentially. The essence of the argument is captured by the following hypothetical example (more extensive analyses will be presented later). Imagine that the memory store contains three traces that decay according to the exponential equation

$$Y = (A)e^{-BX},\ -B < 0, \tag{2}$$

where Y is retrieval probability, X is time, A is the value of Y when $X = 0$, B is the slope (as in Equation 1), and e is the base of natural logs. Thus, for each trace, retrieval probability declines by a constant percentage, over each unit of time. Suppose further that each curve has a different slope, and that the curves combine to form an aggregate curve, equal to the arithmetic average of the three components. Figure 2a, illustrates the example. Though composed of exponential components, the aggregate curve itself is nonexponential; its percent of decline, or loss-percentage, decreases over time. Other functions with this characteristic include the power function and the logarithmic function. However, a regression analysis of the aggregate curve in Figure 2a indicates that the curve fits a power function (R^2 = .99) better than it does a logarithmic function (R^2 = .98), an exponential function (R^2 = .87), or a linear function (R^2 = .76). This kind of observation has led to the hypothesis that forgetting is fundamentally exponential, and that the empirical power function is an artifact of averaging across component curves. Indeed, reanalysis of published data

I thank William Batchelder, Robert Bjork, Bennet Murdock, In Jae Myung, David Rubin, Ryan Tweney, John Wixted, and Steven Sloman, for their helpful comments on the manuscript. Correspondence should be addressed to R. B. Anderson, Department of Psychology, Bowling Green State University, Bowling Green, OH 43403 (e-mail: randers@bgnet.bgsu.edu).

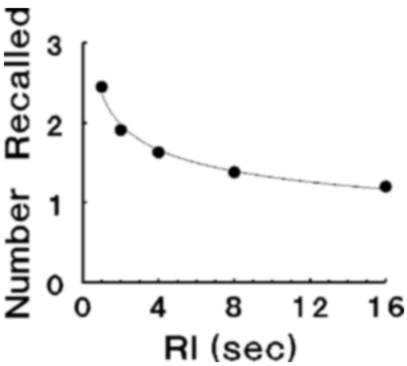

Figure 1. Mean number of items recalled as a function of retention interval (RI). Adapted from R. B. Anderson, Tweney, Rivardo, and Duncan (1997). The data are from the flat test-probability condition in Experiment 2 and are collapsed across five blocks of trials. Solid line shows best-fitting power function.

sets has shown that arithmetic averaging across trial position often exaggerates the fit of the aggregate curve to a power function, relative to an exponential function (R. B. Anderson & Tweney, 1997).

Though tenable, the exponential-component hypothesis raises the question of why forgetting should proceed exponentially at the trace level. Indeed, Newell and Rosenbloom (1981) recognized the same problem with regard to the hypothesis that skill acquisition might obey a power law at the aggregate level, but obey an exponential law at the component-skill level. The present paper offers a refined and expanded version of the exponential-component hypothesis. In this new account, component curves need not be exponential (as in R. B. Anderson & Tweney, 1997; Newell & Rosenbloom, 1981; Wickens, 1998) to produce a power curve at the aggregate level. Instead, the power law may be an emergent property that arises naturally from intercomponent slope variability, irrespective of the components' mathematical form.[2]

Consider the case of linear components. Of course, the arithmetic average of *purely* linear components will itself be linear. However, if one introduces the reasonable assumption that forgetting is range limited (i.e., that retrieval probability never falls below zero), then even linear components can produce an aggregate power function. Figure 2b shows three hypothetical range limited components. Component 1 is a proper linear function (within the time range of the example). However, Components 2 and 3 have sufficiently steep slopes so that performance drops to zero and remains there as time approaches its maximum. The figure also shows that the aggregate curve fits a power function better than it fits a linear, logarithmic, or exponential function (see Table 1 for the R^2 values). Figures 2c and 2d

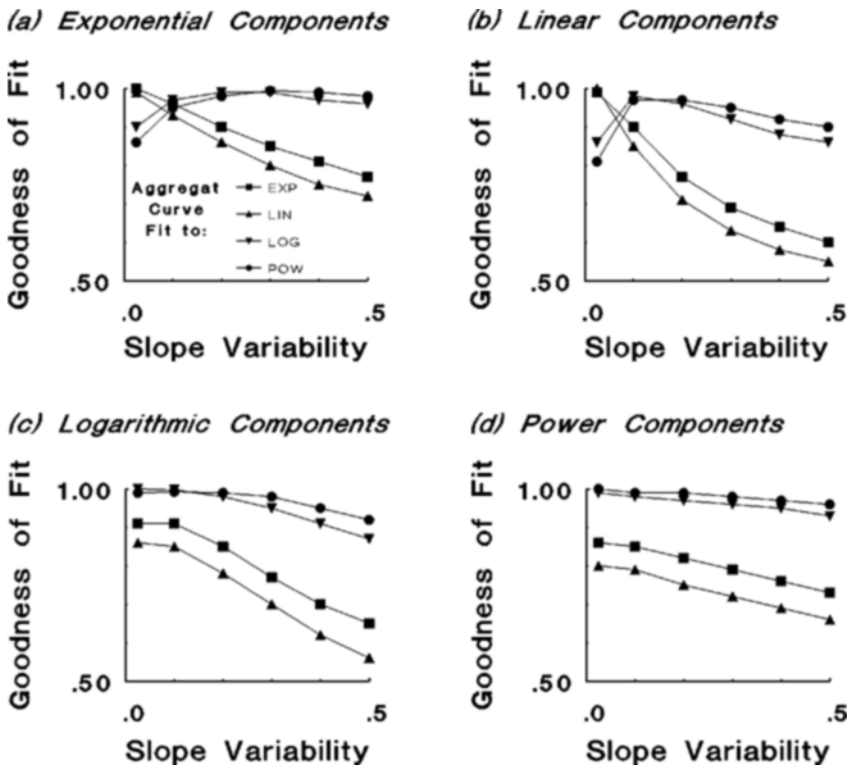

Figure 2. Aggregate forgetting curves derived by arithmetic averaging of three hypothetical component curves. Each panel shows a different type of component curves: exponential, range-limited linear, range-limited logarithmic, and power. Solid lines show best-fitting power functions.

Table 1
Goodness-of-Fit (R^2) of Aggregate Curves (in Figure 2) to Best-Fitting Exponential (Exp), Linear (Lin), Logarithmic (Log), and Power (Pow) Functions

	Mean R^2			
Component Type	Exp	Lin	Log	Pow
Exponential	.87	.76	.98	.99
Linear (range limited)	.74	.65	.92	.95
Logarithmic (range limited)	.74	.59	.90	.97
Power	.71	.60	.89	.96

Note—Functions were fit by entering appropriate log transforms into a linear regression algorithm.

show that an aggregate power curve may also result from the averaging of logarithmic components (range limited, as with the linear case) or power components.[3] Table 2 contains the mathematical representations of the four function types. Note that range limitation is unnecessary for exponential and power functions since both functions necessarily maintain a positive value.

Taken together, the examples demonstrate that intercomponent slope variability can yield aggregate power curves, irrespective of the components' forms (i.e., exponential, range-limited logarithmic, range-limited linear). Consequently, the ubiquity of the power law may reflect the widespread presence of slope variability.

To date, there has been no formal, mathematical proof of power-curve emergence. Laplace series expansions, for example, show that the power function *cannot* be formally expressed as the sum of exponential components (R. B. Anderson & Tweney, 1997; Newell & Rosenbloom, 1981). Moreover, in the case of range-limited linear and range-limited logarithmic components, formal proof via differential calculus seems inapplicable, since not every component is fully differentiable. Consequently, the present paper employs computational simulation, rather than formal mathematical proof, to demonstrate the phenomenon's generality.

METHOD

Each simulation contained 100 memory traces (components) that decayed exponentially, linearly, logarithmically, or in accordance with a power function (in the linear and logarithmic cases, the functions were range limited). All component curves were arbitrarily assigned a value of .9 for the A parameter (Table 2). The simulations did not include manipulation of the A parameter: It is already known that variability in A cannot be a general cause of emergent power curves since, in the case of exponential components, variability in A has no effect on the aggregate fit (R. B. Anderson & Tweney, 1997; Estes, 1956).

Values for B (i.e., the slopes) were generated randomly via the following procedure. First, B was set to a level that would yield a performance level of .5 after 20 units of simulated elapsed time (B was set to .029, .020, .133, and .195, for the exponential, linear, logarithmic, and power components, respectively). Next, in order to generate component curves with differing slopes, random deviates were added to the initial B. The deviates were sampled from a normal population with a mean of zero. The sampling procedure was repeated to produce 100 values of B, defining a forgetting curve for each of 100 traces. To insure that all curves would be descending (as in Figure 2) rather than ascending, B was prevented from taking on a negative value. If B became negative as a result of adding a random deviate, then B was resampled. (Note that the above constraint on B is distinct from the range limitation of the linear and logarithmic functions.) The 100 component curves were arithmetically averaged to produce an aggregate curve. In order to assess the effect of component slope variability on the form of the aggregate curve, the standard deviation of the random deviate distribution (described above) was manipulated across six levels: 0, 0.1, 0.2, 0.3, 0.4, and 0.5. The simulation was replicated 500 times at each level of slope variability.

Although the power law has been observed using various curve-fitting methods, the most frequent method involves appropriately transformed data into a linear regression algorithm R. B. Anderson & Tweney, 1997). Thus, since the linear regression method defines (to a great extent) the power-law phenomenon, that method will be used in the present analyses.

RESULTS AND DISCUSSION

Although other fit indices are available, the present simulations used R^2 as the goodness-of-fit measure. R^2 was computed by linearizing the aggregate curves and then submitting them to a simple linear regression. Linearization was accomplished by transforming the Y axis to $\log Y$ (to assess fit to an exponential function) or by transforming the X axis to $\log X$ (to assess logarithmic fit) by log transforming both axes (to assess power fit). Of course, no linearization (i.e., no transformation) was necessary to assess goodness-of-fit to a linear function.

The incidence of aggregate power curves was assessed by counting the number of times, n (of 500), in which the aggregate curve's fit to a particular function was superior to its fit to any of the three remaining functions. Table 3 shows that the number of aggregate power curves increased as amount of intercomponent slope variability increased, thereby establishing that slope variability at the component level can produce a power function at the aggregate level. In the case of power components, the aggregate curve was best fit by a power function, regardless of slope variability. Thus, although the presence of an aggregate power curve is consistent with the existence of power components, it is also quite consistent with the presence of other types of component curves. Table 4 shows similar results for simulations in which the A parameter is set to .5 instead of .9. Note, however, that Table 4 shows power curves sometimes combining to form aggregate logarithmic curves.

The aggregate power-curve phenomenon was also assessed by calculating the mean goodness-of-fit (defined

Table 2
Mathematical Equations Representing Exponential, Linear, Logarithmic, and Power Functions

Function Type	Equation
Exponential	$Y = (A)e^{-BX}, -B < 0$
Linear	$Y = A - BX, Y > 0, B > 0$
Logarithmic	$Y = A - B \log X, Y > 0, B > 0$
Power	$Y = (A)X^{-B}, -B < 0$

Table 3
Effect of Component Type and Intercomponent Slope Variability on the Number of Cases (*n*) in Which the Aggregate Curve Was Best Fit by an Exponential (Exp), Linear (Lin), Logarithmic (Log), or Power (Pow) Function

Component Type	Slope Variability	*n* Exp	Lin	Log	Pow
Exponential	.0	500	0	0	0
	.1	58	0	442	0
	.2	0	0	495	5
	.3	0	0	68	432
	.4	0	0	0	500
	.5	0	0	0	500
Linear (range limited)	.0	0	500	0	0
	.1	0	0	464	36
	.2	0	0	1	499
	.3	0	0	0	500
	.4	0	0	0	500
	.5	0	0	0	500
Logarithmic (range limited)	.0	0	0	500	0
	.1	0	0	500	0
	.2	0	0	0	500
	.3	0	0	0	500
	.4	0	0	0	500
	.5	0	0	0	500
Power	.0	0	0	0	500
	.1	0	0	0	500
	.2	0	0	0	500
	.3	0	0	0	500
	.4	0	0	0	500
	.5	0	0	0	500

Note—The components' *A* parameters are fixed at 0.9.

here and throughout as mean R^2) of the aggregate curve to the best-fitting exponential, linear, logarithmic, and power functions. Figure 3 shows the effect of the amount of intercomponent slope variability (i.e., the standard deviation of distribution of random deviates) on the mean R^2 averaged across 500 replications.

Overall, the power fit became increasingly superior as intercomponent variability increased, irrespective of component-curve type. Of course, in the case of power-curve components (Figure 3d), the aggregate curve was fit perfectly by a power function when slope variability was zero. Still, the relative goodness of the power fit increased with slope variability. In sum, the simulation results for R^2 were consistent with the results for *n* (i.e., the frequency of aggregate power curves) and with the examples described in the introduction, all of which show that variable component slopes can yield aggregate power curves.[4]

It should be noted that the present simulations indicated a level of power-fit superiority that exceeds levels reported by R. B. Anderson and Tweney (1997). In that study, aggregate power curves emerged primarily when component–slope differences were large and systematic; *random* variability had only a slight impact on the aggregate power fit (also note that R. B. Anderson & Tweney's work dealt exclusively with exponential component curves). The present simulations are unique, however, in that *B* was constrained so that all component curves would be downward sloping. Apparently, such a constraint is not only reasonable (as noted earlier) but is also an important factor in producing power curves from exponential components.

Other Kinds of Distributions

Newell and Rosenbloom (1981) argued that it should be possible to create power curves from uniformly (i.e., rectangularly distributed exponential-component slopes). However, Wixted and Ebbesen (1997) suggested that Weibull-distributed exponentials do not readily yield aggregate power curves. Table 5 provides simulation results that confirm Wixted's point but that also demonstrate the ready emergence of aggregate logarithmic curves and the emergence of a single aggregate power curve. Moreover, Tables 5 and 6 demonstrate that range-limited linear and range-limited logarithmic components, whether Weibull distributed or uniformly distributed, still yield aggregate power curves.

Response Surface Analysis

The present paper focuses on traditional computational *simulation* to show that arithmetically averaged component curves—whether exponential, range-limited logarithmic, or range-limited linear—tend to yield power-like curves at the aggregate level. Very recently, Myung, Kim, and Pitt (2000) used another technique, *response surface analysis* (RSA), to make methodological points concerning power-curve emergence. (They studied emergence from exponential components—not from linear or logarithmic components.)

Table 4
Effect of Component Type and Intercomponent Slope Variability on the Number of Cases (n) in Which the Aggregate Curve Was Best Fit by an Exponential (Exp), Linear (Lin), Logarithmic (Log), or Power (Pow) Function

Component Type	Slope Variability	n			
		Exp	Lin	Log	Pow
Exponential	.0	500	0	0	0
	.1	500	0	0	0
	.2	500	0	0	0
	.3	500	0	0	0
	.4	471	0	0	29
	.5	1	0	188	311
Linear	.0	0	500	0	0
(range limited)	.1	0	0	498	2
	.2	1	0	173	326
	.3	9	0	92	399
	.4	–	–	–	–
	.5	–	–	–	–
Logarithmic	.0	0	0	500	0
(range limited)	.1	0	0	497	3
	.2	0	0	142	358
	.3	0	0	12	488
	.4	0	0	0	500
	.5	0	0	0	500
Power	.0	0	0	0	500
	.1	0	0	0	500
	.2	0	0	120	380
	.3	0	0	52	448
	.4	0	0	21	479
	.5	0	0	6	494

Note—In the case of linear components, when the amount of slope variability was very large, there were cases in which the aggregate curve contained zeros, precluding the log transform method of curve fitting. The components' A parameters are fixed at 0.5.

The RSA approach is applicable to the present work and offers an alternative means of drawing conclusions similar to those obtained via traditional simulation.

Following Myung et al. (2000), two-dimensional response surfaces may be constructed by first simplifying the equations in Table 2 so that the A parameter is omitted (or, equivalently, fixed at 1.0). For example, in the case of an exponential function,

$$Y = (A)e^{-BX},\ -B < 0$$

is simplified to

$$Y = e^{-BX},\ -B < 0.$$

The next step is to compute the value of Y (i.e., the "response" in "RSA") for two arbitrary values of X ($X_1 = 2$ and $X_2 = 10$), with $-B$ also set to an arbitrary value. The resulting values of Y are referred to as Y_1 and Y_2. Thus Y_1 and Y_2 define a point in a two-dimensional space. Next, the operation is repeated for different values of $-B$—keeping the same two values for X_1 and X_2 (in the present analysis, B varies across 1,000 values, from .01 to 10.0). The result is a set of points that define a curve, or surface, in two-dimensional space (see Myung et al., 2000, for a discussion of the surface's formal analytic properties). Figure 4 shows four such surfaces, constructed from component-curve functions that exclude A as a parameter. It is important to note that these surfaces are not forgetting curves—the forgetting curves are only implicitly represented in the figure, as pairs of Y_1 Y_2 values. Rather, the surfaces represent the range of responses (Y values) that can be produced by a given function (i.e., a given type of trace-level forgetting curve).

The surfaces in Figure 4 can be used as analytic tools to demonstrate power-curve emergence: For any of the four surfaces one can visualize the arithmetic averaging of two component curves by choosing any two arbitrary points, i

Table 5
Mean R^2 and Number of Best Fits (n) to Exponential (Exp), Linear (Lin), Logarithmic (Log), and Power (Pow) Functions, for Aggregate Curves With Weibull-Distributed B Components

Component Type	Fitted Function							
	Exp		Lin		Log		Pow	
	R^2	n	R^2	n	R^2	n	R^2	n
Exp	.95	0	.83	0	.99	499	.86	1
Lin	.92	4	.71	0	.96	98	.98	398
Log	.82	0	.72	0	.96	0	.99	500

Note—Mean R^2 is for the entire set of 500 aggregate curves (for each component type). The Weibull population distribution has shape and range parameters of 1.0 and 0.25, respectively, with a mean of -0.25 (sampled values were multiplied by -1.0) and a standard deviation of 0.25. The uniform population has a mean and standard deviation of -0.25 and 0.125, respectively. Linear and logarithmic component curves are range limited.

Table 6
Mean R^2 and Number of Best Fits (n) to Exponential (Exp), Linear (Lin), Logarithmic (Log), and Power (Pow) Functions, for Aggregate Curves With Uniformly Distributed B Components

Component Type	Fitted Function: Exp R^2	Exp n	Lin R^2	Lin n	Log R^2	Log n	Pow R^2	Pow n
Exp	.95	0	.78	0	.98	498	.96	11
Lin	.88	2	.57	0	.87	98	.99	498
Log	.87	1	.75	0	.97	1	.99	499

Note—Mean R^2 is for the entire set of 500 aggregate curves (for each component type). The uniform population distribution has a mean and standard deviation of −0.25 and 0.125, respectively. Linear and logarithmic component curves are range limited.

and j (each point in Figure 4 actually represents an entire component curve) and interpolating a straight line between the two points. The point at the center of the line represents the arithmetic average of the two curves represented by arbitrary points i and j. Thus, Figure 4 can be used as an analytic tool to demonstrate the following consequences of arithmetic averaging: (1) For exponential component functions, the aggregate curve has an enhanced fit to a power function (e.g., Point P); (2) for range-limited linear and logarithmic components, power curves can emerge if some but not all components contain a Y_2 equal to zero; (3) when the range of B is extreme, the aggregate curve tends to fit a power function better than it fits other common functions, even if the power fit is not very good (e.g., Point Q).[5]

Qualifiers

The present analyses assume that at the component level, retention tends *not* to increase over time and tends *not* to take on a negative value. Because of these restrictions, manipulation of interslope variability tends to alter the shape of the distribution of slopes. Thus, it is possible that *slope-distribution shape* is the proximal cause of power-curve emergence, with slope variability being an indirect or even noncritical factor. In principle, it may be possible to manipulate shape while holding variance constant, but then shape would still be confounded with other measures of variability, such as range. The present work addresses the shape issue to some degree by showing power-curve emergence in Weibull and rectangular distributions (in addition to range-limited normal distributions). Futhermore, R. B. Anderson and Tweney (1997) demonstrated a degree of power-curve emergence in non-range-limited distributions.

Conclusions

The present simulations and RSAs support the hypothesis that component curves with highly variable slopes

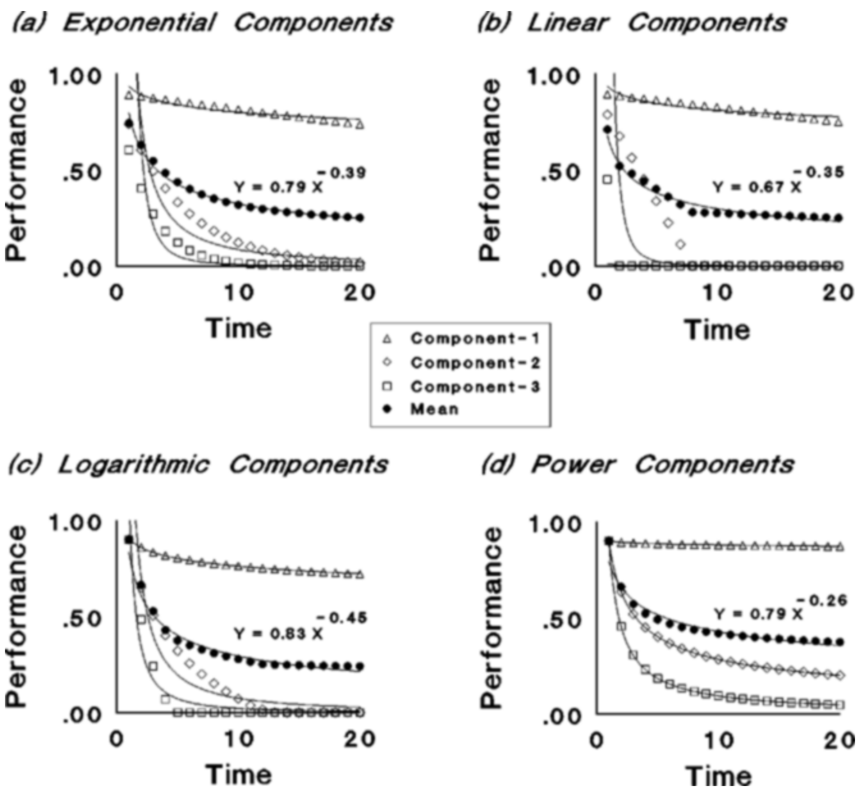

Figure 3. Mean R^2 as a function of intercomponent slope variability (the standard deviation of the random deviate distribution) and of fit type. Each panel shows simulation results for a particular type of component curve.

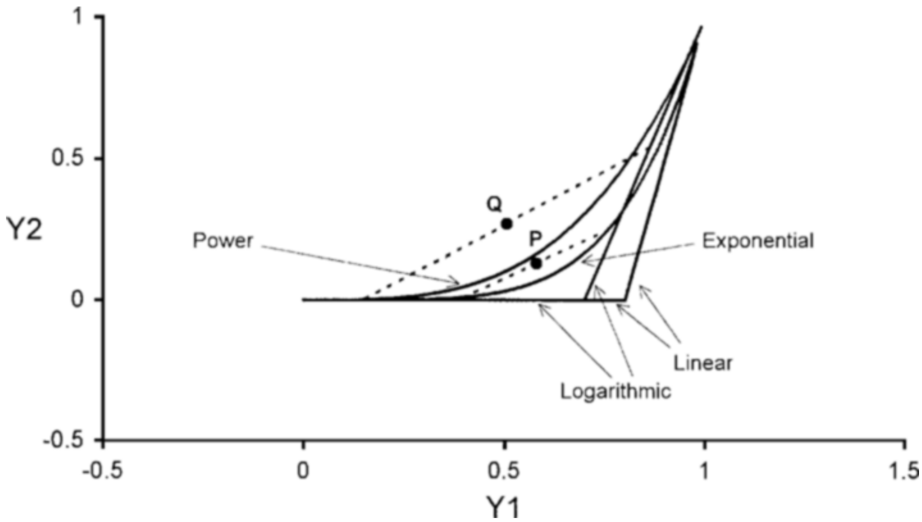

Figure 4. Two-dimensional response surface analysis. Point P represents the arithmetic average of two exponential curves. Point Q represents the arithmetic average of two range-limited logarithmic curves.

tend to produce approximations to power curves at the aggregate level—whether those component slopes are exponential, range-limited linear, or range-limited logarithmic. In addition, the simulations show that when the components themselves are power curves, the aggregate curve tends to be either power or logarithmic, depending on the amount of variability and depending on the value of the *A* parameter.

This conclusion does not deny the power function's descriptive utility. Instead, the function can be reconceptualized as indicating a much deeper generalization: That is, it may be that the power law is ubiquitous because slope variability is ubiquitous. Moreover, occasional or even frequent violations of the power law (i.e., exponential forgetting, Wickelgren, 1972; or exponential skill acquisition, Heathcote, Brown, & Mewhort, 2000) need not undermine the power function's validity. Indeed, adherence to the power law might turn out to be valuable as an index of the magnitude of intercomponent slope variability.

Power-curve emergence also has clear theoretical significance, particularly if the component curves are aggregated *within* a system rather than *across* subjects or experimental conditions.

For example, a multiple-trace memory system, such as Minerva 2 (Hintzman, 1986, 1988), will likely exhibit different decay rates for different traces. Moreover, the Minerva 2 model additively combines individual traces into theoretically meaningful composites (Hintzman called them "echoes"). Thus, a power law of forgetting might arise from the additive combination of forgetting curves across multiple traces.

Emergent power curves also have theoretical significance with regard to production systems in learning. For example, in John Anderson's ACT theory (and its various permutations in J. R. Anderson, 1982, 1993), cognitive and behavioral subprocedures, or productions, can operate sequentially with one another and can thereby produce additive effects on completion time for a given task. Consequently, insofar as different productions might possess very different learning rates, the *power law* of learning could arise solely from variable learning rates across productions. Indeed, Newell and Rosenbloom (1981) suggested such production variability as one possible explanation for the power law. The present findings bolster the plausibility of such an explanation, given that power-curve emergence seems not to require that the component curves have a particular mathematical form (e.g., the components do not have to be exponential). The evidence thus far suggests that sufficient slope variability (and/or the resultant distortion of the distribution of slopes), along with appropriate range limitation of components, may be primary preconditions for power-curve emergence.

REFERENCES

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

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, **2**, 396-408.

Anderson, R. B., & Tweney, R. D. (1997). Artifactual power curves in forgetting. *Memory & Cognition*, **25**, 724-730.

Anderson, R. B., Tweney, R. D., Rivardo, M., & Duncan, S. (1997). Need probability affects retention: A direct demonstration. *Memory & Cognition*, **25**, 867-872.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, **53**, 134-140.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, **7**, 185-207.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review,* **95**, 528-551.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review,* **95**, 492-527.

Myung I. J., Kim, C., & Pitt, M. A. (2000). Toward an explanation of the power law artifact: Insights from response surface analysis. *Memory & Cognition,* **28**, 832-840.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 57-84). Hillsdale, NJ: Erlbaum.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Rubin, D. C. (1982). On the retention function for autobiographical memory. *Journal of Verbal Learning & Verbal Behavior*, **21**, 21-38.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review,* **103**, 734-760.

Stevens, S. S. (1971). Neural events and the psychophysical law. *Science,* **170**, 1043-1050.

Wickelgren, W. A. (1972). Trace-resistance and the decay of long-term memory. *Journal of Mathematical Psychology,* **9**, 418-455.

Wickens, T. D. (1998). On the form of the retention function: Comment on Rubin and Wenzel (1996): A qualitative description of retention. *Psychological Review,* **105**, 379-386.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science,* **2**, 409-415.

Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition*, **25**, 731-739.

NOTES

1. For purposes of simplicity and consistency, the term *slope* is used, throughout, to refer to the *B* parameter.

2. Here and throughout, the term *form* refers to a curve's *general mathematical* form (e.g., exponential, logarithmic)—not the precise form that includes specific values for the parameters.

3. The arithmetic averaging of non-range-limited logarithmic curves produces a logarithmic aggregate curve (Estes, 1956). In addition, geometric averaging of exponential-curve or power-curve components preserves their general mathematical form so that the aggregate curve is also an exponential curve (R. B. Anderson & Tweney, 1997; Estes, 1956).

4. Though the present simulations focused exclusively on forgetting, the results are applicable to any domain in which behavior varies continuously as a function of some independent variable. Such domains include other areas of psychology (e.g., skill acquisition and psychophysics) and areas within economics. The findings may also have applicability in the biological and physical sciences.

5. Having demonstrated the relevance of RSA to the present findings, I must point out that the simulations presented in the main body of this paper involved fitting aggregate curves to *two*-parameter functions (the *A* parameter was fixed for each simulated component but was free to vary in the fitting of the aggregate curves to mathematical functions). In contrast, the two-dimensional RSA assumes one-parameter functions—at both the component level and the aggregate level.

(Manuscript received May 12, 1999;
revision accepted for publication April 28, 2001.)