JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2008, **89,** 119–123 NUMBER 1 (JANUARY)

CONTROL, PREDICTION, ORDER, AND THE JOYS OF RESEARCH

JOHN A. NEVIN

UNIVERSITY OF NEW HAMPSHIRE

JEAB's first decade featured control of individual behavior by operant contingencies, where the experimenter could interact with a subject in real time and obtain fairly immediate evidence of control, with the possibility of direct extension to applied settings. In subsequent decades, emphasis shifted toward long-term parametric studies with quantitative analyses of individual and group data. As a result, the reinforcers for the experimenter shifted from controlling behavior to uncovering and describing order, often using mathematical expressions. The same sorts of reinforcers are available for quantitative descriptions of aggregate behavioral data that may inform public policy.

Key words: order, quantification, model of war duration

The invitation to prepare this essay prompted me to review some early volumes of *JEAB*, so I pulled down the dusty, well-thumbed issues at the top left of my shelves and settled in to revisit them and try to recapture what it felt like to participate in the experimental analysis of behavior 40–50 years ago. What struck me most about *JEAB*'s first decade was an exuberant sense of the power of operant contingencies relating stimuli, responses, and reinforcers. In Volumes 1–10, the various creatures whose behavior came under the control of contingencies included the alligator, chicken, fighting cock, fighting fish, goat, guinea pig, horse, octopus, porpoise, sea lion, turtle, and vulture, as well as the usual suspects—rats, pigeons, laboratory primates, and humans. And although the conventional operants—lever presses and key pecks—were featured in the majority of research reports, contingencies also were applied to preening by pigeons, barking by dogs, mewing by cats, and small-scale covert muscle twitches and galvanic skin responses by humans.

Of greater applied interest were activities such as imitating, reading, stuttering, thumb-sucking, and classroom compliance in humans of diverse ages. Positive reinforcers included access to adjunctive activities such as grooming and fighting, and sensory events such as flashing lights or the sight of another creature, as well as consumables such as food or water. Negative reinforcers included noise, tail-pinch, and wind, as well as bright light and electric shock. Many studies showed that the behavior engendered by response–reinforcer contingencies came under the control of discriminative stimuli, which included interoceptive gastric cues, type of food pellet, electric shock, and a particular rat, as well as the usual lights and tones. And the subjects did heroic things: A rat avoided shocks in a continuous 43-day session, a pigeon pecked for food on various schedules in a 17-day session, and a chimpanzee made it through FR 120,000 (with some help from conditioned reinforcers). The entire enterprise was suffused with a sense of adventure, and it looked as if we could apply the technology of contingencies to pretty much anything.

Another powerful impression was the sense of communal scientific enterprise. The apparatus notes are the essence of communality: Here's a gimmick that worked in my lab—maybe you can use it too—and if it helps you to publish in my research area, that's OK. The notes also conveyed a sense of playful discovery: Their authors described all sorts of clever, easy-to-build or off-the-shelf devices that could be used to detect responses, program reinforcers, deliver shocks, and avoid ''fixed-ratio screwing'' while making cable connections (Verhave, 1958). It was always intriguing to encounter a novel solution to a problem, and it is good to see that the tradition is still alive in 2007.

The first article in the first issue of the next decade was the report of autoshaping the

---

Address correspondence to John A. Nevin, 20 Harborview Lane, Vineyard Haven, MA 02568 (e-mail: jnevin@cisunix.unh.edu). I thank Nora Nevin for her comments on an earlier draft, and Elaine Cyr for her collaboration in the analysis shown in Figure 1.

doi: 10.1901/jeab.2008.89-119

pigeon's key peck by Brown and Jenkins (1968). Because the rate of key pecking could be made to vary over a wide range by schedule contingencies, it appeared to be the ideal operant, so it was puzzling to see it emerge in a respondent conditioning paradigm. And because it was a directed response of the whole organism, as opposed to the usual salivary or eyelid response, autoshaped pecking was puzzling for standard accounts of classical conditioning as well. The initial report triggered a series of studies, and it soon became clear that autoshaped pecking persisted despite reinforcer omission (Williams & Williams, 1968), and that the topography of the response evoked by the reinforcer (e.g., food or water) determined the form of the autoshaped peck (Jenkins & Moore, 1973). Although these results challenged a naïve faith in the universal power of operant contingencies, they led to an appreciation of biological and evolutionary factors in operant–respondent interactions, and suggested that we had to consider the topographical compatibility of CR and UR in order to understand the effects of response-contingent reinforcers.

A very different sort of development was taking place concurrently: The analysis of steady-state performance maintained by various schedules of reinforcement. A number of studies examined performance maintained by variations in the classical schedules—FR, FI, VR, and VI. Some articles in the early issues introduced temporally defined schedules that mimicked some features of response-based schedule definitions, or arranged interlocking schedules to bridge the ratio–interval divide. Dozens of related studies explored the interactions between punishment and reinforcement schedules, or parametric effects of escape and avoidance schedules. Concurrent VI VI schedules revealed spectacular order when the data were expressed as relative measures, and some studies used mathematical expressions to describe their results. All of these developments within *JEAB*'s first decade foreshadowed the parametric, systematic, quantitative analyses of schedule performance that have become so prominent in its later years.

I am not going to review all of the various research topics that emerged in *JEAB*'s early pages that then were explored over subsequent years. Instead, I want to consider what maintains our behavior as practitioners of the experimental analysis of behavior. In ''A case history in scientific method,'' Skinner (1959) noted that ''The organism whose behavior is most extensively modified and most completely controlled in research of the sort I have described is the experimenter himself'' (p. 98). Skinner was referring to analyses like those that appeared in the early volumes of *JEAB*, where the experimenter interacted with the subject over fairly brief time periods: Changes in the experimenter's behavior were prompted by some aspect of the subject's behavior and then reinforced by clear evidence of experimental control. Relatedly, Sidman (2007) described his own excitement in the orderliness of avoidance behavior, and in the seemingly magical emergence of stimulus equivalence, but also noted that ''When we publish our research findings, we are not allowed to communicate the thrill of research … or the exhilaration in the discovery of order'' (p. 312). In fact, I think the thrill of research in *JEAB*'s early years was evident without explicit description: Most of the articles were brief and engaging to read, and the excitement in the discoveries that led to publication could be savored at least vicariously because the procedures and results made contact with the experiences of readers, including students whose rat-lab courses included shaping, discrimination, chaining, and basic schedules.

It is a lot harder for a researcher to experience the excitement of discovery from parametric studies where experiments involve many conditions, some lasting for 100 sessions, and where the richness of the results cannot be appreciated until all those data have been summarized, digested, transformed, and fitted by theoretical models in order to estimate the values of model parameters. The problem, of course, is that the variables that control behavior cannot be identified by real-time give-and-take between subject and experimenter, and ''the exhilaration in the discovery of order'' is distributed over weeks or months of data analysis. At the end of the process, the most thrilling result may be the invariance of a parameter value rather than its systematic change, but it is not easy to convey the excitement of finding invariance after shuffling through a pile of numbers. The experimental analysis of behavior has become a mature science, and its youthful enthusiasm engendered by real-time experimental control may be impossible to recapture.

If scientific behavior depends on its antecedents and its consequences, including the scientist's private delight in the unmasking of order or invariance, then researchers who have long histories of obtaining orderly results are likely to keep on doing similar research in the future. When Skinner (1986) wrote ''Some thoughts about the future,'' the ideas and research projects that he recommended for further exploration were rather direct extensions of work that he and his colleagues had done three or four decades earlier, of the sort he described in ''A case history in scientific method'' (1959). He acknowledged as much. Likewise, my history of engagement in quantitative analyses leads me to suggest that the field has made major advances in establishing orderly relations involving behavior controlled by antecedent stimuli and schedules of reinforcement, but many empirical issues must still be addressed; and that promising models embracing many aspects of discriminated operant behavior have been developed, but they address fairly narrowly defined paradigms. The quantitative analysis of discriminated operant behavior will (or should) continue in the future, and I predict that data will be more elegantly quantified and that unifying models will be developed that have some chance of addressing the complexities of human behavior in the real world.

Society supports the science of behavior to the extent that the science yields findings and applications that can be used to make life better, and we are members of society at large as well as behavioral scientists. For example, if it had turned out that my work on the principles of resistance to change applied only to pigeons pecking lighted keys for food, would I have persisted in this line of work, even if it turned out that these principles could be linked to unique aspects of the pigeon's neurobiology and genetic heritage? I doubt it. For me at least, demonstrations of generality from nonhumans to humans have been important for maintaining my scientific behavior, and such demonstrations are probably crucial for the health of the field as a whole, including its ability to attract new researchers and funding for their work.

In bringing our work to bear on everyday human concerns, there is some tension between a strict technical language that attempts to describe behavioral phenomena in objective ways and the commonsense vocabulary of daily life. Our specialized vocabulary is superb for communication within the behavior-analytic community, but it keeps us from contacting new audiences that could broaden support of our endeavor. Another concern is that our specialized objective vocabulary may stand in the way of our own work. We are organisms who study the behavior of other organisms, and excessive concern with the technical vocabulary may have the effect of preventing empathy with our subjects. After all, the subject is always right, and when a pigeon apparently behaves contrary to expectation based on currently accepted principles or model predictions, one way to figure out what is going on is to put one's self into the pigeon's head, so to speak —and then translate the resulting insights into an objective experimental analysis. Presumably, our everyday vocabulary has been selected for its effectiveness in characterizing experiences that are common to humankind, and perhaps to other creatures as well, so it can be quite helpful in appreciating the subject's perspective on its actions. Why not allow that vocabulary to guide scientific exploration (without weakening standards of evidence or clarity of interpretation) and to make objectively stated results more broadly accessible? In the end, the results are what count, not the words that characterize them.

I noted above that *JEAB* has become substantially more quantitative and mathematical over the course of its 50 years. When results are characterized by equations with clearly defined variables, it seems to me that debates over the proper language of behavior analysis are simply bypassed. If the experimenter arranges x, then the subject does y, where the value of y may be modulated by parameters a, b, and c. The names of the parameters have absolutely no effect on the equations, but may provide ideas for future analyses via their commonsense connotations. Moreover, readers may be more likely to plough through the transformations of the data and the derivation of the equations, and thus truly appreciate the order that emerges, if the problem being

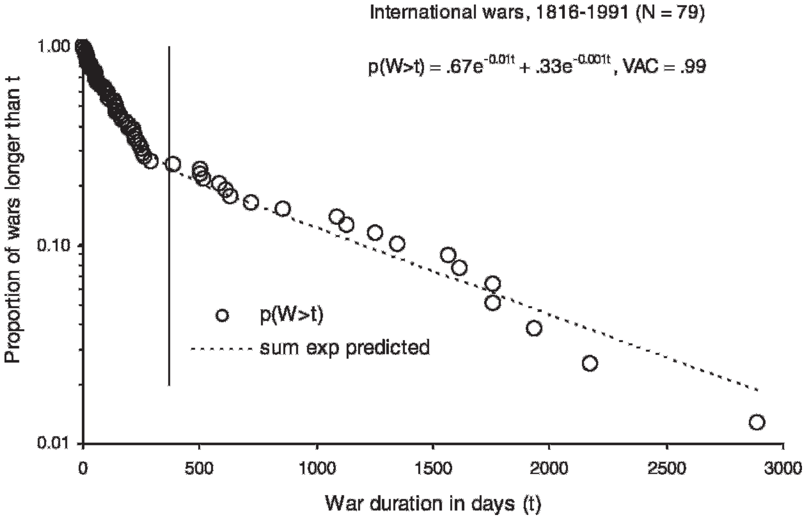

Fig. 1. Survivor plot of the durations of all international wars between 1816 and 1991, based on a data set provided by the Correlates of War project. The data are fitted by the sum of two exponential functions with the parameter values indicated; the *y*-axis is logarithmic. The vertical line marks one year from war onset.

addressed is framed in a way that makes contact with everyday life.

The techniques of quantitative analysis can be extended to important problem areas that lie far beyond the behavior lab. For example, the durations of international wars between 1816 and 1991, taken from data sets compiled by the Correlates of War Project (2000) at the University of Michigan, can be described by the same analytic method that Shull, Grimes, and Bennett (2004) used for interresponse time (IRT) distributions. Figure 1 presents a survivor plot, on a log y axis, of the proportion of international wars that lasted at least t days (the term ''survivor'' is less than felicitous here because the survivors are wars, not the people they affect). There are two linear segments, suggesting that wars tend to end with a constant probability in time which changes abruptly from a fairly high value to a substantially lower value at about one year. The plot looks much like the ''broken stick'' form of the IRT distributions reported by Shull and his colleagues. Accordingly, I fitted the summed-exponential model that they used:

$$\mathrm{P}(\mathrm{W} > \mathrm{t}) = p\mathrm{e}^{-\mathrm{st}} + (1 - p)\mathrm{e}^{-\mathrm{lt}}, \qquad (1)$$

where P(W>t) is the proportion of all wars having a duration greater than t days, the term $p\mathrm{e}^{-\mathrm{st}}$ represents the contribution of ''short'' wars in the steep left-hand segment of the plot, and $(1-p)\mathrm{e}^{-\mathrm{lt}}$ represents the contribution of ''long'' wars in the shallower right-hand segment. The parameters and variance accounted for are given in the figure. Clearly, war durations are at least as orderly as rat's IRTs (cf. Shull et al., Figure 5).

Shull et al. (2004) suggested that the break in the survivor plot of IRTs represented a transition from one sort of activity (high-rate within-bout bursts) to another (bout initiation). That cannot work for war durations, but a different sort of interpretation is suggested by the fact that the break occurs about one year after war onset. The first and subsequent anniversaries provide occasions for political leaders to exhort their people to carry on, such as ''We must honor the brave young men and women who have given their lives for our great cause by continuing the fight until the tenacious enemy is defeated.'' This, of course, is a tragic example of the sunk-cost fallacy, which may be evident in pigeons (Navarro & Fantino, 2005; de la Piedad, Field, & Rachlin, 2006) as well as political leaders in wartime.

The quantitative model described here can be applied to the war in Iraq, currently in its fifth year. Using the parameters in Figure 1, I estimate that the conditional probability of ending the war in Iraq between June 1, 2007 (as I write) and January 1, 2008 (when *JEAB* turns 50) is .04. I also estimate that the probability that the war will last at least 2912 days, the duration of US combat involvement in Vietnam, is .05 (assuming no major changes in public policy).[1] This example shows how a combination of orderly data and a quantitative model permits at least probabilistic prediction, and I confess to feeling a genuine thrill when I entered the summed-exponential model into an Excel spreadsheet and saw the fitted line snap into place on top of the data points in Figure 1.

JEAB was established 50 years ago to publish ‘‘experiments relevant to the behavior of individual organisms,’’ and its first 10 volumes featured the control of ongoing behavior by operant contingencies. Figure 1, by contrast, describes the aggregate historical behavior of nations. It would be marvelous if analyses of this sort, extending quantitative methods and findings from individual behavior to larger domains, could be brought to bear on public policy in matters of supreme importance. Now *that* would be exciting.

REFERENCES

Bennett, D. S., & Stam, A. C. (2006). Predicting the length of the 2003 U.S. Iraq war. *Foreign Policy Analysis, 2,* 101–116.

Brown, P. L., & Jenkins, H. M. (1968). Autoshaping of the pigeon’s key-peck. *Journal of the Experimental Analysis of Behavior, 11,* 1–8.

Correlates of War Project. (2000). *Inter-State War* (Version 3.0) [Data file]. Retrieved June 13, 2007, from http://www.correlatesofwar.org

de la Piedad, X., Field, D., & Rachlin, H. (2006). The influence of prior choices on current choice. *Journal of the Experimental Analysis of Behavior, 85,* 3–21.

Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior, 20,* 163–181.

Navarro, A. D., & Fantino, E. (2005). The sunk cost effect in pigeons and humans. *Journal of the Experimental Analysis of Behavior, 83,* 1–13.

Shull, R. L., Grimes, J. A., & Bennett, A. J. (2004). Bouts of responding: The relation between bout rate and the rate of variable-interval reinforcement. *Journal of the Experimental Analysis of Behavior, 81,* 65–83.

Sidman, M. (2007). The analysis of behavior: What’s in it for us? *Journal of the Experimental Analysis of Behavior, 87,* 309–316.

Skinner, B. F. (1959). A case history in scientific method. In *Cumulative record,* New York: Appleton-Century-Crofts. (Reprinted from Skinner, B. F. [1956]. *American Psychologist, 11,* 221–223.)

Skinner, B. F. (1986). Some thoughts about the future. *Journal of the Experimental Analysis of Behavior, 45,* 229–235.

Verhave, T. (1958). New type of connector plugs and sockets. *Journal of the Experimental Analysis of Behavior, 1,* 86.

Williams, D. R., & Williams, H. (1968). Auto-maintenance in the pigeon: Sustained pecking despite contingent nonreinforcement. *Journal of the Experimental Analysis of Behavior, 12,* 511–520.

---

[1] For a more detailed analysis that includes a number of additional factors that affect war durations, see Bennett and Stam (2006). According to their model, the expected duration of the war in Iraq is 83 months.