Skinner, B.F. (1956)

8

A Case History in Scientific Method

It has been said that college teaching is the only profession for which there is no professional training, and it is commonly argued that this is because our graduate schools train scholars and scientists rather than teachers. We are more concerned with the discovery of knowledge than with its dissemination. But can we justify ourselves quite so easily? It is a bold thing to say that we know how to train a man to be a scientist. Scientific thinking is the most complex and probably the most subtle of all human activities. Do we actually know how to shape up such behavior, or do we simply mean that some of the people who attend our graduate schools eventually become scientists?

Except for a laboratory course which acquaints the student with standard apparatus and standard procedures, the only explicit training in scientific method generally received by a young psychologist is a course in statistics—not the introductory course, which is often required of so many kinds of students that it is scarcely scientific at all, but an advanced course which includes "model building," "theory construction," and "experimental design." But it is a mistake to identify scientific practice with the formalized constructions of statistics and scientific method. These disciplines have their place, but it does not coincide with the place of scientific research. They offer *a* method of science but not, as is so often implied, *the* method. As formal disciplines they arose very late in the history of science, and most of the facts of science have been discovered without their aid. It takes a great deal of skill to fit Faraday with his wires and magnets into the picture which statistics gives us of scientific thinking. And most current scientific practice would be equally refractory, especially in the important initial stages. It is no wonder that the laboratory scientist is puzzled and often dismayed when he discovers how his behavior has been reconstructed in

From *American Psychologist*, 1956, *11*, 221–233.

the formal analyses of scientific method. He is likely to protest that this is not at all a fair representation of what he does.

But his protest is not likely to be heard. For the prestige of statistics and scientific methodology is enormous. Much of it is borrowed from the high repute of mathematics and logic, but much of it derives from the flourishing state of the art itself. Some statisticians are professional people employed by scientific and commercial enterprises. Some are teachers and pure researchers who give their colleagues the same kind of service for nothing—or at most a note of acknowledgment. Many are zealous people who, with the best of intentions, are anxious to show the nonstatistical scientist how he can do his job more efficiently and assess his results more accurately. There are strong professional societies devoted to the advancement of statistics, and hundreds of technical books and journals are published annually.

Against this, the practicing scientist has very little to offer. He cannot refer the young psychologist to a book which will tell him how to find out all there is to know about a subject matter, how to have the good hunch which will lead him to devise a suitable piece of apparatus, how to develop an efficient experimental routine, how to abandon an unprofitable line of attack, how to move on most rapidly to later stages of his research. The work habits which have become second nature to him have not been formalized by anyone, and he may feel that they possibly never will be. As Richter [1] has pointed out, "Some of the most important discoveries have been made without any plan of research," and "there are researchers who do not work on a verbal plane, who cannot put into words what they are doing."

If we are interested in perpetuating the practices responsible for the present corpus of scientific knowledge, we must keep in mind that some very important parts of the scientific process do not now lend themselves to mathematical, logical, or any other formal treatment. We do not know enough about human behavior to know how the scientist does what he does. Although statisticians and methodologists may seem to tell us, or at least imply, how the mind works—how problems arise, how hypotheses are formed, deductions made, and crucial experiments designed—we as psychologists are in a position to remind them that they do not have methods appropriate to the empirical observation or the functional analysis of such data. These are aspects of human behavior, and no one knows better than we how little can at the moment be said about them.

[1] Richter, C. P. Free research versus design. *Science*, 1953, *118*, 91–93.

Some day we shall be better able to express the distinction between empirical analysis and formal reconstruction, for we shall have an alternative account of the behavior of Man Thinking. Such an account will not only plausibly reconstruct what a particular scientist did in any given case, it will permit us to evaluate practices and, I believe, to teach scientific thinking. But that day is some little distance in the future. Meanwhile we can only fall back on examples.

When the director of Project A of the American Psychological Association asked me to describe and analyze my activities as a research psychologist, I went through a trunkful of old notes and records and, for my pains, reread some of my earlier publications. This has made me all the more aware of the contrast between the reconstructions of formalized scientific method and at least one case of actual practice. Instead of amplifying the points I have just made by resorting to a generalized account (principally because it is not available), I should like to discuss a case history. It is not one of the case histories we should most like to have, but what it lacks in importance is perhaps somewhat offset by accessibility. I therefore ask you to imagine that you are all clinical psychologists—a task which becomes easier and easier as the years go by—while I sit across the desk from you or stretch out upon this comfortable leather couch.

The first thing I can remember happened when I was only twenty-two years old. Shortly after I was graduated from college Bertrand Russell published a series of articles in the old *Dial* magazine on the epistemology of John B. Watson's Behaviorism. I had had no psychology as an undergraduate but I had had a lot of biology, and two of the books which my biology professor had put into my hands were Loeb's *Physiology of the Brain* and the newly published Oxford edition of Pavlov's *Conditioned Reflexes*. And now here was Russell extrapolating the principles of an objective formulation of behavior to the problem of knowledge! Many years later when I told Lord Russell that his articles were responsible for my interest in behavior, he could only exclaim, "Good Heavens! I had always supposed that those articles had demolished Behaviorism!" But at any rate he had taken Watson seriously, and so did I.

When I arrived at Harvard for graduate study, the air was not exactly full of behavior, but Walter Hunter was coming in once a week from Clark University to give a seminar, and Fred Keller, also a graduate student, was an expert in both the technical details and the sophistry of Behaviorism. Many a time he saved me as I sank into the quicksands of an amateurish discussion of "What is an image?" or "Where is red?" I soon came into contact with W. J. Crozier, who had studied under Loeb. It had

been said of Loeb, and might have been said of Crozier, that he "resented the nervous system." Whether this was true or not, the fact was that both these men talked about animal behavior without mentioning the nervous system and with surprising success. So far as I was concerned, they cancelled out the physiological theorizing of Pavlov and Sherrington and thus clarified what remained of the work of these men as the beginnings of an independent science of behavior. My doctoral thesis was in part an operational analysis of Sherrington's synapse, in which behavioral laws were substituted for supposed states of the central nervous system.

But the part of my thesis at issue here was experimental. So far as I can see, I began simply by looking for lawful processes in the behavior of the intact organism. Pavlov had shown the way; but I could not then, as I cannot now, move without a jolt from salivary reflexes to the important business of the organism in everyday life. Sherrington and Magnus had found order in surgical segments of the organism. Could not something of the same sort be found, to use Loeb's phrase, in "the organism as a whole"? I had the clue from Pavlov: control your conditions and you will see order.

It is not surprising that my first gadget was a silent release box, operated by compressed air and designed to eliminate disturbances when introducing a rat into an apparatus. I used this first in studying the way a rat adapted to a novel stimulus. I built a soundproofed box containing a specially structured space. A rat was released, pneumatically, at the far end of a darkened tunnel from which it emerged in exploratory fashion into a well-lighted area. To accentuate its progress and to facilitate recording, the tunnel was placed at the top of a flight of steps, something like a functional Parthenon (Figure 1). The rat would peek out from the tunnel, perhaps glancing suspiciously at the one-way window through which I was watching it, then stretch itself cautiously down the steps. A soft click (carefully calibrated, of course) would cause it to pull back into the tunnel and remain there for some time. But repeated clicks had less and less of an effect. I recorded the rat's advances and retreats by moving a pen back and forth across a moving paper tape.

The major result of this experiment was that some of my rats had babies. I began to watch young rats. I saw them right themselves and crawl about very much like the decerebrate or thalamic cats and rabbits of Magnus. So I set about studying the postural reflexes of young rats. Here was a first principle not formally recognized by scientific methodologists: When you run onto something interesting, drop everything else and study it. I tore up the Parthenon and started over.

If you hold a young rat on one hand and pull it gently by the tail, it will

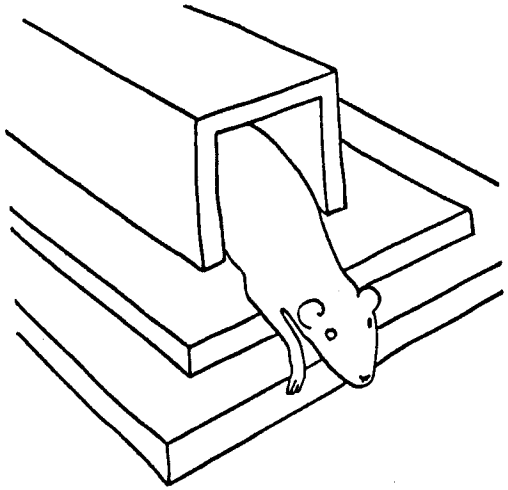

FIG. 1

resist you by pulling forward and then, with a sudden sharp spring wl usually disengages its tail, it will leap out into space. I decided to st this behavior quantitatively. I built a light platform covered with c and mounted it on tightly stretched piano wires (Figure 2). Here w version of Sherrington's torsion-wire myograph, originally designec

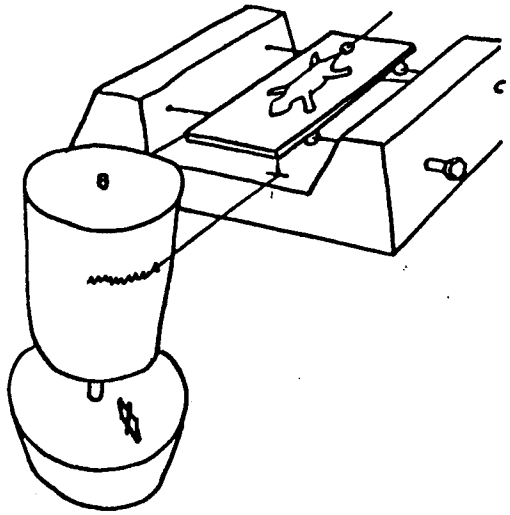

FIG. 2

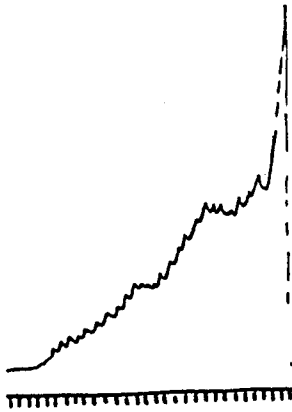

FIG. 3

record the isometric contraction of the *tibialis anticus* of a cat, but here adapted to the response of a whole organism. When the tail of the young rat was gently pulled, the rat clung to the cloth floor and tugged forward. By amplifying the fine movements of the platform, it was possible to get a good kymograph record of the tremor in this motion and then, as the pull against the tail was increased, of the desperate spring into the air (Figure 3).

Now, baby rats have very little future, except as adult rats. Their behavior is literally infantile and cannot be usefully extrapolated to everyday life. But if this technique would work with a baby, why not try it on a mature rat? To avoid attaching anything to the rat, it should be possible to record, not a pull against the substrate, but the ballistic thrust exerted as the rat runs forward or suddenly stops in response to my calibrated click. So, invoking the first principle of scientific practice again, I threw away the piano-wire platform and built a runway, eight feet long. This was constructed of light wood, in the form of a U girder, mounted rigidly on vertical glass plates, the elasticity of which permitted a very slight longitudinal movement (Figure 4). The runway became the floor of a long tunnel, not shown, at one end of which I placed my soundless release box and at the other end myself, prepared to reinforce the rat for coming down the runway by giving it a bit of wet mash, to sound a click from time to time when it had reached the middle of the runway, and to harvest kymograph records of the vibrations of the substrate.

Now for a second unformalized principle of scientific practice: Some ways of doing research are easier than others. I got tired of carrying the rat back to the other end of the runway. A back alley was therefore added (Figure 5). Now the rat could eat a bit of mash at point C, go down the back alley A, around the end as shown, and back home by runway B. The experimenter at E could collect records from the kymograph at D in comfort. In this way a great many records were made of the forces exerted against the substratum as rats ran down the alley and occasionally stopped dead in their tracks as a click sounded (Figure 6).

There was one annoying detail, however. The rat would often wait an inordinately long time at C before starting down the back alley on the next run. There seemed to be no explanation for this. When I timed these delays with a stop watch, however, and plotted them, they seemed to show orderly changes (Figure [illegible]). This was, of course, the kind of thing I was looking for. I forgot all about the movements of the substratum and began to run rats for the sake of the delay measurements alone. But there was now no reason why the runway had to be eight feet long and, as the second

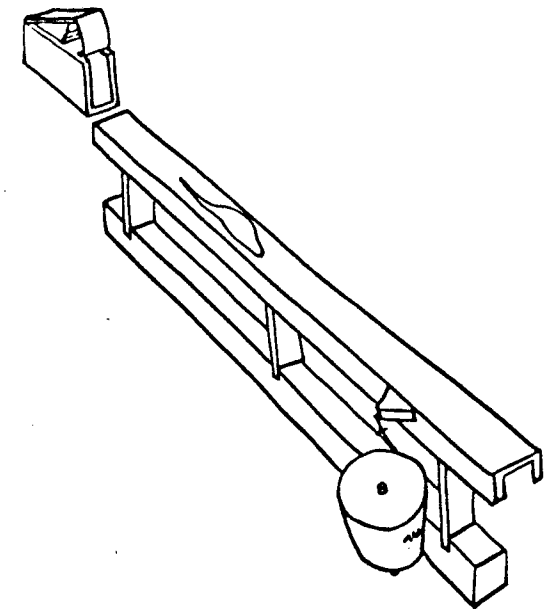

FIG. 4

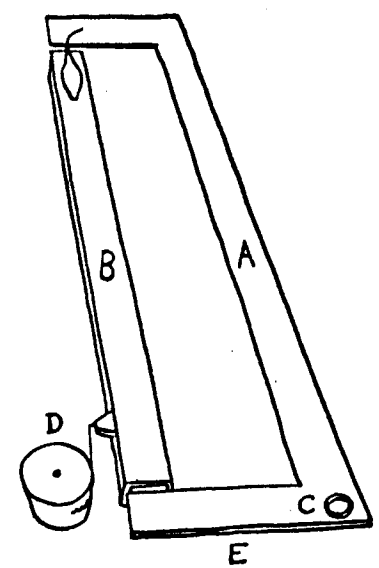

FIG. 5

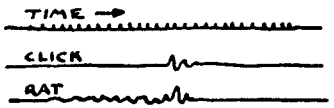

FIG. 6

FIG. 7

principle came into play again, I saw no reason why the rat could not deliver its own reinforcement.

A new apparatus was built. In Figure 8 we see the rat eating a piece of food just after completing a run. It produced the food by its own action. As it ran down the back alley A to the far end of the rectangular runway, its weight caused the whole runway to tilt slightly on the axis C and this movement turned the wooden disc D, permitting a piece of food in one of the holes around its perimeter to drop through a funnel into a food dish. The food was pearl tapioca, the only kind I could find in the grocery stores in reasonably uniform pieces. The rat had only to complete its journey by coming down the homestretch B to enjoy its reward. The experimenter was able to enjoy *his* reward at the same time, for he had only to load the magazine, put in a rat, and relax. Each tilt was recorded on a slowly moving kymograph.

A third unformalized principle of scientific practice: Some people are lucky. The disc of wood from which I had fashioned the food magazine was taken from a storeroom of discarded apparatus. It happened to have a central spindle, which fortunately I had not bothered to cut off. One day it occurred to me that if I wound a string around the spindle and allowed it to unwind as the magazine was emptied (Figure 9), I would get a different kind of record. Instead of a mere report of the up-and-down movement of the runway, as a series of pips as in a polygraph, I would get a *curve*. And

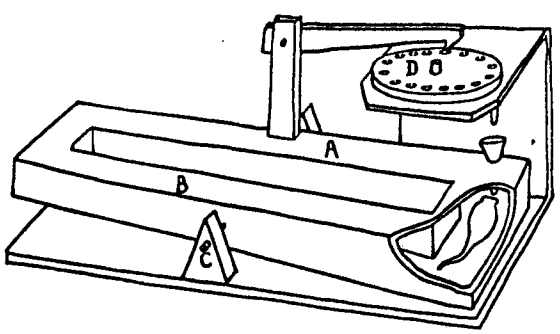

FIG. 8

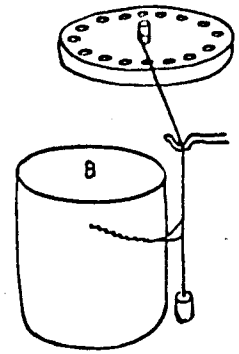

FIG. 9

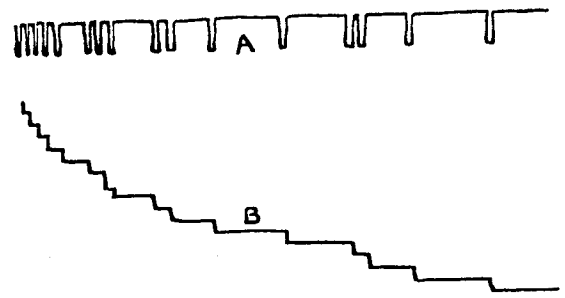

FIG. 10

I knew that science made great use of curves, although, so far as I could discover, very little of pips on a polygram. The difference between the old type of record at A (Figure 10) and the new at B may not seem gr but as it turned out the curve revealed things in the rate of responding, and in changes in that rate, which would certainly otherwise have been missed. By allowing the string to unwind rather than to wind, I had got my curve in an awkward Cartesian quadrant, but that was easily remedied. Psychologists have adopted cumulative curves only very slowly, but I think it is fair to say that they have become an indispensable tool for certain purposes of analysis.

Eventually, of course, the runway was seen to be unnecessary. The rat could simply reach into a covered tray for pieces of food, and each movement of the cover could operate a solenoid to move a pen one step in a cumulative curve. The first major change in rate observed in this way was due to indigestion. Curves showing how the rate of eating declined with the time of eating comprised the other part of my thesis. But a refinement was needed. The behavior of the rat in pushing open the door was not a normal part of the ingestive behavior of *Rattus rattus*. The act was obviously learned but its status as part of the final performance was not clear. It seemed wise to add an initial conditioned response connected with ingestion in a quite arbirary way. I chose the first device which came to hand—a horizontal bar or lever placed where it could be conveniently depressed by the rat to close a switch which operated a magnetic magazine. Ingestion curves obtained with this initial response in the chain were found to have the same properties as those without it.

Now, as soon as you begin to complicate an apparatus, you necessarily invoke a fourth principle of scientific practice: Apparatuses sometimes

break down. I had only to wait for the food magazine to jam to get an extinction curve. At first I treated this as a defect and hastened to remedy the difficulty. But eventually, of course, I deliberately disconnected the magazine. I can easily recall the excitement of that first complete extinction curve (Figure 11). I had made contact with Pavlov at last! Here was a

FIG. 11

curve uncorrupted by the physiological process of ingestion. It was an orderly change due to nothing more than a special contingency of reinforcement. It was pure behavior! I am not saying that I would not have got around to extinction curves without a breakdown in the apparatus; Pavlov had given too strong a lead in that direction. But it is still no exaggeration to say that some of the most interesting and surprising results have turned up first because of similar accidents. Foolproof apparatus is no doubt highly desirable, but Charles Ferster and I in recently reviewing the data from a five-year program of research found many occasions to congratulate ourselves on the fallibility of relays and vacuum tubes.

I then built four soundproofed ventilated boxes, each containing a lever and a food magazine and supplied with a cumulative recorder, and was on my way to an intensive study of conditioned reflexes in skeletal behavior. I would reinforce every response for several days and then extinguish for a day or two, varying the number of reinforcements, the amount of previous magazine training, and so on.

At this point I made my first use of the deductive method. I had long since given up pearl tapioca as too unbalanced a diet for steady use. A neighborhood druggist had shown me his pill machine, and I had had one made along the same lines (Figure 12). It consisted of a fluted brass bed across which one laid a long cylinder of stiff paste (in my case a MacCollum formula for an adequate rat diet). A similarly fluted cutter was then lowered onto the cylinder and rolled slowly back and forth, converting the paste into about a dozen spherical pellets. These were dried for a day or so before use. The procedure was painstaking and laborious. Eight rats eating a hundred pellets each per day could easily keep up with production. One pleasant Saturday afternoon I surveyed my supply of dry pellets and,

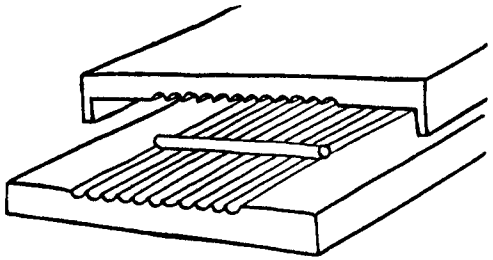

FIG. 12

appealing to certain elemental theorems in arithmetic, deduced that ur
I spent the rest of that afternoon and evening at the pill machine,
supply would be exhausted by ten-thirty Monday morning.

Since I do not wish to deprecate the hypothetico-deductive metho
am glad to testify here to its usefulness. It led me to apply our sec
principle of unformalized scientific method and to ask myself why *e*
press of the lever had to be reinforced. I was not then aware of what
happened at the Brown laboratories, as Harold Schlosberg later told
story. A graduate student had been given the task of running a cat thr
a difficult discrimination experiment. One Sunday the student found
supply of cat food exhausted. The stores were closed, and so, with a l
tiful faith in the frequency-theory of learning, he ran the cat as usual
took it back to its living cage unrewarded. Schlosberg reports that th
howled its protest continuously for nearly forty-eight hours. Unawa
this I decided to reinforce a response only once every minute and to
all other responses to go unreinforced. There were two results: (*a*
supply of pellets lasted almost indefinitely; and (*b*) each rat stab
at a fairly constant rate of responding.

Now, a steady state was something I was familiar with from ph
chemistry, and I therefore embarked upon the study of periodic rein
ment. I soon found that the constant rate at which the rat stabilize
pended upon how hungry it was. Hungry rat, high rate; less hungr
lower rate. At that time I was bothered by the practical problem o
trolling food deprivation. I was working half time at the Medical
(on chronaxie of subordination!) and could not maintain a good sc
in working with the rats. The rate of responding under periodic rein
ment suggested a scheme for keeping a rat at a constant level of depri
The argument went like this: Suppose you reinforce the rat, not at t
of a given period, but when it has completed the number of res

ordinarily emitted in that period. And suppose you use substantial pellets of food and give the rat continuous access to the lever. Then, except for periods when the rat sleeps, it should operate the lever at a constant rate around the clock. For, whenever it grows slightly hungrier, it will work faster, get food faster, and become less hungry, while whenever it grows slightly less hungry, it will respond at a lower rate, get less food, and grow hungrier. By setting the reinforcement at a given number of responses it should even be possible to hold the rat at any given level of deprivation. I visualized a machine with a dial which one could set to make available, at any time of day or night, a rat in a given state of deprivation. Of course, nothing of the sort happens. This is "fixed-ratio" rather than "fixed-interval" reinforcement and, as I soon found out, it produces a very different type of performance. This is an example of a fifth unformalized principle of scientific practice, but one which has at least been named. Walter Cannon described it with a word invented by Horace Walpole: *serendipity*—the art of finding one thing while looking for something else.

This account of my scientific behavior up to the point at which I published my results in a book called *The Behavior of Organisms* is as exact in letter and spirit as I can now make it. The notes, data, and publications which I have examined do not show that I ever behaved in the manner of Man Thinking as described by John Stuart Mill or John Dewey or in reconstructions of scientific behavior by other philosophers of science. I never faced a Problem which was more than the eternal problem of finding order. I never attacked a problem by constructing a Hypothesis. I never deduced Theorems or submitted them to Experimental Check. So far as I can see, I had no preconceived Model of behavior—certainly not a physiological or mentalistic one and, I believe, not a conceptual one. The "reflex reserve" was an abortive, though operational, concept which was retracted a year or so after publication in a paper at the Philadelphia meeting of the APA. It lived up to my opinion of theories in general by proving utterly worthless in suggesting further experiments. Of course, I was working on a basic Assumption—that there was order in behavior if I could only discover it—but such an assumption is not to be confused with the hypotheses of deductive theory. It is also true that I exercised a certain Selection of Facts but not because of relevance to theory but because one fact was more orderly than another. If I engaged in Experimental Design at all, it was simply to complete or extend some evidence of order already observed.

Most of the experiments described in *The Behavior of Organisms* were done with groups of four rats. A fairly common reaction to the book was

that such groups were too small. How did I know that other groups of four rats would do the same thing? Keller, in defending the book, countered with the charge that groups of four were too *big*. Unfortunately, however, I allowed myself to be persuaded of the contrary. This was due in part to my association at the University of Minnesota with W. T. Heron. Through him I came into close contact for the first time with traditional animal psychology. Heron was interested in inherited maze behavior, inherited activity, and certain drugs—the effects of which could then be detected only through the use of fairly large groups. We did an experiment together on the effect of starvation on the rate of pressing a lever and started the new era with a group of sixteen rats. But we had only four boxes, and this was so inconvenient that Heron applied for a grant and built a battery of twenty-four lever-boxes and cumulative recorders. I supplied an attachment which would record, not only the mean performance of all twenty-four rats in a single averaged curve, but mean curves for four subgroups of twelve rats each and four subgroups of six rats each.[2] We thus provided for the design of experiments according to the principles of R. A. Fisher, which were then coming into vogue. We had, so to speak, mechanized the Latin square.

With this apparatus Heron and I published a study of extinction in maze-bright and maze-dull rats using *ninety-five* subjects. Later I published mean extinction curves for groups of twenty-four, and W. K. Estes and I did our work on anxiety with groups of the same size. But although Heron and I could properly voice the hope that "the possibility of using large groups of animals greatly improves upon the method as previously reported, since tests of significance are provided for and properties of behavior not apparent in single cases may be more easily detected," in actual practice that is not what happened. The experiments I have just mentioned are almost all we have to show for this elaborate battery of boxes. Undoubtedly more work could be done with it and would have its place, but something had happened to the natural growth of the method. You cannot easily make a change in the conditions of an experiment when twenty-four apparatuses have to be altered. Any gain in rigor is more than matched by a loss in flexibility. We were forced to confine ourselves to processes which could be studied with the baselines already developed in earlier work. We could not move on to the discovery of other processes or even to a more refined analysis of those we were working with. No matter

[2] Heron, W. T., & Skinner, B. F. An apparatus for the study of behavior. *Psychol. Rec.*, 1939, *3*, 166–176.

how significant might be the relations we actually demonstrated, our statistical Leviathan had swum aground. The art of the method had stuck at a particular stage of its development.

Another accident rescued me from mechanized statistics and brought me back to an even more intensive concentration on the single case. In essence, I suddenly found myself face to face with the engineering problem of the animal trainer. When you have the responsibility of making absolutely sure that a given organism will engage in a given sort of behavior at a given time, you quickly grow impatient with theories of learning. Principles, hypotheses, theorems, satisfactory proof at the .05 level of significance that behavior at a choice point shows the effect of secondary reinforcement—nothing could be more irrelevant. No one goes to the circus to see the average dog jump through a hoop significantly oftener than untrained dogs raised under the same circumstances, or to see an elephant demonstrate a principle of behavior.

Perhaps I can illustrate this without giving aid and comfort to the enemy by describing a Russian device which the Germans found quite formidable. The Russians used dogs to blow up tanks. A dog was trained to hide behind a tree or wall in low brush or other cover. As a tank approached and passed, the dog ran swiftly alongside it, and a small magnetic mine attached to the dog's back was sufficient to cripple the tank or set it afire. The dog, of course, had to be replaced.

Now I ask you to consider some of the technical problems which the psychologist faces in preparing a dog for such an act of unintentional heroism. The dog must wait behind the tree for an indefinite length of time. Very well, it must therefore be intermittently reinforced for waiting. But what schedule will achieve the highest probability of waiting? If the reinforcement is to be food, what is the absolutely optimal schedule of deprivation consistent with the health of the dog? The dog must run to the tank—that can be arranged by reinforcing it with a practice tank—but it must start instantly if it is to overtake a swift tank, and how do you differentially reinforce short reaction times, especially in counteracting the reinforcement for sitting and waiting? The dog must react only to tanks, not to a refugee driving his oxcart along the road, but what are the defining properties of a tank so far as a dog is concerned?

I think it can be said that a functional analysis proved adequate in its technological application. Manipulation of environmental conditions alone made possible a wholly unexpected practical control. Behavior could be shaped up according to specifications and maintained indefinitely almost at will. One behavioral technologist who worked with me at the time (Keller Breland) is now specializing in the production of behavior as a salable commodity and has described this new profession in the *American Psychologist*.[3]

There are many useful applications within psychology itself. Ratliff and Blough have recently conditioned pigeons to serve as psychophysical observers. In their experiment a pigeon may adjust one of two spots of light until the two are equally bright or it may hold a spot of light at the absolute threshold during dark adaptation. The techniques which they have developed to induce pigeons to do this are only indirectly related to the point of their experiments and hence exemplify the application of a behavioral science.[4] The field in which a better technology of behavior is perhaps most urgently needed is education. I cannot describe here the applications which are now possible, but perhaps I can indicate my enthusiasm by hazarding the guess that educational techniques at all age levels are on the threshold of revolutionary changes (see page 171).

The effect of a behavioral technology on scientific practice is the issue here. Faced with practical problems in behavior, you necessarily emphasize the refinement of *experimental* variables. As a result, some of the standard procedures of statistics appear to be circumvented. Let me illustrate. Suppose that measurements have been made on two groups of subjects differing in some detail of experimental treatment. Means and standard deviations for the two groups are determined, and any difference due to the treatment is evaluated. If the difference is in the expected direction but is not statistically significant, the almost universal recommendation would be to study larger groups. But our experience with practical control suggests that we may reduce the troublesome variability by changing the conditions of the experiment. By discovering, elaborating, and fully exploiting every relevant variable, we may eliminate *in advance of measurement* the individual differences which obscure the difference under analysis. This will achieve the same result as increasing the size of groups, and it will almost certainly yield a bonus in the discovery of new variables which would not have been identified in the statistical treatment.

The same may be said of smooth curves. In our study of anxiety, Estes and I published several curves, the reasonable smoothness of which was obtained by averaging the performances of twelve rats for each curve. The individual curves published at that time show that the mean curves do not

[3] Breland, K., & Breland, Marion. A field of applied animal psychology. *Amer. Psychologist*, 1951, 6, 202–204.

[4] Ratliff, F., & Blough, D. S. Behavioral studies of visual processes in the pigeon. Report of Contract N50ri-07663, Psychological Laboratories, Harvard University, September, 1954.

faithfully represent the behavior of any one rat. They show a certain tendency toward a change in slope which supported the point we were making, and they may have appeared to warrant averaging for that reason.

But an alternative method would have been to explore the individual case until an equally smooth curve could be obtained. This would have meant not only rejecting the temptation to produce smoothness by averaging cases, but manipulating all relevant conditions as we later learned to manipulate them for practical purposes. The individual curves which we published at that time point to the need not for larger groups but for improvement in experimental technique. Here, for example, is a curve the smoothness of which is characteristic of current practice. Such curves were shown in the making in a demonstration which Ferster and I arranged at the Cleveland meeting of the American Psychological Association (Figure 13). Here, in a single organism, three different schedules of reinforcement

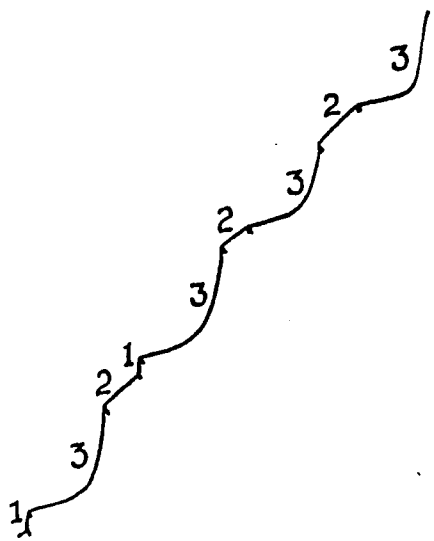

FIG. 13

are yielding corresponding performances with great uniformity under appropriate stimuli alternating at random. One does not reach this kind of order through the application of statistical methods.

In *The Behavior of Organisms* I was content to deal with the over-all slopes and curvature of cumulative curves and could make only a rough classification of the properties of behavior shown by the finer grain. The grain has now been improved. The resolving power of the microscope has been greatly increased, and we can see fundamental processes of behavior in sharper and sharper detail. In choosing rate of responding as a basic datum and in recording this conveniently in a cumulative curve, we make important temporal aspects of behavior *visible*. Once this has happened, our scientific practice is reduced to simple looking. A new world is opened to inspection. We use such curves as we use a microscope, X-ray camera, or telescope. This is well exemplified by recent extensions of the method. These are no longer part of my case history, but perhaps you will permit me to consult you about what some critics have described as a *folie à deux* or group neurosis.

An early application of the method to the behavior of avoidance and escape was made by Keller in studying the light aversion of the rat. This was brilliantly extended by Murray Sidman in his shock-avoidance experiments. It is no longer necessary to describe avoidance and escape by appeal to "principles," for we may *watch* the behavior develop when we have arranged the proper contingencies of reinforcement, as we later watch it change as these contingencies are changed.

Hunt and Brady have extended the use of a stable rate in the study of anxiety-producing stimuli and have shown that the depression in rate is eliminated by electroconvulsive shock and by other measures which are effective in reducing anxiety in human patients. O. R. Lindsley has found the same thing for dogs, using insulin-shock therapy and sedatives. Brady has refined the method by exploring the relevance of various schedules of reinforcement in tracing the return of the conditioned depression after treatment. In these experiments you *see* the effect of a treatment as directly as you see the constriction of a capillary under the microscope.

Early work with rats on caffeine and Benzedrine has been extended by Lindsley with dogs. A special technique for evaluating several effects of a drug in a single short experimental period yields a record of behavior which can be read as a specialist reads an electrocardiogram. Dr. Peter Dews of the Department of Pharmacology at the Harvard Medical School is investigating dose-response curves and the types and effects of various drugs, using pigeons as subjects. In the Psychological Laboratories at Harvard additional work on drugs is being carried out by Morse, Herrnstein, and Marshall, and the technique is being adopted by drug manufacturers. There could scarcely be a better demonstration of the experimental treatment of variability. In a *single* experimental session with a *single* organism one observes the onset, duration, and decline of the effects of a drug.

The direct observation of *defective* behavior is particularly important.

Clinical or experimental damage to an organism is characteristically unique. Hence the value of a method which permits the direct observation of the behavior of the individual. Lindsley has studied the effects of near-lethal irradiation, and the effects of prolonged anesthesia and anoxia are currently being examined by Thomas Lohr in co-operation with Dr. Henry Beecher of the Massachusetts General Hospital. The technique is being applied to neurological variables in the monkey by Dr. Karl Pribram at the Hartford Institute. The pattern of such research is simple: establish the behavior in which you are interested, submit the organism to a particular treatment, and then look again at the behavior. An excellent example of the use of experimental control in the study of *motivation* is some work on obesity by J. E. Anliker in collaboration with Dr. Jean Mayer of the Harvard School of Public Health, where abnormalities of ingestive behavior in several types of obese mice can be compared by direct inspection.

There is perhaps no field in which behavior is customarily described more indirectly than psychiatry. In an experiment at the Massachusetts State Hospital, O. R. Lindsley is carrying out an extensive program which might be characterized as a quantitative study of the temporal properties of psychotic behavior.[5] Here again it is a question of making certain characteristics of the behavior visible.

The extent to which we can eliminate sources of variability before measurement is shown by a result which has an unexpected significance for comparative psychology and the study of individual differences. Figure 14 shows tracings of three curves which report behavior in response to a multiple fixed-interval fixed-ratio schedule. The hatches mark reinforcements. Separating them in some cases are short, steep lines showing a high constant rate on a fixed-ratio schedule and, in others, somewhat longer "scallops" showing a smooth acceleration as the organism shifts from a very low rate just after reinforcement to a higher rate at the end of the fixed interval. The values of the intervals and ratios, the states of deprivation, and the exposures to the schedules were different in the three cases, but except for these details the curves are quite similar. Now, one of them was made by a *pigeon* in some experiments by Ferster and me, one was made by a *rat* in an experiment on anoxia by Lohr, and the third was made by a *monkey* in Karl Pribram's laboratory at the Hartford Institute. Pigeon, rat, monkey, which is which? It doesn't matter. Of course, these three species have behavioral repertoires which are as different as their

[5] Lindsley, O. R. Operant conditioning methods applied to research in chronic schizophrenia. *Psychiat. Res. Rep.*, 1956, 5, 118–139.

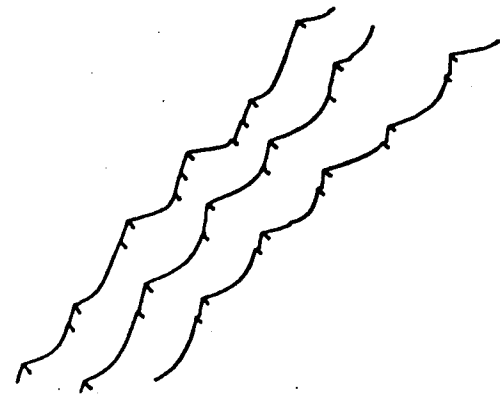

FIG. 14

anatomies. But once you have allowed for differences in the ways in which they make contact with the environment, and in the ways in which they act upon the environment, what remains of their behavior shows astonishingly similar properties. Mice, cats, dogs, and human children could have added other curves to this figure. And when organisms which differ as widely as this nevertheless show similar properties of behavior, differences between members of the same species may be viewed more hopefully. Difficult problems of idiosyncrasy or individuality will always arise as products of biological and cultural processes, but it is the very business of the experimental analysis of behavior to devise techniques which reduce their effects except when they are explicitly under investigation.

We are within reach of a science of the individual. This will be achieved, not by resorting to some special theory of knowledge in which intuition or understanding takes the place of observation and analysis, but through an increasing grasp of relevant conditions to produce order in the individual case.

A second consequence of an improved technology is the effect upon behavior theory. As I have pointed out elsewhere, it is the function of learning theory to create an imaginary world of law and order and thus to console us for the disorder we observe in behavior itself. Scores on a T maze or jumping stand hop about from trial to trial almost capriciously. Therefore we argue that if learning is, as we hope, a continuous and orderly process, it must be occurring in some other system of dimensions—perhaps in the nervous system, or in the mind, or in a conceptual model of behavior. Both the statistical treatment of group means and the averaging of curves encourage the belief that we are somehow going behind the individual

case to an otherwise inaccessible, but more fundamental, process. The whole tenor of our paper on anxiety, for example, was to imply that the change we observed was not necessarily a property of behavior, but of some theoretical state of the organism ("anxiety") which was merely *reflected* in a slight modification of performance.

When we have achieved a practical control over the organism, theories of behavior lose their point. In representing and managing relevant variables, a conceptual model is useless; we come to grips with behavior itself. When behavior shows order and consistency, we are much less likely to be concerned with physiological or mentalistic causes. A datum emerges which takes the place of theoretical fantasy. In the experimental analysis of behavior we address ourselves to a subject matter which is not only manifestly the behavior of an individual and hence accessible without the usual statistical aids but also "objective" and "actual" without recourse to deductive theorizing.

Statistical techniques serve a useful function, but they have acquired a purely honorific status which may be troublesome. Their presence or absence has become a shibboleth to be used in distinguishing between good and bad work. Because measures of behavior have been highly variable, we have come to trust only results obtained from large numbers of subjects. Because some workers have intentionally or unconsciously reported only selected favorable instances, we have come to put a high value on research which is planned in advance and reported in its entirety. Because measures have behaved capriciously, we have come to value skillful deductive theories which restore order. But although large groups, planned experiments, and valid theorizing are associated with significant scientific results, it does not follow that nothing can be achieved in their absence. Here are two brief examples of the choice before us.

How can we determine the course of dark adaptation in a pigeon? We move a pigeon from a bright light to a dark room. What happens? Presumably the bird is able to see fainter and fainter patches of light as the process of adaptation takes place, but how can we follow this process? One way would be to set up a discrimination apparatus in which choices would be made at specific intervals after the beginning of dark adaptation. The test patches of light could be varied over a wide range, and the percentages of correct choices at each value would enable us eventually to locate the threshold fairly accurately. But hundreds of observations would be needed to establish only a few points on the curve and to prove that these show an actual change in sensitivity. In the experiment by Blough already mentioned, the pigeon holds a spot of light close to the threshold throughout

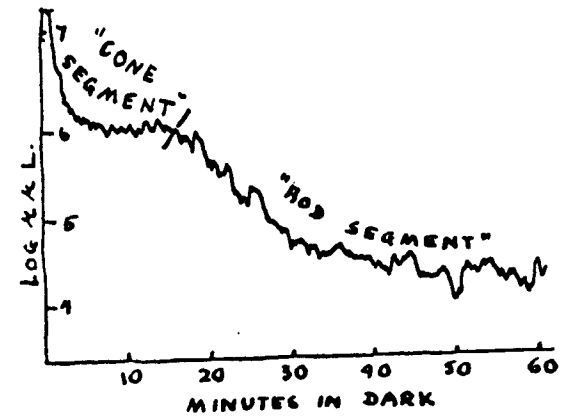

FIG. 15

the experimental period. A single curve, such as the one sketched in Figure 15, yields as much information as hundreds of readings, together with the means and standard deviations derived from them. The information is more accurate because it applies to a single organism in a single experimental session. Yet many psychologists who would accept the first as a finished experiment because of the tables of means and standard deviations would boggle at the second or call it a preliminary study. The direct evidence of one's senses in observing a process of behavior is not trusted.

As another example, consider the behavior of several types of obese mice. Do they all suffer from a single abnormality in their eating behavior or are there differences? One might attempt to answer this with some such measure of hunger as an obstruction apparatus. The numbers of crossings of a grid to get to food, counted after different periods of free access to food, would be the data. Large numbers of readings would be needed, and the resulting mean values would possibly not describe the behavior of any one mouse in any experimental period. A much better picture may be obtained with one mouse of each kind in single experimental sessions, as Anliker has shown.[6] In an experiment reported roughly in Figure 16, each mouse was reinforced with a small piece of food after completing a short "ratio" of responses. The hypothalamic-obese mouse shows an exaggerated but otherwise normal ingestion curve. The hereditary-obese mouse eats slowly but for an indefinite length of time and with little change in rate

[6] Anliker, J., and Mayer, J. Operant conditioning technique for studying feeding patterns in normal and obese mice. *J. Appl. Psychol.*, 1956, 8, 667–670.

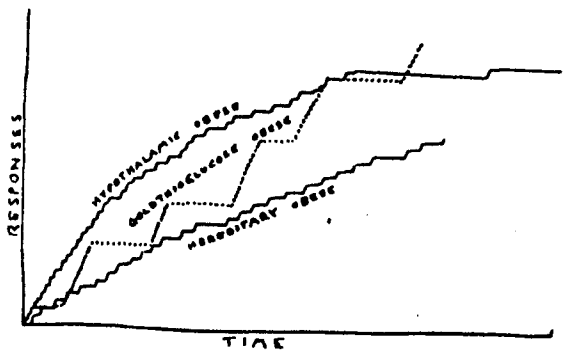

FIG. 16

The gold-poisoned obese mouse shows a sharp oscillation between periods of very rapid responding and no responding at all. These three individual curves contain more information than could probably ever be generated with measures requiring statistical treatment, yet they will be viewed with suspicion by many psychologists because they are single cases.

It is perhaps natural that psychologists should awaken only slowly to the possibility that behavioral processes may be directly observed, or that they should only gradually put the older statistical and theoretical techniques in their proper prospective. But it is time to insist that science does not progress by carefully designed steps called "experiments" each of which has a well-defined beginning and end. Science is a continuous and often a disorderly and accidental process. We shall not do the young psychologist any favor if we agree to reconstruct our practices to fit the pattern demanded by current scientific methodology. What the statistician means by the design of experiments is design which yields the kind of data to which *his* techniques are applicable. He does not mean the behavior of the scientist in his laboratory devising research for his own immediate and possibly inscrutable purposes.

The organism whose behavior is most extensively modified and most completely controlled in research of the sort I have described is the experimenter himself. The point was well made by a cartoonist in the Columbia *Jester* (Figure 17). The caption read: "Boy, have I got this guy conditioned! Every time I press the bar down he drops in a piece of food." The subjects we study reinforce us much more effectively than we reinforce them. I have been telling you simply how I have been conditioned to behave. And of course it is a mistake to argue too much from

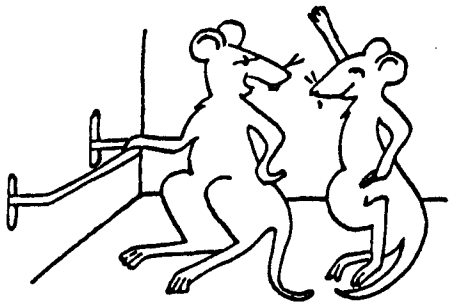

FIG. 17

one case history. My behavior would not have been shaped as it wa
it not for personal characteristics which all psychologists fortunat
not share. Freud has had something to say about the motivation of
tists and has given us some insight into the type of person who a
the fullest satisfaction from precise experimental design and the int
of deductive systems. Such a person tends to be more concerned w
success as a scientist than with his subject matter, as is shown by t
that he often assumes the role of a roving ambassador. If this seems
let me hasten to characterize my own motivation in equally unfla
terms. Several years ago I spent a pleasant summer writing a novel
Walden Two. One of the characters, Frazier, said many things which
not yet ready to say myself. Among them was this:

> I have only one important characteristic, Burris: I'm stubborn. I've h
> one idea in my life—a true *idée fixe* . . . to put it as bluntly as possi
> idea of having my own way. "Control" expresses it, I think. The co
> human behavior, Burris. In my early experimental days it was a frenzied,
> desire to dominate. I remember the rage I used to feel when a predictio
> awry. I could have shouted at the subjects of my experiments, "Behave, da
> behave as you ought!" Eventually I realized that the subjects were alway
> They always behaved as they ought. It was I who was wrong. I had mad
> prediction.

(In fairness to Frazier and the rest of myself, I want to add hi
remark: "And what a strange discovery for a would-be tyrant, t
only effective technique of control is unselfish." Frazier means, of
positive reinforcement.)

We have no more reason to say that all psychologists should beh
I have behaved than that they should all behave like R. A. Fishe
scientist, like any organism, is the product of a unique history. Th

tices which he finds most appropriate will depend in part upon this history. Fortunately, personal idiosyncrasies usually leave a negligible mark on science as public property. They are important only when we are concerned with the encouragement of scientists and the prosecution of research. When we have at last an adequate empirical account of the behavior of Man Thinking, we shall understand all this. Until then, it may be best not to try to fit all scientists into any single mold.