

*THORNDIKE'S PUZZLE BOXES AND THE ORIGINS OF
THE EXPERIMENTAL ANALYSIS OF BEHAVIOR*

PAUL CHANCE

LAUREL, DELAWARE

The year of Thorndike's dissertation on animal intelligence, 1898, may mark the beginning of the field that eventually became known as the experimental analysis of behavior. The dissertation began a major shift in thinking about animal and human learning, provided important methodological innovations, and carried the seeds of later research and theory, particularly by B. F. Skinner. Although Thorndike was an associationist in 1898, the dissertation began the systematic search for fundamental behavioral processes, and laid the foundation for an empirical science of behavior.

Key words: E. L. Thorndike, puzzle boxes, history, animal intelligence, learning, problem solving, associations

When the crude beginnings of this research have been improved and replaced by more ingenious and adroit experimenters, the results ought to be very valuable. (Thorndike, 1898, p. 30f)

Attaching dates to the origin of a field of study is bound to start an argument, but a case can be made for placing the beginning of the experimental analysis of behavior in 1898. That was the year E. L. Thorndike published his famous dissertation, *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*.

It was a banner year for Thorndike. In January he spoke before the New York Academy of Sciences. In June, a summary of his experiments appeared in *Science*, and *Psychological Review Monograph* published his dissertation in its entirety. Finally in December he described his work in an invited address at the seventh annual meeting of the American Psychological Association.

Although Thorndike was only 23 years old as the year opened, his research caused quite a stir. It provoked immediate criticism from members of the old guard. Wesley Mills (1899), an eminent animal psychologist at McGill University, suggested that Thorndike's findings were the result of the contrived nature of his experiments. C. Lloyd Morgan (1900), the leading figure in comparative psy-

chology at the time, also complained about the artificiality of the experiments, and added that Thorndike's animals were less his subjects than his victims. Yet the appearance of the dissertation was widely recognized, as Thorndike's biographer Geraldine Joncich (1968) would later write, as "an undeniably important event" (p. 148).

Joncich may be guilty of understatement: The dissertation helped to establish comparative psychology as an experimental science; began a major shift in thinking about both animal and human learning; provided important methodological innovations in behavioral research; and carried the seeds of later research and theory by many others, most notably B. F. Skinner. I believe it also provides the starting point for the field that would become known as the experimental analysis of behavior. On this 100th anniversary of the dissertation, it is appropriate that we examine the experiments and theory it describes.

Experiments

Thorndike's general experimental method was "to put animals when hungry in enclosures from which they could escape by some simple act" (p. 6; all page references are to Thorndike, 1898). Thorndike typically placed food in view outside of the enclosure and observed how the animal behaved. He also recorded the time that elapsed before the animal escaped. He repeated this procedure over and over again and noted the change in the animal's behavior and in escape times.

The most famous of the dissertation experiments are those with cats in "puzzle box-

The author is editor of "Beyond the Data," a column of commentary on behavioral issues on the Web site of the Cambridge Center for Behavioral Studies (behavior.org). He thanks Diane Chance for help in preparing the figures for this article.

Address correspondence to the author at P.O. Box 248, Laurel, Delaware 19956 (E-mail: pbchance@aol.com).

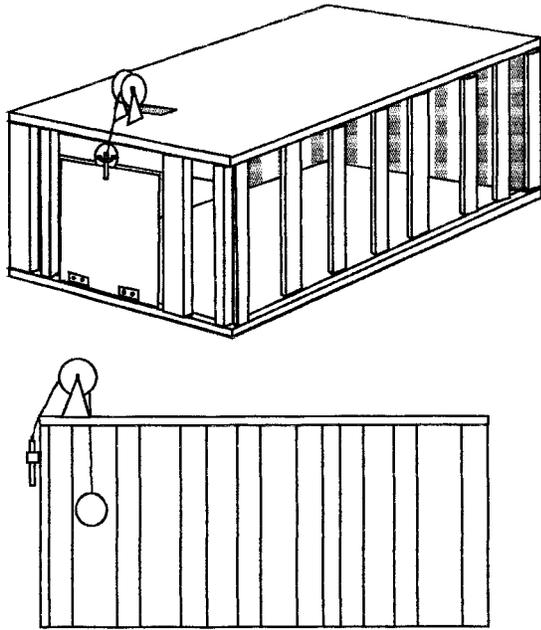


Fig. 1. Box A. Pulling on the loop (lower figure) released a bolt and the door fell open.

es.” There were 15 of these boxes, and they were constructed mainly of wooden slats and hardware cloth. Each box contained a door that the cat could open by manipulating some device. A cat could open the door to Box A, for example, by pulling on a wire loop suspended six inches above the box floor (see Figure 1). The door to Box H could be opened by pushing it aside (see Figure 2). Cats opened the door to Box I by pressing a lever (see Figure 3). (The cat that

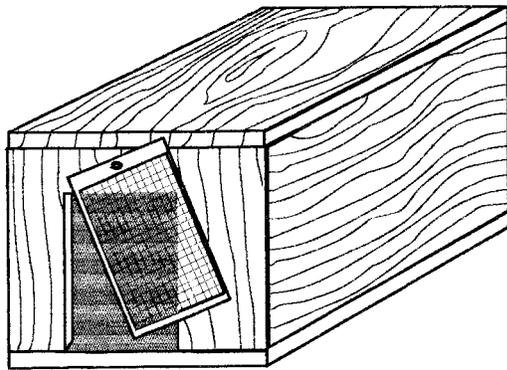


Fig. 2. Box H. The door pivoted on a screw and could be pushed aside to the right or left.

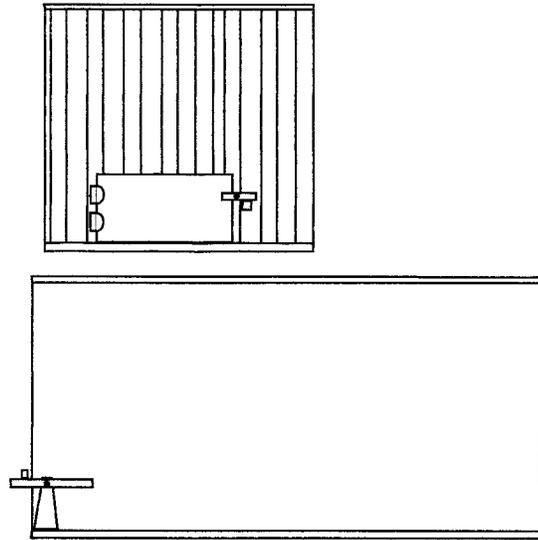


Fig. 3. Box I. Depressing a lever (lower figure) raised a bar that pivoted, allowing the door to swing open. (In the lower figure, the solid side of the box has been removed to reveal the lever.)

first escaped from Box I may well deserve a place in history for being the first in a long line of lever-pressing animals.) Box K, the only box depicted graphically in the dissertation and the one that is most familiar to behaviorists, required the performance of three distinct responses: The cat had to depress a treadle, pull on a string, and push a bar up or down before the door would finally fall open (see Figure 4).

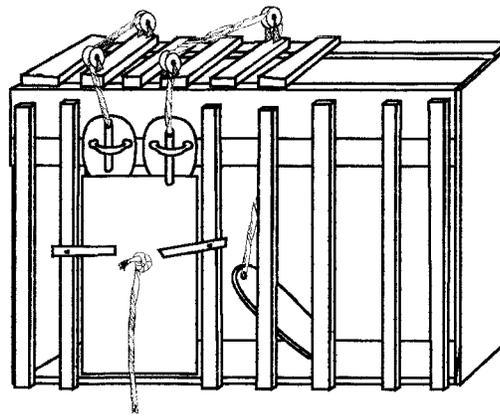


Fig. 4. Box K. The door is held in place by a weight suspended by a string. To open the door, a cat had to depress a treadle, pull on a string, and push a bar up or down. (After Thorndike, 1898, Figure 1, p. 8.)

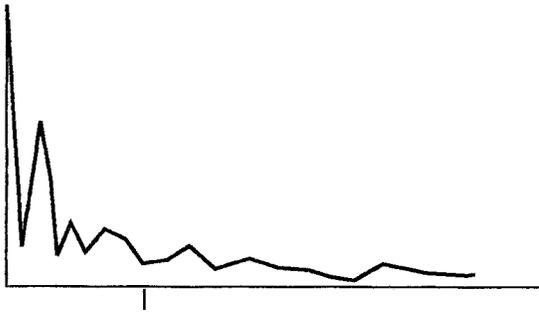


Fig. 5. Performance of Cat 12 in Box A. Box A required pulling on a loop. Trials are depicted on the abscissa, escape times on the ordinate. In this instance, escape times varied from a high of 160 s to a low of 6 s over 24 trials. The short vertical line on the abscissa indicates an interruption in training of about 24 hr. (From Thorndike, 1898, Figure 2, p. 18.)

Thorndike notes that a cat placed in such a box typically

tries to squeeze through any opening; it claws and bites at the bars or wire; it thrusts its paws out through any opening and claws at everything it reaches; it continues its efforts when it strikes anything loose and shaky; it may claw at things within the box. It does not pay very much attention to the food outside, but seems simply to strive instinctively to escape from confinement. The vigor with which it struggles

is extraordinary. For eight or ten minutes it will claw and bite and squeeze incessantly. (p. 13)

Thorndike found that when he put a cat into a given box again and again, the whole demeanor of the animal changed. At first the cat's behavior appeared to be almost random, one might even say chaotic. Gradually, however, it became more orderly, more deliberate, more efficient. "The cat that is clawing all over the box in her impulsive struggle will probably claw the string or loop or button so as to open the door. And gradually . . . after many trials, the cat will, when put in the box, immediately claw the button or loop in a definite way" (p. 13).

In addition to noting gross changes in behavior, Thorndike recorded the time required for the cat to escape and plotted these data on a graph to yield a "time-curve." The time-curve for Cat 12 in Box A, which required pulling on a wire loop, shows a rapid and fairly steady decline in escape time and is typical of the performance of cats in boxes that require a single response (see Figure 5). The time-curve for Cat 4 in Box K, which required three distinct responses, shows slower and more erratic progress (see Figure 6).

Thorndike was interested not only in the

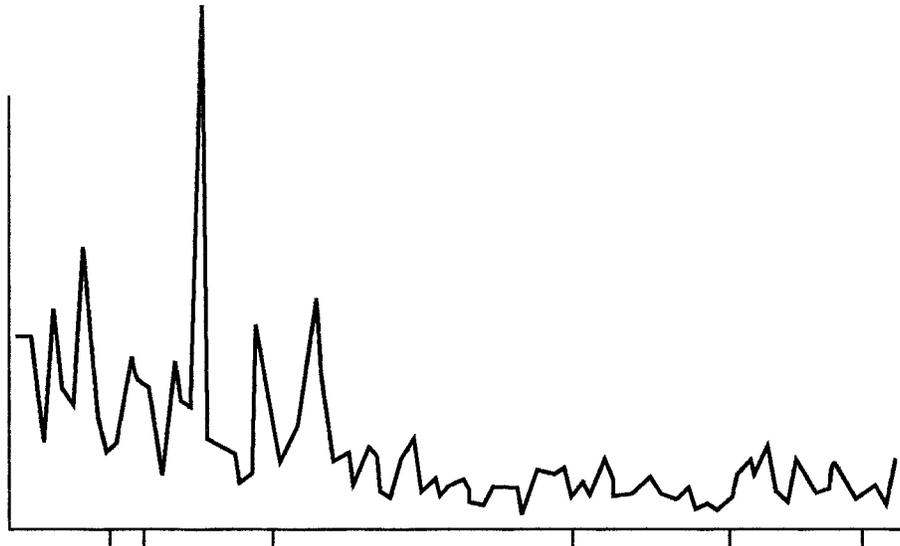


Fig. 6. Performance of Cat 4 in Box K. Box K required three distinct responses. The figure shows escape times on approximately 117 trials over a 7-day period. Progress was slow and erratic. (After Thorndike, 1898, Figure 10, p. 26.)

change in escape times, but in the *rate* of the change: He wrote that "we may take the general slope of the curve as representing very fairly the progress of the association" (p. 16). It was the slope of the curve, in other words, that showed the rate of learning. This interest in curve slope anticipates Skinner's (1938) use of the slope of a cumulative record.

Thorndike also did puzzle box experiments with 3 small dogs of unspecified breed. The nine puzzle boxes for dogs resembled those used for cats. The experimental procedure was also similar except that the period of food deprivation for dogs was shorter. This was because of the "practical necessity that the dogs should be kept from howling in the evening" (p. 32). The dogs showed the same sort of changes in gross behavior noticed in cats, with effective behavior persisting and becoming smoother and more efficient, and ineffective behavior dropping out. The time-curves for dogs also resembled those for cats.

Thorndike did experiments with chicks that showed the same patterns seen in dogs and cats. In some experiments a chick could escape confinement by stepping on a platform, pulling on a string, pecking a door, or pecking a tack. In other experiments the birds escaped from pens, some of which were constructed of books stood on end. To escape from one of the more complicated pens, a bird had to climb a spiral staircase, go through a hole in a wall, walk across a horizontal ladder, and jump off a ledge to reach food and the company of other chicks. Of course, today such a sequence of acts would be called a response chain.

Thorndike found that a confined bird typically "runs back and forth, peeping loudly, trying to squeeze through any openings there may be, jumping to get over the wall, and pecking at the bars or screen" (p. 36). Eventually the ineffective behaviors died out and the bird performed the act required for escape as soon as it was placed in the enclosure.

In most of the dissertation experiments, the animals escaped confinement by manipulating a device that caused a door to fall open. In some experiments, however, escape required Thorndike's intervention. For example, in one experiment Thorndike removed a chick from a box whenever it

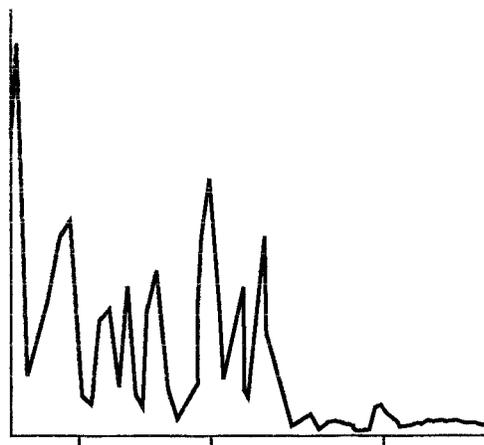


Fig. 7. Cat 5 in Box Z. The cat escaped from the box when it licked itself. The graph depicts about 64 trials over 4 days. The time-curve shows greater variability in performance and slower learning than when escape resulted from manipulation of some part of a box (cf. Figure 5). (After Thorndike, 1898, Figure 6, p. 22.)

preened its feathers. In other experiments, Thorndike opened the door of a box when a cat licked or scratched itself. These efforts were generally successful. The chick that escaped a box by preening, for example, eventually "would whirl his head round and poke it into the feathers as soon as dropped in the box" (p. 27). However, these experiments generally produced greater variability in performance and shallower time-curves (slower rates of learning) than when the means of escape involved manipulating some part of the enclosure (see Figure 7).

Thorndike conducted some experiments on generalization, although he did not use that term. After an animal had learned to escape from one box, Thorndike put it into a different kind of box. He found that the animal took what it learned, as it were, to the next problem. "A cat that has learned to escape from A by clawing," he wrote, "has when put into C or G a greater tendency to claw at things than it instinctively had at the start, and a less tendency to squeeze through holes" (p. 14).

Other experiments involved discrimination. Thorndike noticed that some cats climbed the wire netting in their home cages when he was about to feed them. Could he get control over this behavior by systematically manipulating the environment? Thorn-

dike tested this idea by saying, "I must feed those cats!" just before feeding them. At other times he announced, "I will not feed them," and then did not provide food. Thorndike recorded whether the cat climbed up to the wire netting of its pen after each of the statements. He recorded error data in two frequency graphs, one showing the failure to climb to the netting at the first signal, the other showing climbing the netting at the second signal. The animal learned to respond appropriately to what would now be called the S^D ("I must feed those cats!"; i.e., a stimulus correlated with obtaining food by climbing) in 60 trials, but learning *not* to approach at the S^A ("I will not feed them"; i.e., a stimulus correlated with no food) took much longer—380 trials.

Thorndike also did a number of experiments on observational learning. His basic procedure was to allow a chick, cat, or dog to observe as another of its species escaped from an enclosure, and record the observer's tendency to imitate the model's successful action. For example, a box might have two compartments separated by a wire screen. A cat could watch a model escape from confinement and reach food by pulling on a string. The observer could then escape confinement by pulling a string, but typically it did not. Instead, it went through the same gradual learning process seen in cats that had not had the benefit of a model. In the end, Thorndike concluded that, at least for animals below the level of the primates, "we should give up imitation as an *a priori* explanation of any novel intelligent performance" (p. 62). Later experiments would prove him wrong on this point (e.g., Herbert & Harsh, 1944; Miller & Dollard, 1941).

Theory

Although the dissertation's empirical contributions are of major importance in the history of behavior analysis, the theoretical work, outmoded though it is today, may have been equally important to the development of the field. In Thorndike's day, most people, including scientists devoted to the study of comparative psychology, believed that higher animals such as birds, dogs, and cats learned through "the association of ideas." This meant that they understood, in some primitive way, the logical relationships among

events and used those ideas to reason their way through problems. In accounting for a cat opening a door by manipulating a latch, for example, George Romanes (1882) wrote, "First the animal must have observed that the door is opened by the hand grasping the handle and moving the latch. Next she must reason, by 'the logic of feelings'—'If a hand can do it, why not a paw?' Then strongly moved by this idea she makes the first trial" (quoted in Thorndike, p. 41).

The support offered for this theory, even in scientific circles, consisted largely of anecdotal evidence and casual observation. Thorndike argued that this led to a bias in favor of evidence supporting animal reasoning. "Thousands of cats on thousands of occasions sit helplessly yowling," Thorndike wrote, "and no one takes thought of it or writes to his friend, the professor; but let one cat claw at the knob of a door supposedly as a signal to be let out, and straightway this cat becomes the representative of the cat-mind in all the books. . . . In short, the anecdotes give really the . . . *super-normal* psychology of animals" (p. 4f). There was, Thorndike said, no solid evidence that animals grasped ideas or learned through reasoning, and he set forth his arguments against the theory: First, the behavior of animals in an enclosure is impulsive and apparently random, not systematic and logical. The behavior of a cat in a puzzle box is initially "just a mad scramble to get out" (p. 43); the animal shows no sign of contemplation or thoughtfulness.

Second, the change in an animal's behavior is gradual, not abrupt. A cat becomes more and more likely to step on a treadle when put into the box, and less and less likely to scratch at the slats. If the animal understood the relationship between the pressing of the treadle and the opening of the door, the time-curve would show a sudden fall in escape times. The dozens of time-curves Thorndike produced "show no such phenomenon. . . . The gradual slope of the time-curve, then, shows the absence of reasoning" (p. 45).

Third, the animals show no sign of understanding the relationship between action and consequence even *after* they have learned to escape from a box. Thorndike found, for example, that a cat that had learned to escape a box by pulling on a loop would paw at the

air where the loop had been, even when the loop was no longer there.

Fourth, animals learned only if they performed the necessary act themselves. As noted earlier, Thorndike saw no evidence that animals benefited from observing the performance of a model. He also found that putting the animal's limb through the required motion did not produce learning.

For example, a cat would be put in B . . . and left two minutes. I would then put my hand in through the top of the box, take the cat's paw and with it pull down the loop. The cat would then go out and eat the fish. This would be done over and over again, and after every ten or fifteen such trials the cat would be left in alone. . . . The results . . . show that no animal who fails to perform an act in the course of his own impulsive activity will learn it by being put through it. (p. 68)

Thus, the actual performance of an act, not merely the *idea* of an act, was essential.

Having thus demolished, so he thought, the dominant theory of animal learning, Thorndike then offered an alternative. He proposed that animal learning has nothing to do with reasoning or the association of ideas; it occurs as a result of "trial and accidental success," a phrase that gave a new emphasis to the role of actions and their consequences. (Although Thorndike is typically said to have studied "trial-and-error" learning, he preferred the phrase "trial and success.")

Thorndike theorized that a cat is innately equipped with various "action impulses" when placed in a box. It has, for example, the impulse to bite the bars and to scratch at objects within the box. If pulling a loop results in escaping from a box, the impulse to pull on loops is strengthened; if the impulse to scratch at bars does not result in escape, that impulse is weakened. There is no association of ideas, said Thorndike, only the association of sensations, such as the sight of a loop, and the impulse to perform certain acts, such as clawing at things. Thorndike suggested a mechanical "stamping in" of these associations.

As this summary of the theory implies, Thorndike distinguished between impulse and act. An impulse was the *feeling* of performing an act, as distinct from the act itself. Although Thorndike would later speak of responses rather than impulses, in the disser-

tation it was the *feeling* of acting, and not the act itself, that concerned him. He wrote that the important thing was "the feeling of the doing"; the act itself was "a secondary affair" (p. 15).

The resulting view of learning requires no ideas in the head of the animal:

The cat that is clawing all over the box in her impulsive struggle will probably claw the string or loop or button so as to open the door. And gradually all the other non-successful impulses will be stamped out and the particular impulse leading to the successful act will be stamped in by the resulting pleasure, until, after many trials, the cat will, when put in the box, immediately claw the button or loop in a definite way. (p. 13)

It may surprise some readers to learn that this statement is as close as Thorndike gets to a formal expression of the law of effect in the dissertation. In fact, the phrase "law of effect" never appears in this work.

Thorndike anticipated that his theory would be challenged on the grounds that some instances of learning are too remarkable to be the result of an essentially random process. The behavior of a cat that opens a latched gate, for example, seems too complicated to be the result of accidental success. "The whole substance of the argument vanishes," Thorndike wrote, "if, as a matter of fact, animals do learn those things by accident. *They certainly do*" (p. 40). If animals in carefully controlled experiments can learn, through trial and accidental success, to perform an act, then there is no reason to suppose they acquire the act in any other way in more natural settings.

Assessing the Dissertation

Anyone who is interested in the study of behavior as a natural science must recognize Thorndike as a seminal figure: He replaced isolated anecdotal reports with repeated experimental observation; introspection with systematic description; and subjective impressions with numerical data. He was probably the first to study the process of learning in a truly systematic way, and was one of the first to capture learning in the slope of a line graph. At a time when others speculated about animal cogitations, Thorndike showed that a particular cat in a particular box took

a certain number of seconds to perform the precise act required for escape.

Thorndike began the search for fundamental behavioral processes, and noticed that those processes seemed remarkably similar in different species. The chicks learned more slowly than the dogs and cats, but “the biggest part of the difference . . . is not referable so much to any difference in intelligence as to a difference in their bodily organs and instinctive impulses” (p. 38).

Thorndike recognized not only what would today be called reinforcement, but extinction, generalization, discrimination, and response chaining. He realized that effective responses are “selected by success” (p. 14), and understood that three elements are essential to the description of behavior: the situation, the act, and the consequence. He introduced the lever press and the disk peck as dependent variables (although the disk was a tack suspended by a string), and emphasized individual rather than group data.

Thorndike rejected Cartesian dualism, and issued this challenge: “To the interactionists I would say: ‘Do not any more repeat in tiresome fashion that consciousness *does* alter movement, but get to work and show when, where, in what forms and to what degrees it does so’ ” (p. 104). This challenge is still being issued to dualists today.

In all of these contributions, and others, we can see the roots of *The Behavior of Organisms* (Skinner, 1938) and of much of the work that followed it. Yet it would be wrong to leave the impression that Thorndike was a Skinnerian 6 years before Skinner was born. He was not. He was as much an associationist as Romanes, except that in his view it was not ideas that were associated, but sensations and impulses. And despite his criticism of the subjectivism of others, he was himself, at least in 1898, a mentalist concerned with the “inner nature” of the animal’s associations; the facts “as observed from the outside” were of limited interest to him (p. 38).

He was also guilty on occasion of the same sort of careless reasoning that he criticized in other animal researchers. Thorndike tells us, for example, that he repeated his experiments with a given animal “until the animal had formed a perfect association between the sense-impression of the interior of that box and the impulse leading to the successful

movement. When the association was thus perfect, the time taken to escape was, of course, practically constant and very short” (p. 6). This is clearly circular: There was no evidence of any association between sense-impressions and impulses except the “practically constant and very short” escape times from which it was inferred.

We can also see, with the telescopic lens provided by a hundred years of research, that Thorndike failed to appreciate the significance of some of his own findings. He was puzzled, for example, by the fact that when escape was the result of licking or scratching, there was “a noticeable tendency . . . to diminish the act until it becomes a mere vestige of a lick or scratch. . . . The licking degenerates into a mere quick turn of the head with one or two motions up and down with tongue extended. Instead of a hearty scratch, the cat waves its paw up and down rapidly for an instant” (p. 28). Later research by others would suggest that the behavior “degenerated” only because the requirement for reinforcement degenerated.

Similarly, although Thorndike recognized the selective effects of consequences, and knew that animals could perform complex response chains, he seems to have had no idea that an experimenter might shape behavior or build chains by carefully arranging consequences.

Finally, if we judge Thorndike’s dissertation against current standards, we must admit that it would be unlikely to win a doctorate in behavior analysis today, or even in general experimental psychology. His experimental methods, his data analysis, his theoretical arguments—all major contributions of his day—would be unacceptable.

Were Thorndike alive, I do not believe he would take offense at this assessment. Indeed, he anticipated it. He wrote in the dissertation, “When the crude beginnings of this research have been improved and replaced by more ingenious and adroit experimenters, the results ought to be very valuable” (p. 30f). Whether there have been any more ingenious and adroit experimenters than Thorndike is a matter for debate, but it is clear that we should not hold him to a standard arrived at after a hundred years of the experimental analysis of behavior. Instead, we should recognize that it was his great contri-

butions that laid the foundation for that new science.

REFERENCES

- Herbert, M. J., & Harsh, C. M. (1944). Observational learning by cats. *Journal of Comparative Psychology*, *37*, 81–95.
- Jonçich, G. (1968). *The sane positivist: A biography of Edward L. Thorndike*. Middletown, CT: Wesleyan University Press.
- Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. New Haven, CT: Yale University Press.
- Mills, W. (1899). The nature of animal intelligence and the methods of investigating it. *Psychological Review*, *5*, 262–274.
- Morgan, C. L. (1900). *Animal behavior*. London: Edward Arnold.
- Romanes, G. J. (1882). *Animal intelligence*. New York: Appleton.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.
- Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplement* *2* (4, Whole No. 8).

