
J. B. Watson's Approach to Learning:

Why Pavlov? Why Not Thorndike?

Jacob L. Gewirtz
Florida International University

Throughout his writing career, John B. Watson showed a continuing partiality for Pavlov's stimulus-substitution and avoidance/escape paradigm. For his diverse conceptual and practical goals, including his infant training conception ("...give me a dozen healthy infants..."), Thorndike's response learning-by-consequences paradigm seemed to fit Watson's variegated needs far better. In this analysis, the attempt is made to propose solutions to the mystery of why Thorndike's learning model was ignored so completely, while Pavlov's model was favored so consistently, by Watson.

In his early writings, John Broadus Watson (1914, 1919) gave to *instincts* a key role. As in the writings of his contemporaries (e.g., McDougall, 1908), instincts accounted for much behavior termed unlearned. In his subsequent writing, Watson (1919) conceived that instincts/reflexes are present early in life, but that they are soon displaced by learned habits. Subsequently, Watson (1924, 1925) concluded that what theretofore had been called instincts were the products of learning. Watson wrote of this conclusion at around the time that influential authors, like the sociologist L. L. Bernard (1924), concluded that what were theretofore termed instinctive behaviors were the direct outcomes of learning processes.

When dealing with, or explaining, diverse learning phenomena, Watson followed existing conventions. For learning, he emphasized particularly the functionalist principles of contiguity, frequency, and recency. Perhaps because it emphasized the contiguity and frequency principles to which he was partial in his approach to learning, and was clearly scientific, Watson had seemed to show some continuing partiality for the *learning* model that he often emphasized during the same years as he presented an early version of his dramatically-new orientation to the psychology of the period, that he termed *behaviorism*: it was the Pavlovian (1928) reflex-based paradigm termed *respondent (classical)* conditioning. Indeed, following the Yerkes and Morgulis (1909) *Psychological Bulletin* article that introduced Pavlovian learning into the American psychological literature, Watson's 1915 presidential address to the American Psychological Association (published in the 1916 *Psychological Review*) was influential in establishing the Pavlovian learning paradigm in American psychology. Moreover, Watson continued his emphasis on the Pavlovian conditioning paradigm to the end: In the last versions of his book *Behaviorism* (1924, 1925, 1930), Watson devotes Chapter II, How to Study Human Behavior, almost entirely to the utility

of Pavlov's stimulus-substitution conditioning paradigm in the study of human behavior. He did not emphasize the learning paradigms of functionalism to which he was earlier exposed as a graduate student at Chicago.

Given that Watson's approach increasingly involved environmental causation, in contrast to an early emphasis on instinctive behavior (1914), it is not surprising that he singled-out *learning qua training* as the environmental-behavior interaction process underlying the critical behavior changes required by, and validating, the extreme environmentalism as a central tenet of his behaviorism at its peak (Watson & Watson, 1928). What is surprising in retrospect is that Watson often appeared to have based his central behavior-change engine upon such a limited learning model as Pavlov's, that is superimposed on a short litany of sometimes-atypical reflexes, and that emphasizes at its foundation only stimulus substitution (the CS for the US). While Pavlov (1928) had credited Thorndike with being the first to show a bold accurate approach to the immense task of objective research on animal learning, Watson's emphasis nearly-always appeared to be on the Pavlovian learning paradigm that necessarily ignores the important and wide-ranging role of behavior change, in which it is the consequences associated with the training instances that comprise the bulk of the learning resulting from the systematic interaction of the individual's behavior with features of the environment.

One well-known example from his later writings, but one that only very-generally (and perhaps incorrectly) is taken to imply an instrumental-learning model comes from one of Watson's (1924 - 1925) later writings in which his, by then extreme, environmentalism peaked:

...give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select: doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and the race of his ancestors. I am going beyond my facts and I admit it, but so have the advocates of the contrary...Please note that when this experiment is made I am to be allowed to specify the way the children are to be brought up and the type of world they have to live in. (p. 10)

In the frame outlined, the focus of the present analysis is to speculate about why Watson consistently placed so much of his emphasis upon Pavlovian stimulus-substitution, avoidance, and

escape learning when, in hindsight, the Thorndikian-learning model based on the consequences of behavior has seemed more appropriate for much of Watson's later emphasis on training, as exemplified in the paragraph quoted above. The Thorndikian-learning conception had been available in the literature for two decades when Watson had begun his long-term emphasis on the generic use of the Pavlovian-conditioning model.

Thorndike's Learning Model: The Importance of the Consequence Produced by Behavior

In 1898, in the *Psychological Review Monograph Series*, E. L. Thorndike published a report of his doctoral work, which he had carried out at Columbia University under the mentorship of J. M. Cattell, after having taken his Masters degree at Harvard under William James. The monograph was titled *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*. In that monograph, Thorndike was the first experimental researcher to report how nonreflexive, nonspecies-specific, trial-controlled animal-behavior latency could be modified by its consequences. Small hungry domestic animals (cats for the most part) were individually placed in barred chambers (puzzle boxes). If the animal exhibited the behavior leading to escape (pulling a string, stepping on a platform, and/or reaching through the bars and turning a latch on the door front), the puzzle box's door would open and the animal could exit and have access to and/or eat the food placed outside the door as well as to what the freedom implied. Thorndike's conclusion was that, in the course of its exploration of the puzzle box, the animal exhibited the first instance of an appropriate behavior unit by chance and that, across trials, escape latencies would decrease systematically.

Thorndike attributed this performance-improvement pattern to the strengthening of an S-R connection between events in the box: i.e., the stimulus, and the gross behavior that opened the door permitting exit; i.e., the response. In a later work titled *Animal Intelligence*, Thorndike (1911) developed further his 1898 notions and emphasized, in detail, his conception of the *Law of Effect*. Thus, the systematic strengthening of the S-R resulted from the *Law of Effect*, which held that the greater the contingent (satisfaction) the greater the strengthening, and the greater the contingent (dissatisfaction, discomfort), the greater the weakening, of the bond. Satisfaction is denoted by a state the organism does nothing to avoid and often to attain or preserve (e.g., escape, food for the hungry animal), while dissatisfaction / annoyance is denoted by a state that the organism typically avoids or escapes (e.g., electric shock). From his early work, Thorndike concluded that learning was gradual, did not involve consciousness or mental processes like thought, and involved the same processes across mammalian species. Early on, Thorndike's conception of the basis of learning emphasized both exercise (frequency of learning occasions such as pairings) and effect (positive or negative consequences); in the late 1920s, Thorndike came to emphasize primarily consequences. (Within Skinner's conception of the *Empirical Law of Effect*, reinforcers and punishers came to displace satisfaction and dissatisfaction as the engine of change in responses, i.e., learning,

resulting from consequences.)

Thorndike's functionalism, his controlled research on animal behavior, his emphasis on the homogeneity of learning processes across mammalian species (including the human), and the absence in Thorndike's research of an emphasis on consciousness, are all factors that should have made much of Thorndike's work, and the results of that work, of interest and of importance to Watson. After all, Watson did emerge from a context of functionalism at Chicago, and Thorndike was a functionalist who worked, not on introspectively-derived human consciousness, but on animal behavior that denoted learning. Furthermore, Thorndike had written a positive and sympathetic review of Watson's *Manifesto of Behaviorism* (1913). In addition, Dr. Peter Harzem (personal communication, May 27, 1999), who is preparing a biography of J. B. Watson, has informed me that, in the Watson Archive at the Johns Hopkins University, he had located a friendly letter that Watson wrote in 1919, inviting Thorndike to collaborate in an important research project. In this context, it appears unlikely that Watson would have ignored Thorndike's learning-by-consequences paradigm out of pettiness or pique and our search for bases of Watson ignoring Thorndike's seemingly more-appropriate learning model for Pavlov's must continue.

Some Possible Bases for Watson's Ignoring the Thorndikian-Learning Paradigm

There are a number of plausible speculative *negative* reasons why, singly or in some combination, Thorndike's learning paradigm might not have appealed to Watson. *First*, there is a hedonistic mentalism implicit in Thorndike's Law of Effect. *Second*, the "satisfying" and "dissatisfying" states of affairs of Thorndike's Law of Effect were unparsimonious (especially so for a contiguity theorist like Watson), if not also teleological in the special sense of the future influence of current behavior consequences. *Finally*, for Watson, the proper level of analysis for psychological science was "molecular" rather than "molar", as the terms, first advanced by Broad, were applied by E. C. Tolman (1928) in connection with his molar "purposive behavior." The term molecular implied smaller units of analysis, with behaviors being parts of reflexes or small movements like those of muscles or glands, and being explained by under-the-skin factors. The term molar implied larger units of analysis and larger behavior units like actions that were organized in functional contexts with environmental objects and events.

For Watson, molecular (rather than molar) units were requisites for proper scientific description and explanation. With Watson's strong molecular bias, the clearly-molar units of the Thorndike paradigm could have been seen by Watson as unscientific. For the very same reason, the proper molecular units of Pavlov's paradigm, the absence in Pavlov's work of hedonism, mentalism, or teleology, and the presence there of parsimony and an emphasis on contiguity and frequency, all would have recommended Pavlov's learning paradigm to Watson as proper science. These criteria may have trumped the likely utility of Thorndike's response learning-by-consequences paradigm for Watson's radical purpose of ignoring genotype while manipu-

lating environmental factors through training-qua-learning processes. In this frame, it is interesting to note that the post-Watson waves of behavioristic theories, primarily of learning processes, have emphasized molar, rather than molecular, units of analysis. These include the approaches of Tolman, Hull, and Skinner.

REFERENCES

Bernard, L. L. (1924). *Instinct: A study in social psychology*. New York: Henry Holt.
McDougall, W. (1908). *Introduction to social psychology*. London: Methuen.
Pavlov, I. P. (1928). *Lectures on conditioned reflexes*. New York: Liveright.
Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review*, Monograph Supplement, 2 (8).

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.
Tolman, E. C. (1928). Purposive behavior. *Psychological Review*, 35, 524-530.
Watson, J. B. (1914). *Behavior: An introduction to comparative psychology*. New York: Henry Holt. Reprinted in 1967 with an Introduction by R. J. Herrnstein. New York: Holt, Rinehart & Winston.
Watson, J. B. (1924, 1925). *Behaviorism*. New York: People's Publishing Co.
Watson, J. B. (1930). *Behaviorism*. New York: W. W. Norton.
Watson, J. B. (1916). The place of the conditioned reflex in psychology. *Psychological Review*, 23, 89-116.
Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist*. Philadelphia: Lippincott.
Watson, J. B., & Watson, R. R. (1928). *The psychological care of the infant and child*. New York: Norton.
Yerkes, R. M., & Morgulis, S. (1909). The method of Pawlow in animal psychology. *Psychological Bulletin*, 6, 257-273.

Author's note

Correspondence address: J. L. Gewirtz, Department of Psychology, Florida International University, Miami, FL 33196. E-mail address: gewirtz@fiu.edu