Chapter 1

Association and Cognition

A. J. Wills

In April 1903, Pavlov presented his early work on the conditioned reflex to the International Congress of Medicine, Madrid (Pavlov, 1903/1928a). Two of the, no doubt numerous, events that celebrated the centenary of this event were the Spring meeting of the Experimental Psychology Society at Exeter (for which I was the local organizer) and the annual Associative Learning Symposium organized by Professors Rob Honey and John Pearce of Cardiff University.

As the date of the Experimental Psychology Society (EPS) conference approached, my thoughts turned increasingly to the nature of the man and the work we were about to celebrate, and how both were considered today. I repeatedly recalled a conversation with a colleague about a year earlier that started with him saying, "Associative learning ... isn't that something they did in the 1950s when they thought pigeons could be trained to guide missiles?"

It's probably best that this particular opinion remains anonymous but, in less extreme forms, it probably reflects the attitude of a considerable minority of today's professional psychologists. Breaking down the opinion into its constituents, it seems to center on the ideas that the study of associative learning is (a) entirely concerned with nonhuman animals, and is (b) part of an episode in the history of psychology that ended with the "cognitive revolution" of the 1960s and 1970s. Both components of this opinion are fundamentally incorrect, as I indicate later in this chapter. However, first there is a need to define terms.

WHAT IS ASSOCIATIVE LEARNING?

In some senses, there is a tension in talking about Pavlov and "associative learning" in the same breath. The term *associative* was in use throughout Pavlov's career, but it was not a term that he favored. For example, in his Madrid lecture he deliberately eschews the term associative, stating that "habits and associations ... are also reflexes, but condi-

tioned reflexes" (Pavlov, 1903/1928a). This perhaps explains, in part, the reaction of another one of my to-remain-anonymous colleagues. I asked her whether she thought Pavlov would have approved of a human associative learning symposium being held in his honor, particularly given that much of his work was with nonhuman animals. Her reply was, "Human testing no problem, I'd guess. It's the arty-farty cognitive empire-building 'conditioning is just one corner of the empire of associative learning' that he might balk at."

This concern can, I think, be partially addressed by a proper definition of terms. In the context of the title and content of this book, "associative learning" is intended to be a description of a class of problem facing the organism. It is not a statement about the nature of the internal processes by which the organism solves that problem. It is also not a statement about the suitability or otherwise of particular theoretical constructs to explain the organism's behavioral reactions. For example, in 1932 Pavlov concluded a discussion of teaching a dog to raise its paw in response to a verbal command in the following way: "Why this is merely simple association, as psychologists usually assume, and by no means not an act of intelligence, of ingenuity-even if of elementary things-remains unclear to me" (p. 124).

In this statement, Pavlov is presumably using the term *simple association* to refer to certain specific process theories discussed earlier in that paper. Hence, the term *association* is being used to describe a specific theory of which he is critical. However, there seems little doubt that the appropriate behavior of the dog in this situation is contingent upon it detecting a reliable associative relationship between a particular type of sound emitted by the trainer and the action that is rewarded. In this manner, the situation Pavlov describes necessarily involves associative learning, but may or may not involve the formation of the "simple associations" assumed by a particular process theory. The confusion inherent in using such similar terms for such different concepts perhaps leads to the conclusion that a more distinct terminology should be used. The preferred nomenclature of this volume is to use "associative learning" to describe a class of problem facing the organism, and to use (somewhat) more theoretically specific terms to describe and differentiate process accounts (e.g., connectionist models, controlled processes).

Associative Learning and Humans

Let's return to the misconception that the study of associative learning is wholly concerned with nonhuman animals. There can be absolutely no doubt that Pavlov considered his work to be of great importance to the understanding of the human condition. Perhaps the most striking manifestation of this view was the 2 months in the summer of 1918 that Pavlov spent studying psychiatric patients. This work resulted in a tentative theory of catalepsy derived by analogy to experiments he had previously performed with dogs (Pavlov, 1919/1928c). A large part of

his last 7 years of research were also dedicated to the application of his findings to psychiatry. Another example of Pavlov's direct interest in human behavior is his article on the "reflex of purpose" (Pavlov, 1916/1928d). In this he suggests that activities such as stamp collecting are natural results of this reflex (Pavlov was a keen stamp collector, by the way). Pavlov also suggests that suicide can result from the chronic inhibition of the reflex of purpose. Against a context of such liberal extrapolations, it would seem unlikely that Pavlov could have had a principled objection to the study of associative learning in humans.

More generally, a minimal knowledge of the history of psychology draws one to the inevitable conclusion that the study of associative learning has always been deep at the heart of human experimental psychology. From early examples such as the "Little Albert" experiments of Watson and Rayner (1920), through studies of the modification of the galvanic skin response (Cook & Harris, 1937), the learning of preferences (Razran, 1938), simulated medical diagnosis (Smedslund, 1963), stereotype formation (Hamilton & Gifford, 1976), right up to recent work on causal reasoning (Cheng, 1997) and the complexities of retrospective revaluation (e.g., Dickinson & Burke, 1996), psychology has seldom strayed far from the consideration of this basic form of activity.

"Associative" (Connectionist) Theories of Human Behavior

If one accepts that associative learning is, and is likely to remain, a topic of core interest in psychology, this does not necessarily mean that "associative" theories (recall our earlier distinction) are of any current use in the explanation of human behavior. One of the most overstated and misunderstood aspects of the history of psychology is the idea that there was a kind of "cognitive revolution" in the 1960s and 1970s. In the "dark ages" that preceded this revolution, behaviorists ruled with an iron first, disallowing any explanation of behavior that hinted at the presence of an internal state. After the revolution, psychologists were once again free to invoke any mental states and processes they liked and everyone was much happier and more productive. It sometimes seems to be assumed that associative theories vaporized at around that time. Recent histories of psychology sometimes add a final chapter. In this part of the story, everyone reads McClelland and Rumelhart's *Parallel Distributed Processing* (Rumelhart, McClelland, & the PDP Research Group, 1986) and has another conversion experience because here is an approach that permits internal states but that is constrained by certain types of neurophysiological data.

This caricatured history is wrong in every important respect. For example, it combines the experimental methodology of behaviorism (reliance on observable behavior) with its philosophy (inappropriateness of mental states as a subject for study). A cursory glance at any modern textbook of cognitive psychology will demonstrate that the methodol-

ogy of behaviorism survived the cognitive revolution. Similarly, when Pavlov explains that his approach focuses the whole of attention "upon the correlation between external phenomena and the reaction of the organism" (Pavlov, 1903/1928a, p. 50), it is the methodology rather than the philosophy that he is endorsing. This distinction is supported by the clear line he draws between his approach and that of investigators such as Watson: "The practical American mind ... found that it is more important to be acquainted with the exact outward behavior of man than with guesses about his internal states ... I and my co-workers hold another position" (Pavlov, 1903/1928a, p. 40).

What was Pavlov's position on internal states? One clear illustration can be found in the article "Natural Science and the Brain" (Pavlov, 1909/1928b). In this, he proposed that the conditioned reflex developed through the action of two types of brain process. The first of these was a "temporary union, i.e., the establishment of a new connection in the conducting paths" (p. 122). This idea is foreshadows the concept of the "associative link" or "connection weight" employed in many contemporary connectionist (or "associative") models. The second type of brain process was the action of "analyzers." The concept of analyzers corresponds closely to the "nodes" of a contemporary connectionist system.

It's worth just underlining the implications of this article. In 1909, Pavlov (1909/1928b) proposed a theory of associative learning that is recognizably a form of connectionist theory—the sort of approach that a caricatured history of psychology attributes to postcognitive research almost 80 years later. One standard response to this objection is that although associative theories were popular for a time, the work of Minsky and Papert (1969) led to psychologists abandoning associative theory until Rumelhart, Hinton, and Williams (1986) invented back-propagation. One minor problem with this account is that back-propagation was discovered at least 12 years earlier (Werbos, 1974). A more major problem is the simple fact that associative theories were *not* abandoned. One can trace a steady development of associative theory from Thorndike (1898) and Pavlov, through (to name but a few) Spence (1936), Hull (1943), Konorski (1948), Hebb (1949), Rescorla and Wagner (1972), Mackintosh (1975), Grossberg (1976), Pearce and Hall (1980), Wagner (1981), and Pearce (1987). Much of this theoretical work predates *Parallel Distributed Processing* and even some much later work (e.g., Kruschke, 1992, 2001).

"I Tell My Students Pavlov Was a Genius and They Just Gape at Me"

This is another quote from a colleague. I think it reflects a lack of understanding (on the students' part) of the nature and extent of Pavlov's contribution to cognitive psychology. Pavlov was one of the pioneers of the objective, scientific study of associative learning. He was also a pioneer of a theoretical approach to explaining behavior that eventually be-

came known as connectionism. This combination alone makes him a critical figure in the development of cognitive psychology as we know it today. Additionally, people tend to forget the breadth of phenomena within classical conditioning that are attributable to him. Pavlov's body of work goes far beyond the basic demonstration of the development of a conditioned response. Pavlov also demonstrated extinction, and showed through reminder treatments that extinction was not simply unlearning. He provided experimental demonstrations of generalization decrement, context-specificity, and the importance of contiguity. All of this appears in his first public presentation in 1903. By 1912 he had developed the contemporary concept of conditioned inhibition (although his tests for the presence of conditioned inhibition are now generally acknowledged as inadequate).

What makes this all the more remarkable is the fact that Pavlov was over 50 before he started work on the conditioned reflex. In fact, it was not until he was 21 that he decided against becoming a priest, and started to train as a physician. He became a physician at 30, a professor of pharmacology at 41, and a professor of physiology at 46. He won the Nobel prize at 55 for his work on the digestive glands. His most famous book, *The Conditioned Reflexes* (Pavlov, 1927), was published when he was 78.

There is not the space in this volume to provide a detailed biography (Gantt, 1928, provides an excellent introduction). Nevertheless, some aspects of Pavlov's personality are worth mentioning briefly. In one respect, Pavlov was a stereotypical academic. He is reported to have had extraordinarily poor dress sense and that on the rare occasions he bought his own clothes "his choice of colors ... made his friends laugh and his family angry" (Gantt, 1928, p. 14). In other respects, he is much harder to pigeonhole. Pavlov was a keen gymnast in his youth, a keen gardener throughout his life, and a proponent of the importance of physical exercise. He had a very high level of manual dexterity, being a highly accomplished surgeon, and he also had an exceptionally good memory. He was widely regarded as a clear, energetic, and popular lecturer. This comes across in his collected lectures (Gantt, 1928) despite the barriers of the intervening century and the translation from Russian into English. Within the context of his times, Pavlov was also distinguished by his regard for animal welfare, pioneering postoperative care for experimental animals and developing alternatives to vivisection. In 1924, he resigned his professorship in protest at student expulsions. He didn't smoke or drink, but did swear copiously. He worked 7 days a week, but took at least 2 months off each summer.

OUTLINE OF THIS BOOK

This book emerged from the work presented at the "New Directions in Human Associative Learning" symposium (Experimental Psychology Society meeting, Exeter, England, April 10–11, 2003) and some of the

human work presented at the "Associative Learning Symposium" (Gregynog, Wales, April 15–17, 2003). Chapters 3, 6, 7, 10, and 11 were from the Experimental Psychology meeting, whereas chapters 2, 4, and 8 emerged from the Associative Learning Symposium.

This book has been divided into three sections. In the first section (chaps. 1–4) the reader is introduced to some recent data and controversies in the study of associative learning. In the second section (chaps. 5–8) the focus shifts to recent developments in the formal theories of how associative learning occurs. All three chapters are written from the perspective of connectionist modeling and are contributed by a group of authors who have a broadly similar outlook and background. The final section (chaps. 9–11) turns to some of the more applied work on human associative learning; specifically, its application to depression and to the development of preferences (e.g., product preferences). The first chapter of each section provides a short introduction to the material covered.

OUTLINE OF THIS SECTION

This first section of the book begins with a chapter by Andy Baker, Robin Murphy, and Rick Mehta, in which they introduce two well-known theories of human associative learning. First up is Patricia Cheng's PowerPC theory, which is one example of the class of *statistical* (or *normative*) accounts. Next up is Bob Rescorla and Alan Wagner's eponymous theory. Rescorla–Wagner theory is one example of the class of *associative* (or *connectionist*) accounts. Baker, Murphy, and Mehta then go on to review some of the previous attempts to determine whether PowerPC or Rescorla–Wagner is the better theory of human causality judgments. Five of these studies (from Rick Mehta's doctoral thesis) are presented in detail, and are concerned with the role of causal order (i.e., whether causes are observed prior to effects, or effects prior to causes). Neither PowerPC nor Rescorla–Wagner adequately accounts for the causal-order effects presented. Additionally, the authors lead us to the conclusion that the comparison of a specific statistical account with a specific connectionist account tells us little about the relative merits of the class of theory from which each come. Indeed, their final conclusion is that statistical and connectionist accounts exist at different levels of explanation and hence should not be considered as being directly in competition. Statistical accounts help us understand what might need to be computed in order for us to behave adaptively. Associative accounts provide potential mechanisms for these computations.

In contrast, Jan De Houwer, Stefaan Vandorpe, and Tom Beckers argue (chap. 3) that associative accounts as a class fail to provide a mechanism for certain known and important phenomena in human associative learning. They attribute this failure to the fact that associative theories are typically accounts of automatic (i.e., unintentional or unconscious) processes. Human behavior presumably also results from

nonautomatic processes, in other words processes that are intentional, conscious, or effortful. Jan and colleagues review the evidence that human associative learning (a) requires awareness of the learned associations, (b) is effortful in the sense that it disrupts demanding secondary tasks, and (c) can result from both direct experience and verbal instruction. They also point out that human associative learning can show generalizations to novel situations of a type difficult to explain without recourse to abstract rules.

Finally, they review evidence that suggests deductive reasoning can play an important role in determining causal judgments. Perhaps the most striking of these demonstrations is the presence of higher order retrospective revaluation effects in human associative learning. This phenomenon (which is described more fully in the chapter itself) is naturally explained by the recursive application of a deductive rule but is particularly problematic for many associative accounts. De Houwer and colleagues conclude that an understanding of controlled processes is crucial to our understanding of human associative learning. As they readily admit, this conclusion introduces an additional challenge to theorists. Theories of human associative learning that include controlled processes are currently highly descriptive and hence poorly specified. This is in stark contrast to associative theories of automatic processes, which have reached a high level of formal specification (see chaps. 5, 6, 7, and 8). Jan and colleagues' chapter was almost certainly the most controversial in the process of internal peer review this book went through, and seems likely to stimulate debate for some time to come.

In chapter 4, Jason Tangen, Lorraine Allan, and Hedyeh Sadeghi combine some of the themes of chapters 2 and 3. They start with a brief review of certain basic effects in human associative learning, and of the Rescorla–Wagner theory. Then, similarly to Andy Baker's chapter, they discuss the challenges causal-order effects present to both normative and associative accounts of human associative learning. Jason Tangen's suggested resolution of this problem is that normative accounts and associative accounts represent different strategies that can be applied by the human participant. From this starting point, Jason reports a series of experiments designed to investigate the conditions under which an associative or a normative strategy is most likely to be applied.

REFERENCES

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Cook, S. W., & Harris, R. E. (1937). The verbal conditioning of the galvanic skin reflex. *Journal of Experimental Psychology, 21*, 202–210.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *The Quarterly Journal of Experimental Psychology, 49B*, 60–80.

Gantt, W. H. (1928). Ivan P. Pavlov: A biographical sketch. In W. H. Gantt (Ed.), *Lectures on conditioned reflexes* (pp. 11–34). London: Lawrence & Wishart.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: Part I. Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23,* 121–134.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12,* 392–407.

Hebb, D. O. (1949). *The organization of behavior.* New York: Wiley.

Hull, C. L. (1943). *Principles of behavior.* New York: Appelton–Century–Crofts.

Konorski, J. (1948). *Conditioned reflexes and neuron organization.* Cambridge, England: Cambridge University Press.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45,* 812–863.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276–298.

Minsky, M. L., & Papert, S. A. (1969). *Perceptrons: An introduction to computational geometry.* Cambridge, MA: MIT Press.

Pavlov, I. P. (1928a). Experimental psychology and psycho-pathology in animals. In W. H. Gantt (Ed.), *Lectures on conditioned reflexes* (Vol. 1, pp. 47–60). London: Lawrence & Wishart. (Original work published 1903).

Pavlov, I. P. (1928b). Natural science and the brain. In W. H. Gantt (Ed.), *Lectures on conditioned reflexes* (Vol. 1, pp. 120–130). London: Lawrence & Wishart. (Original work published 1916).

Pavlov, I. P. (1928c). How psychiatry may help us to understand the physiology of the cerebral hemispheres. In W. H. Gantt (Ed.), *Lectures on conditioned reflexes* (Vol. 1, pp. 287–293). London: Lawrence & Wishart. (Original work published 1919).

Pavlov, I. P. (1928d). The reflex of purpose. In W. H. Gantt (Ed.), *Lectures on conditioned reflexes* (Vol. 1, pp. 275–281). London: Lawrence & Wishart. (Original work published 1916).

Pavlov, I. P. (1927). *Conditioned reflexes.* New York: Dover.

Pavlov, I. P. (1932). The reply of a physiologist to psychologists. *Psychological Review, 39*(2), 91–127.

Pearce, J. M. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review, 94,* 61–73.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552.

Razran, G. (1938). Conditioning away social bias by the luncheon technique. *Psychological Bulletin, 35,* 693.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research* (pp. 64–99). New York: Appleton–Century–Crofts.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 318–364). Cambridge, MA: MIT Press.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing.* Cambridge, MA: MIT Press.

Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology, 4*(3), 165–173.

Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review, 43,* 427–449.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review, 8.*

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Lawrence Erlbaum Associates.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Werbos, P. J. (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences.* Unpublished doctoral dissertation, Harvard University, Cambridge, MA.