5 Associations

Matthew J. C. Crump

2021-09-01

This chapter covers associationism and early empirical work on associative learning processes in humans and non-human animals.

The previous chapter covered research on cognitive abilities spanning the 1860s to the 1960s, from the perspective of individual differences psychology. For example, in addition to developing mental tests, Alfred Binet wrote more broadly about his school of "Individual Psychology" (Nicolas et al., 2014) that was aimed at using measurements of variation in human attributes to make society more efficient. However, the individual differences school of psychology was not the only approach to studying cognitive abilities. Instead, methods of experimental psychology were also used to ask and answer questions about cognitive abilities during the same time period. Experimental psychologists used similar tests and procedures, but they focused on how task performance varied across experimental manipulations rather than across different individuals. The results of the experiments were used to evaluate theories about how cognitive processes work. This chapter will focus on associationism and early empirical work on associative learning processes in humans and non-human animals that were part of the experimental psychology tradition.

0.0.1 Evolving ideas on cognition

There is a long history of ideas about cognition that predates and informs later psychological approaches to cognition. By focusing on the early experimental research conducted from the

1890s to 1920s, it is possible to refer to elements of this intellectual history and discuss how it shaped and guided cognitive research. Some of the really big claims and themes about cognition that appear in the introductions of early research manuscripts remain relevant today.

We are about to visit the laboratories of early researchers studying associative learning in humans and animals. First, we will go back to James McKeen Cattell's lab. In addition to developing the mental tests discusses in the last chapter, Cattell was also asking questions about how people used and developed association. Then, we turn to the labs of Edward Thorndike (1874-1949) and Ivan Pavlov (1849-1936). Thorndike was in America and Pavlov was in Russia. They were both independently using different experimental approaches to test theories and claims about animal cognition. Before visiting each lab the next sections provide philosophical and historical context that motivated the research questions.

0.0.1.1 Humans and animals

So far we have discussed cognition in terms of human animals. What about cognition in other animals? Before Thorndike and Pavlov there wasn't much experimental research on animal cognition. Since then, whole fields of animal and comparative cognition have been developed. Although this textbook focuses on human cognition, I will attempt to integrate animal cognition as much as possible. And, rather than using "human-animal", and "non-human animal", I will refer to human and/or animal cognition throughout the book as a shorter way to identify the subjects of the research. Last, discussions about which words to use to refer to humans and animals point to background ideas motivating Thorndike and Pavlov's work.

Just like the long human history of ideas about the mind, there is a long history of ideas about animals, animal minds, and the human-animal relationship. Anthropocentrism is a collection of beliefs and traditions that center humans as the most important species. For example, in early Judaic and Christian writings, a supreme supernatural being created humans in their image and set humans on a higher level than other animals. These ideas both distinguish humans from animals (as being different kinds of entities), and place humans in a hierarchy of quality above other animals. Judeo-Christian concepts predate psychology by thousands of years, and even though psychology as a natural science proceeded to construct different views on humans and animals, related hierarchical notions remain deeply ingrained. For example, "higher-order" cognition is often reserved for cognitive abilities argued to be special to humans, like inferential reasoning. And, "lower-order" cognition is often reserved for basic cognitive abilities that are common among animals.

Animism is a set of beliefs ascribing spiritual essences to things, including animals. Many fables, folklore, and religious texts have animal characters imbued with cognitive abilities and other powers. Relatedly, the idea that animals can be understood as if they were humans is also criticized as an inaccurate form of anthropomorphism.

Mystic ideas about humans and animals can be viewed as claims about cognition. One claim is about whether or not human cognition can be explained at all. For example, the idea that people were created in the image of a fundamentally unexplainable supernatural being suggests that the supernatural parts of people are inherently unexplainable. Another claim is about types and qualities of cognitive abilities. For example, the idea that animals possess vital spiritual essences like humans, could suggest that the mental abilities of some animals are similar to people. Experimental Psychology arose out of the scientific tradition to put claims to the test by collecting evidence bearing on the claims, including claims about human and animal cognition.

0.0.1.2 Philosophy

The experiments of Cattell, Thorndike, and Pavlov inspired by philosophical debates in epistemology between rationalist and empiricist views of knowledge.

Rationalism involved views that some knowledge was innate and existed separately from experience and information gained through sense organs. Furthermore, ultimate truths about reality were argued to depend on logic and reason. If the universe was a fundamentally logical place, then the process of accurate reasoning alone would be enough to deduce the ultimate truths.

Empiricism additionally emphasized a role for observation and evidence collection in knowledge creation. For example, humans were argued to acquire knowledge through their sensory experience of the world. Empiricism invited further questions about how people created knowledge from their sensory experience. These "how" questions inspired early experimental psychologists.

The Associationist School included empiricist philosophers who speculated further on the nature of mental processes that were responsible for producing knowledge from experience. In 1689, John Locke wrote "An Essay Concerning Human Understanding" and argued against the rationalist/nativist idea that people were born only with innate knowledge about the world (Locke, 1847). Locke advocated that people acquired knowledge by learning about the world through their experiences.

Locke's An Essay Concerning Human Understanding is available to read on the internet archive

In 1739, David Hume wrote "A Treatise of Human Nature" and put forward a role for associations in the learning process (Hume, 1896). For example, he wrote that "when the mind, therefore, passes from the idea or impression of one object to the idea or belief of another, it is not determined by reason, but by certain principles, which associate together the ideas of these objects, and unite them in the imagination." Hume suggested that acts of cognition involve a process of association that works according to "certain principles". What the principles are, and how they work, is still a primary focus of the modern cognitive sciences.

Hume's A Treatise of Human Nature is available to read on the internet archive

Philosophers also made proposals about the principles guiding the process of association. *The principle of contiguity* states that strength of association depends on the proximity of events in space and time. Events that are closer to each other are associated more strongly. *The principle of similarity* states that more similar events will develop stronger associations than less similar events. *The principle of frequency* is that events that co-occur more frequently will be associated more strongly than events that co-occur less frequently. *The recency principle* suggests stronger associations for recent events than more remote events. In general, these philosophical principles of association have held up quite well and are often components of modern theories.

0.0.1.3 Natural Science

Lesser known work from Robert Hooke (1635-1703), a natural scientist, went much further than Hume's "certain principles". Robert Hooke, coined the word "cell" and was the first person to observe a micro-organism under a microscope. In 1705, Hooke's posthumous works were published and they contained his model of how human memory could operate as a physical system (Hooke, 1705). The model was not entirely physical, because it allowed some role for "immaterial" forces, and it was largely forgotten until fairly recently (Hintzman, 2003). However, despite not having a major historical impact, Hooke's model was a clear attempt to develop a well-specified mechanistic explanation of cognition [1]. The aim of generating a clear explanation of how cognition works, especially in the form of a mechanistic model, remains a current goal of the cognitive sciences.

0.0.1.4 Evolution

A final historical backdrop was Darwin's theory of evolution, published in 1859 in a book titled "On the origin of species" (Darwin, 1859). Darwin's theory connected all life on Earth and explained the origins of species in terms of natural selection processes. The critical ingredients for organisms to evolve over generations included:

On the origin of species is available to read on the internet archive

1. The ability to reproduce
2. Random mutations that produce heritable variations in the traits and behavior of the organism, and
3. Environmental selection pressures.

Organisms that survived in their environment would tend to pass on heritable traits to their offspring. Organisms that perished would be less likely to have offspring and pass on heritable

[1] We will examine models of human memory in greater detail in the upcoming chapters on memory

traits. Random mutations also occur in a generation that may or may not convey a survival advantage. And, an organisms environment can change, which can change whether or not an existing trait remains useful for survival. In this way, the traits of organisms slowly drift and change as some traits are more likely to persist or fade across generations.

Different from anthropocentric views described earlier, the theory of evolution clearly places humans and animals in the same explanatory ballpark. People are a species of animals with an evolutionary family tree, where every person is descended from their parents. Other animals also have their own evolutionary family trees and lineages of descendants. Furthermore, if one goes far enough back in the family tree all animals could be descended from common ancestors.

Evolutionary theory motivates cognition as an inclusive topic for humans and animals. First, humans are animals; so, by definition human cognition is a specific case of animal cognition. There are a wide variety of animals that range in their physical size and capabilities, and it is implausible that a natural evolutionary process would only bestow cognitive processes onto humans and not other animals. Instead, evolutionary theory implies that cognitive processes among animals also evolved over time. In this sense, the field of cognition is as large and diverse as all the ways that cognitive processes have evolved differently across species. At the same time, evolution is known to produce similar solutions across species (e.g., eyes), and it is possible that animals share related cognitive processes that work on the basis of similar principles.

In summary, the backdrop to labs we visit next was western European cultural beliefs about humans and animals, western philosophical debates about psychological constructs like human knowledge, a desire to apply the rigors of natural science methodology to problems in psychology, a convincing theory of evolution suggesting that animals could be used as subjects to gain knowledge about evolutionary basic cognitive processes, and a wide open playing field where there was very little existing empirical work.

0.0.1.5 Associative Claims

Associationism is the topic of this chapter and the philosophical theme motivating the research we will discuss from the labs of Cattell, Thorndike, and Pavlov. Associationist claims about cognition are specific enough that they can be evaluated with evidence. As a result, it becomes possible to use the scientific method to assess claims being made about how cognition works. Let's identify a few really basic claims that could be evaluated.

1. People have associations between concepts
2. New associations can be learned
3. Some associations are stronger than others

First, could you think of examples from your experience that would provide evidence for these claims? Have you ever learned a new association between one thing and another? Are some ideas more strongly associated with others in your experience?

Personally, I've learned many things that could involve associations. For example, I didn't know how to type on a computer keyboard when I was born, I learned how to do it. Today, I barely think about what my fingers are doing when I type because the words I'm thinking are strongly associated with the finger movements I need to make to type the sentence I want to write.

Using everyday experiences it is easy to demonstrate for yourself that some associations are stronger than others. For example, think of a fruit that begins with the letter "A". How long did that take?

Think of any word that has a letter A in the 5th position of the word? Did that take longer?

If you thought word "apple" straight away, but took much longer to think of a word with letter "A" in the fifth position, then it seems that the letter A is more strongly associated with some words than others.

Although, you could come up with additional examples from your experience that would be consistent with the claims, consider the challenge of creating a laboratory-based demonstration capable of generating credible evidence relevant to evaluating a claim.

One goal of a laboratory-based demonstration is to communicate a method and controlled set of circumstances that other people could also use. This would others to reproduce the demonstration and verify the results.

The following examples are early laboratory demonstrations that were used to evaluate associationist claims.

0.0.2 Cattell's Associative Reaction times

In the previous chapter, James McKeen Cattell was introduced as psychologist conducting mental ability testing, even before Binet. Cattell also used experimental psychology methods to gather evidence about associative processes in people.

0.0.2.1 Naming times

In one set of studies he measured how long it took people to see and name objects (Cattell, 1886; CATTELL, 1886). The ability to identify an object by naming it out loud was presumed to involve an association between perceiving the object and the action needed to utter the object's name. One of Cattell's findings was that people took twice as long to read words that "have no connexion" compared to words that were composed in a sentence.

For example, Cattell's subjects were twice as fast to read a regular sentence versus a scrambled sentence [2]

> Read each of the WORDS below as fast as you can
>
> 1. *Regular Sentence*: The candy at the store was red and very tasty
> 2. *Scrambled Sentence*: very at red and The tasty the was store candy
>
> *Were you slower to read the words in the scrambled sen-*

[2] These are not original stimuli used by Cattell

tence?

Similarly, when the task was to read individual letters one a time, Cattell's subjects were twice as fast to read letters when they occurred in words compared to when they did not. This general finding was later re-discovered and termed the word-superiority effect in 1969 (Reicher, 1969).

Read each of the LETTERS as fast as you can

1. *Letters in words*: very at red and The tasty the was store candy

2. *Scrambled letters*: tyshe noact aer ead ta rrdth des etyv Twnasy

Was it harder to read the scrambled letters?

Let's relate Cattell's general findings to the associative claims from earlier:

1. People have associations between concepts
2. New associations can be learned
3. Some associations are stronger than others

The fact that people can read words or letters at all provides some evidence for the first claim. For example, letters on a page are an arbitrary symbol system with culturally accepted rules for pronunciation. When people read aloud, they are interpreting the visual format of words and letters on a page and transforming them into vocal speech act. This is consistent with the very general claim that people have associations between visual and auditory formats for words and letters.

The fact that participants could read Cattell's words and letters is consistent with the second claim too. People are not born knowing how to read. Reading skill is acquired with practice. The fact that people can learn to read provides evidence that people can learn new associations. Of course, these general observations about people's ability to read could already be made without Cattell's experiments.

Cattell's results were also consistent with the claim that some associations are stronger than others. For example, it appears that reading time does not just involve the time it takes to recognize and say a word or letter. If this was strictly the case, then people would take the same amount of time to read words and letter no matter if they occurred in sentences or words. Instead, people were faster to read words and letters when they appeared in familiar contexts, like sentences and words. This suggests that information from the surrounding context facilitates reading, possibly through some kind of association.

0.0.2.2 Association reaction times

In the following year, Cattell applied reaction time methods to the problem of measuring component processes that may be involved in associative processes (Cattell, 1887). He wrote:

> "[In previous work..] 0.4 seconds was needed to see and name a word. When the physiological factors and the time taken up in seeing the word were eliminated, it was found that about 0.1 seconds was spent in finding the name belonging to the printed symbol. The time was longer for letters, which we do not read as often as words, and still longer (about .25 sec) for colours and pictures. I called the time passing, while the motor expression was being found, a 'Will-time'."

Reaction time methods are elaborated upon in the next chapter. For now, it is worth pointing out two ways that reaction time measurements can help us understand how cognitive processes might work. First, it is possible to measure reliable differences in reaction times when people complete particular tasks. For example, Cattell estimated that people need about 400 milliseconds to see and read a word, and that pictures took a little bit longer. Second, it is possible to speculate about the individual components of processing that account for the total reaction time. For example, Cattell divides up the reaction time for a word into physiological factors and time taken up be seeing the word, followed by the time taken to search and recall the name of the word, and finally the time taken to form the action to say the word. In other words, Cattell is proposing that the inner-workings of cognitive processes used to accomplish a whole task like reading a word may be separated and indirectly measured in terms of their temporal processing times.

In the remainder of his paper, Cattell gave subjects an association task that involved 1) receiving a cue/prompt, and 2) responding with a known association to the cue. For example, participants were shown a picture and asked to name it in their first or second language. Cattell found that people were faster at picture naming in their first language than second language. He explained this finding in terms of practice: people had more practice naming objects in their first than second language. However, this raises more questions about why practice changes how fast an object can be named.

Another study in the paper simply measured the times associated with remembering different kinds of facts given a cue. For example, one cue was a city name and subjects were asked to remember the country it was in. Another cue was a month name, and subjects had recall what season it was in, or the preceding month, or the following month. Cattell reported average reaction times and showed consistent differences depending on the conditions of the cue and associative response. Cattell speculated that the pattern of different reaction times reflected corresponding differences in the mental operations needed to carry out each task. Tasks that required more mental operations to complete were assumed to take longer than tasks requiring fewer mental operations to complete.

0.0.3 Thorndike's puzzle boxes

Edward Thorndike (1874-1949) was a student of Cattell's who took some of the first experimental approaches to investigating associative processes in non-human animals [3].

Thorndike viewed previous claims about animal cognition as belonging to two wildly different camps, with almost no middle

Figure 1: Edward Thorndike (1874-1979).

[3] Thorndike also followed Cattell in becoming a strong proponent and public figure in the American eugenics movement.

ground. In one camp, some animals might be very nearly human, potentially being able to reason and form associations on par with people. In the other camp, animals might be simple reflex machines and nothing more. Thorndike was interested in producing evidence using laboratory methods that could be used to assess claims about animal cognition.

Here are a few choice quotes from Thorndike's doctoral thesis (Thorndike, 1898) that illustrate his thinking and approach to the study of animal intelligence.

> "We do not know how delicate or how complex or how permanent are the possible associations of any given group of animals."

> "We say that the kitten associates the sound"kitty kitty" with the experience of nice milk to drink, which does very well for a common-sense answer. It also suffices as a rebuke to those who would have the kitten ratiocinate about the matter, but it fails to tell what real mental content is present. Does the kitten feel "sound of call, memory-image of milk in a saucer in the kitchen, thought of running into the house, a feeling, finally, of 'I will run in'?" Does he perhaps feel only the sound of the bell and an impulse to run in, similar in quality to the impulses which make a tennis player run to and fro when playing? The word association may cover a multitude of essentially different processes, and when a writer attributes anything that an animal may do to association his statement has only the negative value of eliminating reasoning on the one hand and instinct on the other…To give to the word a positive value and several definite possibilities of meaning is one aim of this investigation."

> "Surely every one must agree that no man now has a right to advance theories about what is in animals' minds or to deny previous theories unless he supports his thesis by systematic and extended experiments. My own theories… will doubtless be opposed by many. I sincerely hope they will, provided

the denial is accompanied by actual experimental work. In fact I shall be tempted again and again in the course of this book to defend some theory, dubious enough to my own mind, in the hope of thereby inducing some one to oppose me and in opposing me to make the experiments I have myself had no opportunity to make yet."

To summarize these quotes, Thorndike was interested in settling debates about animal intelligence using laboratory techniques and the scientific method. This would involve creating reproducible situations in the lab where animal behavior could be manipulated and observed between conditions of an experiment. Thorndike argued that his methods would produce evidence relevant to evaluating claims of animal intelligence. He suggested his methods could be used by other researchers to verify his findings. And, that his methods could be improved upon by others to provide more stringent tests of his theories and claims. Thorndike's method of assessing associations in animals involved puzzle boxes, or the modern equivalent of an escape rooms for animals.

0.0.3.1 Thornike's basic methodology

Thorndike conducted experiments on cats, dogs, and chicks. His experimental apparatus was a puzzle box, like the one depicted in Figure 2.

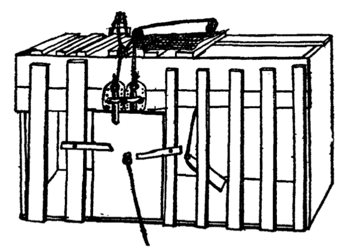


FIG. 1.

Figure 2: A drawing of one of Thorndike's puzzle boxes.

Thorndike made different puzzle boxes with similar fundamental features. First, animals could be placed inside the box. Second, animals could escape from the box by solving the puzzle. For example, pulling a latch or a hoop on a string would unlock a door allowing the animal to escape. Animals were typically deprived of food and they were hungry before they were placed in the box. When they escaped they were given food as a reward. Animals were given practice attempts to get out of the box and Thorndike measured the amount of time needed to get out of the box for each attempt.

0.0.3.2 Putative Mental Components of Association

Thorndike's methods produced clear findings. His first main finding was that the animals could figure out the tricks and escape from the box. His second main finding was that the animals became faster at escaping with practice.

After making these demonstrations, Thorndike then considered questions such as: How were the animals solving each problem? How were they getting faster? What kinds of associations were involved?

Thorndike sought to achieve clarity regarding the types of associations that might be involved. He speculated on the logical stages and types of associations that his animals might have learned. As an example, he wrote:

> There might be in an association, such as is formed after experience with one of our boxes, the following elements:

1. Sense-impression of the interior of the box, etc.
2. Discomfort and desire to get out.
3. Representation of oneself pulling the loop.
4. Fiat comparable to the human "I'll do it."
5. The impulse which actually does it.
6. Sense-impression of oneself pulling the loop, seeing one's paw in a certain place, feeling one's body in a certain way, etc.
7. Sense impression of going outside.
8. Sense impression of eating, and the included pleasure. Also between 1 and 4 we may have
9. Representations of one's experience in going out,
10. Of the taste of the food, etc.

This list details experiences and impulses that presumably occur from the time when an animal enters the box, to the time after it has escaped and is eating the food reward. The general associationist explanation would be these components are associated with one another, possibly in a chain, such that triggering one of the elements would cause a chain reaction, and successively trigger the next associated element, thereby allowing the animal to proceed through the puzzle box and get the food reward.

0.0.3.3 Experimental questions about associative processes

Thorndike also went on to conduct experiments with his puzzle boxes to test ideas about animal learning. These experiments involved manipulations intended to modify some aspect of the learning process.

0.0.3.3.1 Imitation Learning

One question was whether or not an animal could learn how to escape from a puzzle box just by watching another animal escape first. This question could be answered with an experiment. One group was allowed to watch another animal escape first, the other group was not allowed to watch. The empirical question was whether the group that was allowed to watch would solve the puzzle box faster than the other group.

Thorndike found that cats, dogs, and chicks did not benefit from watching other animals solve the puzzle box. This didn't definitely rule out imitation as a possible source of knowledge. But, it suggested to Thorndike that some elements of the associations needed to be experienced directly to be learned. Thorndike mentioned plans to test whether an ape would benefit from imitation learning, but he was unable to conduct the experiment because the ape was too hard to handle.

0.0.3.3.2 Mental Representation

Thorndike was optimistic his methods could help settle questions of mental representation in animals. For example, do animals have internal images that are used in a network of associations? Thorndike proposed the following experiment:

> "The only logical way to go at this question and settle it is, I think, to find some association the formation of which requires the presence of images, of ideas. You have to give an animal a chance to associate sense-impression A with sense-impression B and then to associate B with some act C so that the presence of B in the mind will lead to the performance of C. Presumably the representation of B, if

> present, will lead to C just as the sense-impression B did. Now, if the chance to associate B with A has been improved, you ought, when the animal is confronted with the sense-impression A, to get a revival of B and so the act C. Such a result would, if all chance to associate C with A had been eliminated, demonstrate the presence of representations and their associations."

0.0.3.3.3 General concept formation

Thorndike attempted experiments to determine the ability of his animals to form general concepts about the puzzle boxes. Consider a kitten who has learned to escape from a puzzle box by pulling a loop on a string. What kind of knowledge has the kitten acquired that mediates this newfound ability. Has the kitten developed general concepts about how to escape, or is the learning more specific?

It is possible that the kitten has learned something very, very specific. For example, the learning could involve the specific spatial layout of the box, along with the specific location of the loop, and the specific actions that need to happen to pull the loop in that location in that box. Alternatively, maybe the kitten learned a more general concept, something like "to escape the box, go around and find the loopy thing, and then tug it".

Thorndike tested the specificity of learning in **transfer** experiments. He had animals learn to escape from a first box and then tested their ability to escape from new boxes. The empirical question was whether learning to escape from the first box would help animals escape from the next box. For example, if a kitten learned to pull a loop on the left side of the box to get out, would they then quickly escape out of a box with the loop on the right side?

Thorndike found evidence of **positive transfer**, which means that the training experience conveyed a benefit on the transfer test. His animals learned to escape from a new puzzle box faster when it was similar to a puzzle box they had learned previously.

Although evidence of positive transfer is consistent with a claim that the animals were learning general concepts about the puzzle box, Thorndike favored the view that animals were not learning general concepts, and instead were learning about specific details of the puzzle box that happened to transfer well to similar boxes. In other words, he attempted to explain the phenomena of transfer without assuming that kittens had human-like general reasoning abilities.

0.0.3.4 Further associative questions

Thorndike developed lab methods to generate evidence capable of addressing questions about animal cognition. He hoped other researchers would also use methods from experimental psychology to challenge and extend his work. He also entertained questions about associations that ought to be studied in the future:

1. Delicacy and permanence of associations: How fragile are some associations, how long do associations last after they have been formed?
2. Complexity of associations (Thorndike intended to rank intelligence of animals as a function of the complexity of associations they could acquire)
3. Number of associations: How many associations do different animals have?
4. Inhibition of instinct by habit: Can an animal learn to override an instinctual behavior through associative learning?
5. Role of attention: does the formation of an association depend on attending to sense-impressions?

0.0.4 Pavlov's Classical Conditioning

While Thorndike studied associative processes in animals, Ivan Pavlov was simultaneously conducting independent research on the same topic in Russia. Pavlov was not aware of Thorndike's work until much later, and Pavlov approached the topic of associative learning using different methods. Pavlov's experiments were translated to English in a 1927 lecture series (Pavlov, 1927).

Thorndike approached animal learning from the perspective of experimental psychology, whereas Pavlov was skeptical of psychology as a natural science and approached questions about animal learning from the perspective of a physiologist. Whereas Thorndike considered "psychic" phenomena, such as mental images, in his explanations of animal learning, Pavlov was more interested in measuring observable phenomena such as behavior and physical substances, like secretions, apparently produced by brain processes. Here are a few quotes from Pavlov:

> "The cerebral hemispheres stand out as the crowning achievement in the nervouse development of the animal kingdom. These structures in the higher animals are of considerable dimensions and exceedingly complex, being made up in man of millions upon millions of cells–centres or foci of nercous activity– varying in size, shape, and arrangement, and connected with each other by countless branchings from their individual processes. Such complexity of strucgture naturally suggests a like complexity of function, which in fact is obvious in the higher animal and in man. Consider the dog, which has been for so many countless ages the servant of man. Think how he may be trained to perform various duties, watching hunting, etc. We know that this complex behaviour of the animal, undoubtedly involving the highest nervous activity, is mainly associated with the cerebral hemispheres. If we remove the hemispheres in the dog (Goltz, 1892), the animal becomes not only incapable of performing these duties but also incapable even of looking after itself."
>
> "In astounding contrast with the unbounded activity of the cerebral hemispheres stands the meagre content of present-day physiological knowledge concerning them. Up to the year 1870, in fact, there was no physiology of the hemispheres; they seemed to be out of reach of the physiologist".

Pavlov wrestled with whether or not psychology should be invoked to explain brain functions. He suggested psychological perspectives were imprecise and prone to non-physical interpretations that implicated the brain "special psychical activity". For example, Pavlov wrote that "there is no need for the physiologist to have recourse to psychology", and that "investigations of the physiological activities of the hemispheres should lay a solid foundation for a future true science of psychology."

Figure 3: Réne Descartes (1596-1650)

0.0.4.1 Descartes' reflex

This section provides some philosophical context prior to our discussion of Pavlov's experiments. Pavlov connected his work to philosophical ideas about humans and animals that were popularized by René Descartes (1596-1650). Descartes was a rationalist philosopher who advanced a dualist perspective of the mind. According to Descartes, humans dually possess a physical body and an immaterial soul. The body is a physical machine that obeys the laws and principles of physical machinery. However, the soul does not operate according to physical laws. Descartes' human soul was also a mysterious quality that set humans on a higher plane than animals, which were viewed as inferior– animals had physical bodies, but lacked a soul. Although Descartes advocated for dualism, his ideas about how human and animal bodies behave like complicated machines were inspirational to physiologists (Fearing, 1929), who would later become proponents of explaining cognition solely in terms of physical processes.

One of Descartes' core ideas was the concept of a reflex, which involves a clear-cut line of cause and effect between an impending stimulus or impulse, and a subsequent effect. A common example is the knee-jerk reflex shown in Figure 4, where tapping a knee in the right spot can cause a leg to automatically kick up a little bit.

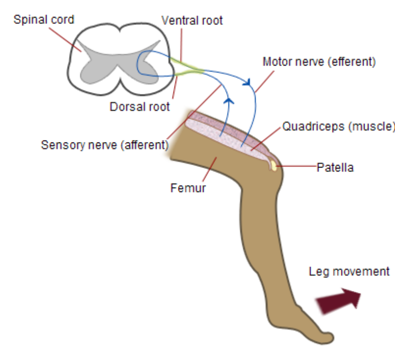
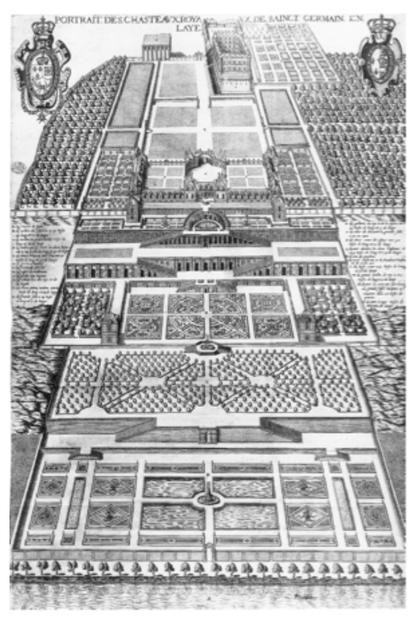


Figure 4: A depiction of the knee-jerk, or Patellar reflex.

In describing humans and animals as physical machines, Descartes was inspired by the gardens at Saint-Germain-en-Laye (Vaccari & Philosophy Documentation Center, 2012) shown in Figure 5. The gardens were a marvel of hydraulic engineering. They contained an extensive network of pipes



connected to fountains, and even controlled statues that had moving parts.

Descartes made an analogy between the pipes in the garden and the physiology of the body and brain. For example, in Figure 6, Descartes drew the brain as a complex system of hydraulic pipes that were connected in a network of cause-and-effect reflexes. In the garden, water pushed through a pipe could cause a statue to move at the other end. Likewise, in the body and brain, liquids pushed through the nervous system would cause movements, coming in the form of behaviors and reflexes.

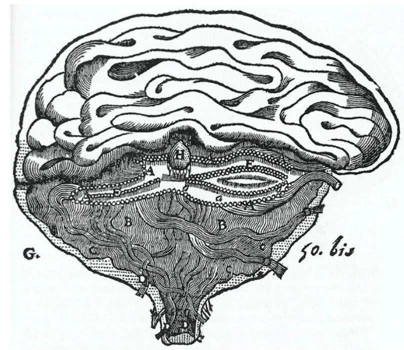


Figure 6: The brain as a complicated system of pipes.

0.0.4.2 Pavlov's liquid

Pavlov considered the possibility that the brain could be an extremely complicated system of reflexes, with many input pathways connected to output pathways. Of Descartes' idea he wrote,

> "Our starting point has been Descartes' idea of the nervous reflex. This is a genuine scientific conception, since it implies necessity. It may be summed up as follows: An external or internal stimulus falls on some one or other nervous receptor and gives reise to a nervous impulse; this nervous impulse is transmitted along nerve fibres to the central nervous system, and here, on account of existing nervous connections, it gives rise to a fresh impulse which passes along outgoing nerve fibres to the active organ, where it excites a special activity of the cellular structures. Thus, a stimulus appears to be connected of necessity with a definite response, as cause with effect. It seems obvious that the whole activity of the organism should conform to definite laws."

Pavlov was interested in discovering the so-called laws of stimulus-response pathways in the brain that he envisioned would ultimately govern human and animal behavior. In contrast to Thorndike's measurements of behavior, Pavlov was focused on physiological measurements. He observed that organs near the brain were involved in secreting various liquids. For example, when a dog smells food, it may begin to salivate. The stimulus sensation of smelling food caused a cascade of events– like water moving through a complicated system of pipes– culminating in a salivation response. In his laboratory, Pavlov was studying saliva responses in dogs when he noticed these stimulus-response pathways were not as simple as hard-wired reflexes. Instead, Pavlov discovered what is now termed “classical conditioning”, a form of association learning for creating new stimulus-response pathways.

The next section reviews a very small subset of classical conditioning phenomena: simple acquisition, extinction, and spontaneous recovery. It is worth noting that Pavlov’s results inspired a much larger discipline centered on associative learning phenomena that is beyond the scope of this book.

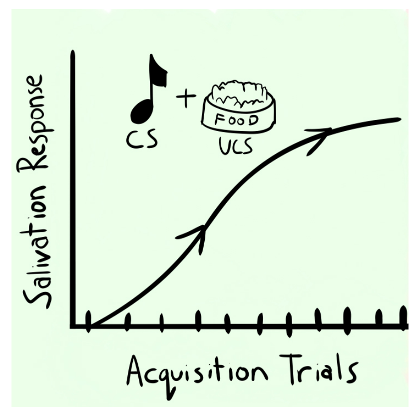


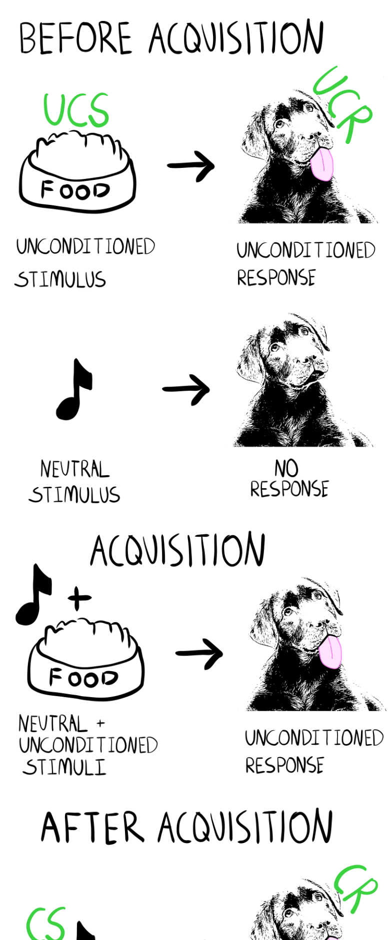
Figure 7: A simple acquisition procedure

0.0.4.3 Simple acquisition and conditioning terminology

In Pavlov’s simple acquisition procedure (shown in Figure 7), a dog was housed in a controlled laboratory setting and underwent numerous “acquisition trials”. On each trial, the dog was presented with a perceptual stimulus, like a loud tone, followed by a reward, like some meat powder. The meat powder was a stimulus that normally caused the dog to salivate. Pavlov discovered that, over the course of acquisition trials, the dog would begin to salivate in response to the loud tone, which was consistently paired with the food reward.

Simple acquisition is an example of a classical conditioning phenomenon that results from the systematic pairing of particular kinds of stimuli and their responses. Before discussing other examples, let’s define terms commonly used to describe stimuli and responses in these procedures: the unconditioned stimulus (UCS), the unconditioned response (UCR), the conditioned stimulus (CS) and the conditioned response (CR).

An unconditioned stimulus (UCS) evokes a response without prior learning from an experiment. For example, a food stimulus, such as pellets or meat powder, can initiate a saliva response in animals, and serve as an effective unconditioned stimulus even if the animal hasn’t had experience with the particu-



lar food source. Importantly, an animal like a dog does not rely on any learning during the experiment to acquire the ability to salivate in response to food. Dogs arrive to the experiment with the ability to salivate in response to the smell of food.

An unconditioned response (UCR) is an "automatic" or default response to an unconditioned stimulus. For example, salivating is an unconditioned response to smelling food. Another example of an unconditioned stimulus is a small puff of air delivered to an eyelid. In this case, a common unconditioned response is to blink: the air puff is the UCS, and blinking is the UCR.

A conditioned stimulus (CS) begins as a neutral perceptual stimulus that does not evoke the unconditioned response. For example, a loud tone or a bright light could become a conditioned stimulus. Importantly, before any acquisition trials, a loud tone would be considered a neutral stimulus because it would not evoke an unconditioned response like salivation.

During the acquisition trials the neutral stimulus is paired with the unconditioned stimulus. For example, the loud tone (neutral stimulus) is paired is paired with food (UCS). Across the acquisition trials the neutral stimulus (the tone) becomes the conditioned stimulus (CS) if it successfully begins to trigger the salivation response.

The conditioned response (CR) is the newly learned response evoked by the conditioned stimulus. For example, after a dog has acquired the association between hearing a tone and getting a food reward, the dog will begin to salivate in response to the tone. In this case, the conditioned stimulus (the tone) now evokes a conditioned response (salivating).

0.0.4.3.1 Explaining simple acquisition

In a simple acquisition procedure, the animal appears to learn a new association between a stimulus and response. However, even this most basic phenomenon of simple acquisition is not so simple to explain. What kind of association was learned during simple acquisition?

Consider some possibilities. Perhaps the tone made the animal expect to receive food, and it was the the expectation for food that evoked salivation. Perhaps the neutral stimulus evokes a vivid mental image of eating delicious food, and mental simulation of eating delicious food causes the salivation response. This explanation is complicated because it invokes mental simulation, which itself is not well-understood, as part of the explanation. As an alternative, the neutral stimulus may become directly associated with the salivation response. In this case, the neutral stimulus may not need to cause intermediate mental images of eating delicious food, but can somehow initiate the salivation through a direct connection. It's also possible that both kinds of associations were learned and involved in producing the response in different ways.

0.0.4.4 Extinction

Can associations be unlearned? Pavlov addressed this question using an extinction procedure. Figure 9 shows an acquisition phase followed by an extinction phase. In the acquisition phase, the tone (CS) is paired with food (UCS), and over trials, the dog begins to salivate when it hears the tone. The acquisition phase verifies that a new stimulus-response association was learned. The next extinction phase attempts to erase the learning from tehq acquisition phase.

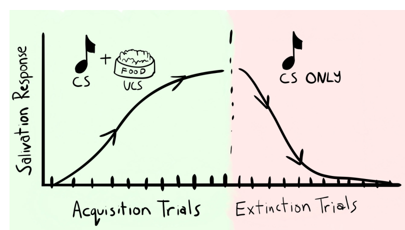


Figure 9: Acquisition followed by extinction.

In the extinction phase, the tone (CS) is presented alone without any reward. At the beginning of the extinction phase the dog shows the salivation response (UCR) when it hears the tone (CS). Throughout the extinction phase the dog will hear the tone many times, but the tone will not be paired with food reward.

The phenomenon of extinction occurs when the dog reduces or ceases to salivate in response to the tone. In other words, extinction is a reduction in the conditioned response (salivation) to the conditioned stimulus (tone).

0.0.4.4.1 Explaining extinction

The phenomena of extinction is very well-established. Animals who have acquired a new conditioned response to a conditioned stimulus will reliably show a reduction in the conditioned response following extinction training. Although the phenomena is well-established, the possible explanations of extinction are not as straightforward. For example, if an association was learned during acquisition, what happened to the association during extinction?

Was the original association somehow weakened? If the associative bond between a conditioned stimulus and response was weakened, this could explain why animals show reduced conditioned responses across extinction trials.

Perhaps the original association was not weakened, but the animal learned a brand new association during the extinction phase. For example, during acquisition the animal could learn that the conditioned stimulus predicts the occurrence of food reward. Likewise, during the extinction phase the animal could learn that the conditioned stimulus predicts the absence of food reward. If both of the associations are learned, they may compete with one another to control whether the animal responds by salivating.

0.0.4.5 Spontaneous Recovery

This chapter ends with Pavlov's finding of spontaneous recovery. This finding helps demonstrate how additional empirical evidence can be useful in evaluating theoretical explanations of associative learning.

Consider the idea that acquisition training causes an associative bond to form between a conditioned stimulus and response. The bond could be like a string connecting the stimulus to the response. On this view, the learning process establishes a string-like connection between the conditioned stimulus and response. When the stimulus appears, it "pulls" out the connected response.

Pavlov discovered several other conditioning phenomena beyond the scope of this chapter. You can read the original translation of his lectures on conditioned reflexes on archive.org

What happens to the string-like associative bond during the extinction phase? One possibility is that extinction causes the associative bond to deteriorate, weaken, and disconnect. In the metaphor, if extinction learning was completely successful in breaking the connection, then the previous association would be completely gone because the associative strands would be erased. As a result, after complete extinction learning, the conditioned stimulus would never again evoke the conditioned response.

Pavlov's discovery of spontaneous recovery was a clue that extinction learning was more complicated that snipping a string. Spontaneous recovery refers to the phenomenon that an extinguished conditioned stimulus can sometimes show a spontaneous recovery and evoke the conditioned response at a later time. For example, a dog who had learned to salivate in response to a tone, and then received extinction training, would no longer salivate in response to the tone. Spontaneous recovery occurs when, sometime later, the dog appears to "randomly" or "spontaneously" begin salivating in response to the tone again, even though that response was apparently extinguished.

According to the string metaphor of associative bonds, spontaneous recovery should not be possible because the connection between the conditioned stimulus and response should have been broken. Instead, spontaneous recovery invites alternative explanations.

One possibility is that extinction does deteriorate an existing associative bond, but never completely, so there is always a small remaining connection. As a result, the remaining connection, even if it is highly deteriorated, may still be enough to support the occasional learned response.

Another possibility is that extinction is caused by a separate form of learning that does not weaken or disconnect association formed during acquisition. For example, extinction learning could involve learning to suppress the conditioned response. On this view, a conditioned stimulus could first acquire an association that triggers a conditioned response, and then during extinction acquire another separate association that suppresses the conditioned response. After the extinction phase is over, it is possible that the suppressive response deteriorates over time, allowing the original learned association to spontaneously recover.

A third possibility is that learning is highly context-sensitive. As a result, extinction of a conditioned response may occur more strongly in the environment where the extinction training occurred, and the spontaneous recovery of the response may be more likely in other environments not associated with the extinction training.

0.0.5 Conclusions

This chapter introduced associationist ideas about cognition through early philosophy and early experimental psychology work on animals by Thorndike and Pavlov.

The associationist philosophers were developing early process theories of how cognition works. A process theory includes a recipe of how individual parts interact together to produce some output. For associationists, the big claim was that cognition involved associative learning processes. These learning processes were assumed to create knowledge about the world through the sensory process of experiencing events and objects in the world.

Early experimental psychologists like Thorndike and Pavlov gave additional credence to the concept of associations as an acceptable unit of scientific study. Their laboratory methods, especially Pavlov's, inspired whole branches and schools of psychology interested in associative learning processes, some of which continued to the modern day. Throughout this textbook we occasionally return to associative learning research, which has succeeded in identifying numerous empirical phenomena and creating detailed mathematical process models of the association formation process; all of which are relevant to cognition–especially, if, as the associationists claimed, cognition is fundamentally about learning associations.

The next two chapters visit major perspectives that lead into the emergence of modern cognitive psychology in the 1960s and 70s. Chapter 6 covers the school of behaviorism, and Chapter 7 covers the introduction of information theory to psychology.

0.0.6 Appendix

0.0.6.1 References

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, *11*(41), 63–65. https://doi.org/b6fr5r

CATTELL, J. M. (1886). THE TIME TAKEN UP BY CEREBRAL OPERATIONS. *Mind*, *os-XI*(42), 220–242. https://doi.org/c5z6fm

Cattell, J. M. (1887). Experiments on the association of ideas. *Mind*, *12*(45), 68–74. https://doi.org/d9bv2m

Darwin, C. (1859). *On the origin of species: By means of natural selection*. John Murray, Albemarle Street. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=D847cGsAAAAJ&citation_for_view=D847cGsAAAAJ:n8FNryW2AHIC

Fearing, F. (1929). René Descartes. A study in the history of the theories of reflex action. *Psychological Review*, *36*(5), 375–388. https://doi.org/cwgz2m

Goltz, F. (1892). Der hund ohne grosshirn (The dog without a cerebrum). *Archiv Für Die Gesamte Physiologie Des Menschen Und Der Tiere*, *51*(11), 570–614.

Hintzman, D. L. (2003). Robert Hooke's model of memory. *Psychonomic Bulletin & Review*, *10*(1), 3–14. https://doi.org/fd9wxx

Hooke, R. (1705). *The posthumous works of Robert Hooke: With a new introduction by Richard S. Westfall*. New York: Johnson Reprint Corp.

Hume, D. (1896). *A treatise of human nature (1793)*. Clarendon Press.

Locke, J. (1847). *An essay concerning human understanding (1690)*. Kay & Troutman.

Nicolas, S., Coubart, A., & Lubart, T. (2014). The program of individual psychology (1895-1896) by Alfred Binet and Victor Henri. *LAnnee Psychologique*, *Vol. 114*(1), 5–60. https://www.cairn.info/revue-l-annee-psychologique1-2014-1-page-5.htm

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (A. G. V., Ed.). London: Oxford University Press. http://books.google.com/books?hl=en&lr=&id=cknrYDqAClkC&oi=fnd&pg=PA1&dq=pavlov&ots=KAnkq8Z5Jb&sig=9io8iEUzHMM3wDB0J7cMjT-bEcY

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimen-*

tal Psychology, *81*(2), 275–280. https://doi.org/10.1037/h0027768

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, *2*(4), i–109. https://doi.org/bk48z2

Vaccari, A., & Philosophy Documentation Center. (2012). Dissolving Nature: How Descartes Made Us Posthuman. *Techné: Research in Philosophy and Technology*, *16*(2), 138–186. https://doi.org/fz97nk