Making Sense of Animal Conditioning

Frances K. McSweeney

Abstract

Operant and classical conditioning provide powerful techniques for understanding and controlling animal behavior. In classical conditioning, behavior changes when an arbitrary stimulus predicts the occurrence of an important stimulus. The animal's behavior towards the arbitrary stimulus changes as a result. In operant conditioning, the frequency of a response is changed by consequences that follow that response. This chapter briefly summarizes some of the characteristics of behavior undergoing conditioning. Topics include: the basic conditioning procedures, sign-tracking, classical conditioning with drug stimuli, the definition of a reinforcer, shaping, differences between reinforcement and punishment, schedules of reinforcement, acquisition, extinction, generalization, discrimination, higher order conditioning, and schedule-induced behavior.

Introduction

Classical and operant conditioning provide two powerful techniques for understanding and controlling animal behavior. In classical conditioning, behavior towards an arbitrary stimulus changes when that stimulus predicts that an important stimulus will occur. In operant conditioning, a response is followed by a consequence (e.g., a reinforcer or punisher) and the response increases or decreases in frequency as a result. This chapter briefly examines the basic conditioning procedures and some of the characteristics of behavior undergoing conditioning.

Classical Conditioning

The discovery of classical conditioning is usually attributed to Ivan Pavlov (1927). Pavlov briefly turned on a metronome and then presented food to a dog. After a few pairings of the metronome with food, the dog salivated when the metronome was presented alone. This procedure is often described by stating that when an arbitrary stimulus (the metronome, called a "conditioned stimulus" or CS) is followed by an important stimulus (food, called an "unconditioned stimulus" or US), a part of the response that is elicited by the US (e.g., salivation, called the "unconditioned response" or UR) is evoked by the CS. The response that occurs when the CS is presented alone (e.g., salivation) is called the "conditioned response" or CR (Figure 1).

Classical conditioning as studied by Pavlov is of some practical interest. For example, fears or phobias may be learned when a stimulus (e.g., a snake) precedes a frightening event (e.g., someone screams; Watson and Rayner 1920). Classical conditioning may facilitate digestion because stimuli that predict food may help to prepare the body for digestion of that food (Woods and Strubbe 1994). Classical conditioning is also thought to play a role in the development of learned preferences for and aversions to foods (Garcia and Koelling 1966). In the case of flavor aversions (Launchbaugh et al. this volume) the flavor of the plant is the CS and plant allelochemical(s) is the US which elicits illness (the UR) resulting in the future avoidance of the plant (CR). Therefore, it may play a role in understanding the feeding patterns of livestock and wildlife.

However, other aspects of Pavlov's procedure reduce the practical usefulness of classical conditioning. For example, Pavlov studied reflexive responses (e.g., salivation) while you may be more interested in "voluntary" behaviors (e.g., coming when called). He also studied salivation while his animals were immobilized by suspending them in a hammock; a practice of little relevance to understanding the behavior of free-ranging animals. In Pavlov's experiment, the same response served as the CR and the UR. That is, dogs salivated when food was presented (the UR) and they learned to salivate to the metronome that predicted food (the CR). If the CR must be identical to the UR, then the domain of classical conditioning is limited. For example, you could only use classical conditioning to train a response if you could find a US that automatically elicited that response. In many cases, this may be impossible. Luckily, these assumptions about classical conditioning are incorrect. In fact, classical conditioning probably plays a larger role in the behavior of free-ranging animals than is commonly assumed. (For more information, see Rescorla 1988).

Frances.K. McSweeney is Edward R. Myer Professor of Psychology, Department of Psychology, Washington State University, Pullman, WA, 99164-4820.

Presented in "Grazing Behavior of Livestock and Wildlife." 1999. Idaho Forest, Wildlife & Range Exp. Sta. Bull. #70, Univ. of Idaho, Moscow, ID. Editors: K.L. Launchbaugh, K.D. Sanders, J.C. Mosley.

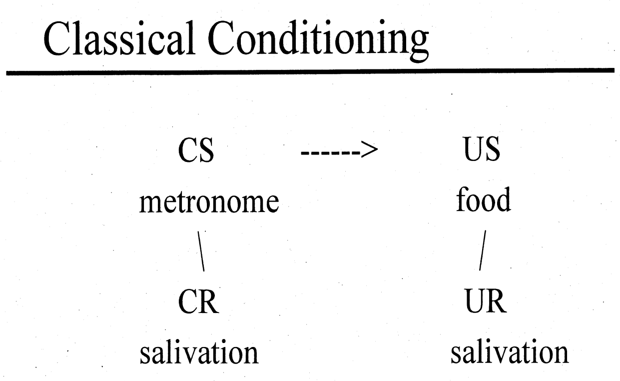

Figure 1. The basic classical conditioning procedure. A conditioned stimulus (CS) is followed by an unconditioned stimulus (US) and the CS acquires the ability to evoke the response (CR) that was formerly emitted to the US alone (UR).

Sign-tracking

Hearst and Jenkins (1974) formulated a principle that they called "sign-tracking". Sign-tracking states that, "Animals approach and contact the best predictor of reinforcers and they withdraw from stimuli that signal the absence of reinforcement." Notice that sign-tracking and the understanding of classical conditioning given earlier both describe how behavior changes when an arbitrary stimulus predicts an important stimulus. However, the two formulations differ in several ways.

According to Hearst and Jenkins, the behavior that is learned is movement in the environment (approach or withdrawal), not a reflexive response (e.g., salivation). The biologically important stimulus (US) is identified as a reinforcer, a term that will be defined later. Behavior also changes when the arbitrary stimulus (CS) predicts the reinforcer, not when the CS is followed by the reinforcer. To date, no generally-accepted definition of "predict" has been offered. However, you will understand Hearst and Jenkin's argument if you understand that prediction is a looser relation between the CS and US than temporal following. For example, the sight of clouds may predict rain even though you rarely get rained on immediately after you see a cloud. These differences make sign-tracking more useful in practice than the traditional view of classical conditioning. For example, I almost paid a heavy price once for underestimating the power of sign-tracking. I was visiting a wildlife park in Australia where a vending machine sold kangaroo chow. Unfortunately, the machine made a loud noise when it operated and that sound (CS) predicted the availability of food (US). As sign-tracking would predict, the kangaroos ran towards the food machine as soon as they heard it operating, an undesirable event for those standing by the machine.

Drugs as USs

Although Pavlov measured the same response as his CR and UR, we now know that these responses need not be identical. Sign-tracking provides one example of the CR (approach) differing from the UR (whatever is evoked by the US, e.g., salivation). The study of drugs as USs provides another example. In this case, the CR may be the opposite of the UR. To give one example, Siegel (1977) used morphine as a US. He showed that an arbitrary stimulus (e.g., a light or tone) that was followed by a morphine injection eventually evoked a CR that was opposite to the UR evoked by the morphine itself. For example, morphine is a pain killer (the UR). In contrast, animals become hypersensitive to pain during a CS that predicts morphine (the CR).

Siegel went on to argue that classical conditioning may contribute to the build up of tolerance for drugs and to the withdrawal symptoms that are observed when drugs are not delivered. This can be more easily understood if we describe the UR to morphine as a "high" (a pleasant state) and the CR to morphine as a "low" (an unpleasant state). As will be discussed (see Acquisition), classically conditioned responses gradually become stronger with each successive pairing of the CS and US. If a conditioned "low" becomes stronger with each successive morphine injection, then more and more of the drug will be required to overcome this "low" and produce the desired high. This is known as developing tolerance. If the CSs that accompany a drug injection (e.g., time of day, sight of the needle) occur without the drug, then the animal will experience only the CR (ie., a low) without the high produced by the US. This low will contribute to withdrawal symptoms.

These findings have several implications for people who deliver drugs to animals. First, if tolerance has developed to a drug, be careful not to give that drug unless the stimuli that usually predict a drug injection (e.g., time of day, method of injection) are also present. Those CSs help to prepare the animal's body to deal with the assault of the drug. That is, they send the body into a state opposite to that produced by the drug. As a result, the drug is less disruptive when it is delivered. Siegel showed that a dose of drug to which an animal has developed tolerance may kill the animal if it is delivered in the absence of the protection provided by its usual CSs.

This tolerance, attributed to stimuli that foreshadows the administration of a drug, may also be relevant to animals exposed to poisonous plants. The flavor, odor, or sight of the toxic plant may serve as CS's that activate detoxification systems or signal metabolic tolerance mechanisms in the animal. This may in part explain why animals can often increase consumption of toxic plants without apparent deleterious effects.

Operant Conditioning

Operant conditioning refers to the fact that behavior changes as a result of its consequences (Figure 2). B. F. Skinner is the most famous student of operant conditioning (Skinner 1938). Because of the power of operant techniques, they form the basis for a multimillion dollar business devoted to training animals for performances in movies, at fairs, etc. Operant techniques are also used to correct animal behavior problems (e.g., for pets or farm animals). Finally, the techniques are used to answer questions of importance to those interested in animal welfare (Foster et al. 1997). For example, they can help to determine what animals "like" and "dislike".

Operant Conditioning

- Reinforcement
  S : R ----> Reinforcer ; R increases
- Punishment
  S: R ----> Punisher ; R decreases

Figure 2. Basic attributes of the operant conditioning procedure, a particular stimulus (S) response (R) pair, the frequency of the response will either increase or decrease depending on whether the events or condition following the response are positve (reinforcer) or negative (punisher).

Positive reinforcement

The principle of positive reinforcement states that a response that is followed by a reinforcer will increase in frequency (Figure 2). Notice that you cannot reinforce a response unless you can identify a reinforcer. Over the years, many definitions for the term "reinforcer" have been tried and all have failed. For example, reinforcers have been defined as substances that are physiologically needed (e.g., food, water), but there are many reinforcers that are not physiologically needed (e.g., watching television, going to the movies). Reinforcers have been defined as stimuli that reduce tension (e.g., sexual behavior), but in many cases, stimuli that increase tension also serve as reinforcers (e.g., watching a scary movie, riding a roller coaster).

Because of these failures, a reinforcer is technically defined as any stimulus that increases the frequency of a response that it follows. This is an undesirable definition because it makes the principle of positive reinforcement circular. That is, the principle now reads, a response that is followed by any stimulus that increases the frequency of a response that it follows will increase in frequency. We can live with this definition because we can identify a stimulus as a reinforcer in one situation (e.g., by showing that it increases the frequency of one response that it follows) and then test the principal of positive reinforcement in another situation (e.g., ask whether that reinforcer will also increase the frequency of other responses).

In practice, many stimuli will serve as reinforcers for nonhuman animals (e.g., food, water, petting, access to conspecifics for herd animals). Others will be useful with humans (praise, money, the opportunity to watch television). If you are having difficulty identifying a reinforcer, try the Premack Principle (e.g., Premack, 1959). Premack argued that the opportunity to perform any high probability response would reinforce any low probability response. The probability of a response was measured by examining what the animal would do when it had free time. Therefore, you can find a reinforcer by observing what an animal does often and using access to that behavior as a reinforcer. According to Premack, if a child reads more than he watches television, then reading will serve as a reinforcer for television watching if, for some reason, you wanted him to watch more TV.

Shaping

You may have noticed that you cannot reinforce a response until that response occurs. Shaping by successive approximations is a procedure that can be used to produce a response so that you can reinforce it. During shaping, you reinforce closer and closer approximations to the desired response. For example, if you want to teach your dog to sit up, you could begin by following any movement by a reinforcer. Then you might reinforce only movements that involved some transfer of the dog's weight to its back paws. Then you might reinforce only movements that involved weight transfer to the back paws plus lifting the forepaws off the ground. By judiciously chosing which behaviors to reinforce and when to alter the reinforced response, you should quickly have your dog sitting up.

The four basic conditioning procedures

You can use operant conditioning to either increase (reinforcement) or decrease (punishment) the frequency of a response. The frequency of a response may change when the response produces something (positive) or when it escapes or avoids something (negative). It is called "positive reinforcement" when a response increases in frequency because it produces something (e.g., sheep walk into the corral when called because they receive feed). Negative reinforcement occurs when a response increases because it escapes or avoids something (e.g., a flock of sheep move into the corral to avoid getting nipped by the sheep dog). Positive punishment occurs when a response decreases in frequency because it produces something (e.g., a cow stops touching the electric fence with her nose because she gets shocked). Negative punishment occurs when a response decreases in frequency because it prevents something that would otherwise occur (e.g., your horse stands still after a ride because moving about delays the removal of the saddle and bridle).

Large organizations (e.g., governments, armed services, universities) control your behavior largely through negative reinforcement. For example, you probably pay your income taxes on time to avoid a fine (negative reinforcement) rather than because you receive a thank you note from the President (positive reinforcement). Positive reinforcers often cost money, but negative reinforcers often do not. Nevertheless, I recommend that you use positive reinforcement and negative punishment to alter behavior whenever possible. The other alternatives, negative reinforcement and positive punishment, involve the delivery of an aversive stimulus. Delivering aversive stimuli can have undesirable consequences. For example, they may elicit aggression. In contrast, positive reinforcement and negative punishment involve the delivery or withdrawal of a positive stimulus which should elicit fewer undesirable behaviors. To give only one example, if an animal is attacking other animals in a herd, a better way to decrease the frequency of attack might be to isolate the animal for a while (negative punishment) rather than to follow attack by a shock from a cattle prod (positive punishment).

Schedules of reinforcement

In a continuous reinforcement procedure (CRF), every occurrence of a response is followed by a reinforcer. CRF is rarely used because it is expensive if the reinforcer costs money. The frequent delivery of reinforcers also disrupts behavior. Therefore, CRF is used to initially teach a response but a schedule of partial reinforcement is used as the response becomes stronger.

In a partial reinforcement procedure (PRF), some instances of a response are not followed by a reinforcer. There are several schedules of PRF, but the most useful may be the fixed (FR) and variable (VR) ratio schedules. In an FR x schedule, a reinforcer is delivered after every x occurrences of a response. For example, in a piece work factory, you might be paid (a reinforcer is delivered) every time you complete 10 widgets. This would be an FR 10 schedule. In a VR x schedule, a reinforcer is delivered after every xth occurrence of the response on the average. For example, a pigeon foraging for grain does not find grain (the reinforcer) each time it pecks the ground (the response), but it does find grain after some variable number of pecks. FR and VR schedules control behavior somewhat differently. Responses occur at a high steady rate if they are reinforced on a VR schedule. In contrast, animals pause after receiving a reinforcer and then later respond at a relatively steady rate when responding on an FR schedule. The pause is longer the larger the number of responses required for reinforcement. In fact, if the ratio requirement becomes too large, the animal may stop making the response. This is called ratio strain. To avoid ratio strain, the number of responses required for a reinforcer should be increased gradually, rather than in large increments.

Characteristics of Conditioned Behavior

Acquisition

Operantly and classically conditioned responses do not appear full blown the first time they occur. Instead, they are gradually acquired as the response is repeatedly followed by the reinforcer or as the CS repeatedly predicts the US. The strength of a conditioned response usually increases as a negatively-accelerated function of experience with the CS-US or reinforcer- response relation (Figure 3).

Extinction

Extinction means that a response that has been classically conditioned will return to its baseline frequency if the relation between the CS and US is broken. This may be done in either of two ways. The US may be removed entirely or the CS and US may be presented randomly with respect to each other. A response that has been operantly conditioned also returns to its baseline frequency if the relation between the response and the reinforcer is broken. Again, this relation may be

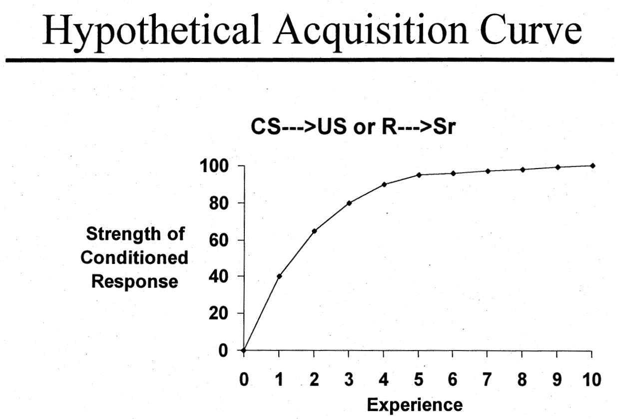

Figure 3. A hypothetical acquisition curve. The strength of a conditioned response increases as a negatively accelerated function of experience with the CS-US or response-reinforcer relation.

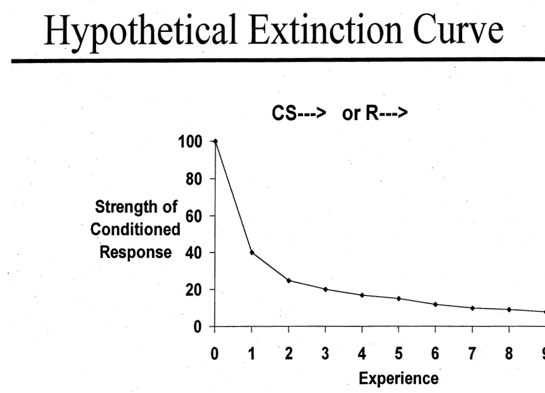

Figure 4. A hypothetical extinction curve. The strength of a conditioned response decreases with experience that the CS no longer predicts the US or the response no longer produces the reinforcer.

broken by removing the reinforcer entirely or by presenting the reinforcer randomly with respect to the response. For example, a deer may return to a specific location in its home range to eat a relished plant (reinforcer). However, the deer will return to this place less often if the plant is removed (the reinforcer was removed) or the plant begins to appear randomly throughout its home range (the reinforcer is presented randomly with respect to location).

A hypothetical extinction curve appears in Figure 4. Theoretically extinction, as punishment, can be used to decrease the frequency of an undesirable response. However, its use in practice may be limited. You can only extinguish behavior that has been conditioned. You can only use extinction if you can identify all of the reinforcers that support the undesirable behavior and can control delivery of those reinforcers. Most behaviors are partially rather than continuously reinforced and extinction is slower for partially than for continuously reinforced behavior. Responses undergoing extinction may also increase in frequency for a brief time at the start of extinction, an undesirable consequence if you're trying to eliminate the response.

Generalization

Generalization refers to the fact that a CR that occurs to one CS will also occur to other stimuli that resemble the CS that was originally paired with the US. The greater the resemblance between the new stimulus and the original CS, the stronger the conditioned response to the new stimulus. For example, if you're stung (US) by a bee (CS), you may learn to fear (CR) other flying insects and your fear will be stronger the more closely the insect resembles a bee. A hypothetical generalization gradient appears in Figure 5.

A response that has been reinforced in the presence of one stimulus will also occur in the presence of other stimuli that resemble the original stimulus. Again, the stronger the resemblance between the new stimulus and the original one, the stronger the response to the new stimulus. For example, a deer may learn to limit intake of big sagebrush because it contains essential oils which have several deleterious digestive consequences. If the deer encounters a new species of sagebrush, such as three-tip sage, it may avoid eating it. The deer may generalize its avoidance of big sage to the newly enountered sage because they both contain similar essential oils which give them a similar odor and taste.

Discrimination

During a classical conditioning discrimination procedure, a stimulus is followed by a US (CS+) and another stimulus is not followed by a US (CS-). The CR will occur to CS+ but not to CS-. During an operant discrimination procedure, a response is reinforced in the presence of one stimulus (S+) and not in the presence of another stimulus (S-). The response will occur in the presence of S+, but not in the presence of S-. For

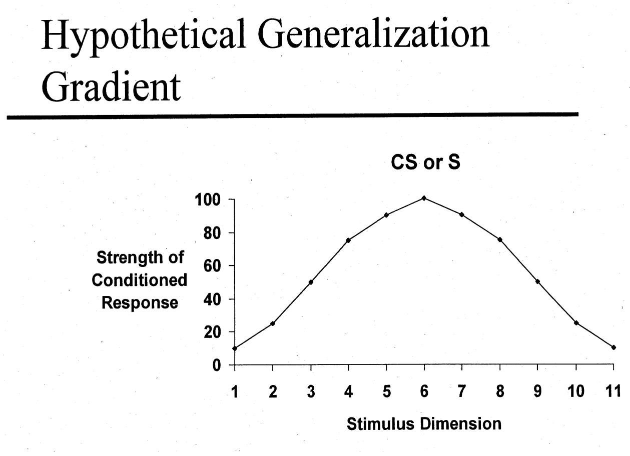

Figure 5. A hypothetical generalization gradient. The strength of a conditioned response decreases as stimuli become more dissimilar to the stimulus that was actually involved in classical or operant conditioning. The x-axis is a stimulus dimension (e.g., the brightness of a light, the loudness of a tone). The stimulus that was present during conditioning appears in position 6.

example, a herd of cows may learn that running to a vehicle (a response) results in getting feed (a reinforcer). They may further learn that feed only comes from the red feed truck when its horn is blaring (S+), but not from other pickups that drive through the pasture(S-).

Discrimination procedures provide a useful technique for asking questions of non-human animals or nonverbal people (e.g., infants). You may have heard that dogs do not "see colors". How do we know? Part of the answer comes from discrimination training. Suppose you reinforce sitting up by giving the dog a treat in the presence of anything red, but not in the presence of anything green. If the dog can see colors, then you will quickly have a dog that sits up when a red, but not a green, stimulus is presented. When this experiment is done properly, dogs do not develop a discrimination.

Higher-order Conditioning

Some stimuli serve as USs or reinforcers from birth with no additional training. These stimuli are called primary reinforcers or USs. They include biologically important stimuli, such as food and water. Other stimuli acquire their ability to act as reinforcers or USs through experience. These stimuli are called secondary, or higher-order, reinforcers or USs. Money provides the most obvious example of a secondary reinforcer.

Stimuli acquire the ability to act as secondary reinforcers in many ways, two of which will be described. First, stimuli that can be exchanged for primary reinforcers will act as secondary reinforcers. Such stimuli are called "tokens". For example, money acquires the ability to act as a reinforcer because it can be exchanged for food, drink and other primary reinforcers. Second, classical conditioning pairing of a stimulus with primary USs or reinforcers will produce a secondary reinforcer or US. Therefore, a bell that is used to summon animals for feeding will gain the ability to act as a reinforcer itself.

The ability of these stimuli to act as secondary reinforcers or USs will extinguish if their relation to the primary reinforcer or US is broken. Therefore, money would gradually lose its ability to reinforce if it was no longer exchangeable for goods and the bell would lose its ability to reinforce if it was presented often without food.

Schedule-induced Behavior

A final oddity of behavior undergoing reinforcement will be mentioned because you may sometimes encounter it. Falk (1971) gave hungry rats food (a reinforcer) when they pressed a lever (a response). In this experiment, food was delivered once every minute on average. When water was also available, Falk noticed that rats drank approximately 50% of their body weight in water over the course of a two hour experimental session. He called this behavior polydipsia (much drinking) and observed that it was counterproductive because the rat was wasting calories by heating a large amount of water to its body temperature and then excreting it.

Later studies showed that animals will perform many other behaviors in excess when reinforcers are spaced in time. These behaviors are called "adjunctive" or "schedule-induced". They include aggression, eating non-food substances (pica), running in a wheel, defecation, escape from the schedule of reinforcement and drug consumption.

Because adjunctive behaviors are excessive and often maladaptive, they have served as models for a variety of problematic behaviors. I cannot discuss these models in detail, but if you suspect that a problematic behavior is schedule-induced, one way of reducing that behavior is to identify the schedule of reinforcement that is maintaining the behavior and to change the interval between successive deliveries of the reinforcer. For example, I once received a telephone call from a woman whose dog was biting her when she fed it. One among many potential explanations for this behavior is that the bite was an aggressive response that occurred because food was spaced in time. In that case, the woman should change the interval between meals to reduce biting.

You can test yourself on the preceding material by explaining how biting could also be a classically or operantly conditioned response. What would you do to eliminate biting if you thought it was classically or operantly conditioned?

References

Falk, J. L. 1971. The nature and determinants of adjunctive behavior. Physiol. Behav. 6:577-588.

Foster, T. M., W. Temple and A. Poling. 1997. Behavior analysis and farm animal welfare. Behav. Analyst. 20:87-96.

Garcia, J. and R. A. Koelling. 1966. Relation of cue to consequences in avoidance learning. Psychon. Sci. 4:123-124.

Hearst, E. and H. M. Jenkins. 1974. Sign-tracking: The Stimulus-reinforcer Relation and Directed Action. Psychonomic Society. Austin, TX.

Pavlov, I. P. 1927. Conditioned Reflexes. Oxford University Press. London, England.

Premack, D. 1959. Toward empirical behavioral laws: 1. Positive reinforcement. Psychol. Rev. 66:219-233.

Rescorla, R. A. 1988. Pavlovian conditioning: It's not what you think it is. Amer. Psychol. 43:151-160.

Siegel, S. 1977. Morphine tolerance acquisition as an associative process. J. Exp. Psychol: Anim. Behav. Proc. 3:1-13.

Skinner, B. F. 1938. The Behavior of Organisms. Appleton-Century-Crofts. New York.

Watson, J. B. and R. Rayner. 1920. Conditioned emotional reactions. J. Exp. Child Psychol. 3:1-14.

Woods, S. C. and J. H. Strubbe. 1994. The psychobiology of meals. Psychon. Bull. Rev. 1:141-155.