

OPERANT CONDITIONING

How does operant conditioning occur?

There are two kinds of behavior that all organisms are capable of doing: involuntary (reflexive) and voluntary. If Inez blinks her eyes because a bee flies close to them, that's a reflex and totally involuntary. But if she then swats at the bee to frighten it, that's a voluntary choice. She *had* to blink, but she *chose* to swat. Classical conditioning is the kind of learning that occurs with reflexive, involuntary behavior. The kind of learning that applies to voluntary behavior is called **operant conditioning**, which is both different from and similar to classical conditioning.

THORNDIKE'S LAW OF EFFECT

Edward L. Thorndike (1874–1949) was one of the first researchers to explore and attempt to outline the laws of learning voluntary responses, although the field was not yet called operant conditioning. Thorndike placed a hungry cat inside a “puzzle box” from which the only escape was to press a lever located on the floor of the box. Thorndike placed a dish of food *outside* the box, so the hungry cat is highly motivated to get out. Thorndike observed that the cat would move around the box, pushing and rubbing up against the walls in an effort to escape. Eventually, the cat would accidentally push the lever, opening the door. Upon escaping, the cat was fed from a dish placed just outside the box. The lever is the stimulus, the pushing of the lever is the response, and the consequence is both escape (good) and food (even better). The cat did not learn to push the lever and escape right away. After a number of trials (and many errors) in a box like this one, the cat took less and less time to push the lever that would open the door. It's important not to assume that the cat had “figured out” the connection between the lever and freedom—Thorndike kept moving the lever to a different position, and the cat had to learn the whole process over again. The cat would simply continue to rub and push in the same general area that led to food and freedom the last time, each time getting out and fed a little more quickly. Based on this research, Thorndike developed the **law of effect**: If an action is followed by a pleasurable consequence, it will tend to be repeated. If an action is followed by an unpleasant consequence, it will tend not to be repeated (Thorndike, 1911). This is the basic principle behind learning voluntary behavior. In the case of the cat in the box, pushing the lever was followed by a pleasurable consequence (getting out and getting fed), so pushing the lever became a repeated response.

So did Thorndike call this operant conditioning?

No, but Thorndike's important and groundbreaking work began the study of what would eventually become **operant conditioning**.

B. F. SKINNER: THE BEHAVIORIST'S BEHAVIORIST

B. F. Skinner (1904–1990) was the behaviorist who assumed leadership of the field after John Watson. He was even more determined than Watson that psychologists should study only measurable, observable behavior. In addition to his knowledge of Pavlovian classical conditioning, Skinner found in the work of Thorndike a way to explain all behavior as the product of learning. He even gave the learning of voluntary behavior a special name: *operant conditioning* (Skinner, 1938). Voluntary behavior is what people and animals do to *operate* in the world. When people perform a voluntary action, it is to get something they want or to avoid something they don't want. So voluntary behavior, for Skinner, is **operant** behavior, and the learning of such behavior is

operant conditioning. The heart of operant conditioning is the effect of consequences on behavior. Thinking back to the section on classical conditioning, learning a reflex really depends on what comes *before* the response—the unconditioned stimulus and what will become the conditioned stimulus. These two stimuli are the *antecedent* stimuli (antecedent means something that comes before another thing). But in operant conditioning, learning depends on what happens *after* the response—the consequence. In a way, operant conditioning could be summed up as this: “If I do this, what’s in it for me?”

Important concepts in operant conditioning

“What’s in it for me?” represents the concept of **reinforcement**, one of Skinner’s major contributions to behaviorism. The word itself means “to strengthen,” and Skinner defined reinforcement as anything that, when following a response, causes that response to be more likely to happen again. Typically, this means that reinforcement is a consequence that is in some way pleasurable to the organism, which relates back to Thorndike’s law of effect. The “pleasurable consequence” is what’s “in it” for the organism. (Keep in mind that a pleasurable consequence might be something like getting food when hungry or a paycheck when you need money, but it might also mean *avoiding* a tiresome chore, like doing the dishes or taking out the garbage. I’ll do almost anything to get out of doing the dishes, myself!) Going back to Thorndike’s puzzle-box research, what was in it for the cat? We can see that the escape from the box and the food that the cat received after getting out are both *reinforcement* of the lever-pushing response. Every time the cat got out of the box, it got reinforced for doing so. In Skinner’s view, this reinforcement is the reason that the cat learned anything at all. In operant conditioning, reinforcement is the key to learning. Skinner had his own research device called a “Skinner box” or “operant conditioning chamber”. His early research often involved placing a rat into one of these chambers and training it to push down on a bar to get food.

PRIMARY AND SECONDARY REINFORCERS

The events or items that can be used to reinforce behavior are not all alike. Let’s say that a friend of yours asks you to help her move some books from the trunk of her car to her apartment on the second floor. She offers you a choice of 2500 rupees or a candy bar. With this money, you could buy more than one candy bar. Now pretend that your friend offers the same deal to a 3-year-old child who lives downstairs for carrying up some of the paperback books: 2500 rupees or a candy bar. Which reward will the child more likely choose? Most children at that age have no real idea of the value of money, so the child will probably choose the candy bar. The money and the candy bar represent two basic kinds of **reinforcers**, items or events that when following a response will strengthen it. The reinforcing properties of money must be learned, but candy gives immediate reward in the form of taste and satisfying hunger. A reinforcer such as a candy bar that fulfills a basic need like hunger is called a **primary reinforcer**. Examples would be any kind of food (hunger drive), liquid (thirst drive), or touch (pleasure drive). Infants, toddlers, preschool-age children, and animals can be easily reinforced by using primary reinforcers. (It’s not a good idea, however, to start thinking of reinforcers as rewards—freedom from pain is also a basic need, so pain itself can be a primary reinforcer when it is *removed*. Removal of a painful stimulus fills a basic need just as eating food when hungry fills the hunger need.)

A **secondary reinforcer** such as money, however, gets its reinforcing properties from being associated with primary reinforcers in the past. A child who is given money to spend soon realizes that the ugly green paper can be traded for candy and treats—primary reinforcers—and so money becomes reinforcing in and of itself.

Secondary reinforcers do indeed get their reinforcing power from the process of classical conditioning. After all, the pleasure people feel when they eat, drink, or get a back rub is an automatic response, and any automatic response can be classically conditioned to occur to a new stimulus. In the case of money, the candy is a UCS for pleasure (the UCR) and the money is present just before the candy is obtained. The money becomes a CS for pleasure. The praise, or more specifically the tone of voice, becomes the CS for pleasure. Although classical and operant conditioning often “work together,” as in the creation of secondary reinforcers, they are two different processes.

POSITIVE AND NEGATIVE REINFORCEMENT

Reinforcers can also differ in the way they are used. Most people have no trouble at all understanding that following a response with some kind of pleasurable consequence (like a reward) will lead to an increase in the likelihood of that response being repeated. This is called **positive reinforcement**, the reinforcement of a response by the *addition* or experience of a pleasurable consequence, such as a reward or a pat on the back. But many people have trouble understanding that the opposite is also true: Following a response with *the removal or escape* from something *unpleasant* will also increase the likelihood of that response being repeated—a process called **negative reinforcement**. Remember the idea that pain can be a reinforcer if it is removed? If a person’s behavior gets pain to stop, the person is much more likely to do that same thing again—which is part of the reason people can get addicted to painkilling medication.

SCHEDULES OF REINFORCEMENT

FIXED INTERVAL SCHEDULE OF REINFORCEMENT

If you receive a paycheck once a week, you are familiar with what is called a **fixed interval schedule of reinforcement**, in which a reinforcer is received *after* a certain, fixed interval of time has passed. A fixed interval schedule of reinforcement does not produce a fast rate of responding. Since it only matters that at least *one* response is made *during* the specific interval of time, speed is not that important. This is similar to the way in which factory workers speed up production just before payday and slow down just after payday (Critchfield et al., 2003). Paychecks aren’t the only kind of fixed schedule that people experience. When do you study the hardest? Isn’t it right before a test? If you know when the test is to be given, that’s like having a fixed interval of time that is predictable, and you can save your greatest studying efforts until closer to the exam. (Some students save *all* of their studying for the night before the exam, which is not the best strategy.)

VARIABLE INTERVAL SCHEDULE OF REINFORCEMENT

Students don’t know exactly what day they might be given a surprise test, so the best strategy is to study a little every night just in case there is a test the next day. Surprise tests are good examples of a **variable interval schedule of reinforcement**, where the interval of time after which the individual must respond in order to receive a reinforcer (in this case, a good grade on the test) changes from one time to the next. Once again, speed is not important, so the rate of responding is slow but steady. Another example of a variable interval schedule might be the kind of fishing in which people put the pole in the water and wait—and wait—and—wait, until a fish takes the bait, if they are lucky. They only have to put the pole in once, but they might refrain from taking it out for fear that just when they do, the biggest fish in the world would swim by.

FIXED RATIO SCHEDULE OF REINFORCEMENT

In ratio schedules, it is the number of responses that counts. In a **fixed ratio schedule of reinforcement**, the number of responses required to receive each reinforcer will always be the same number. The rate of responding is very fast, especially when compared to the fixed interval schedule on the left, and there are little “breaks” in the response pattern immediately after a reinforcer is given. Fixed schedules—both ratio and interval—are predictable, which allows rest breaks. In human terms, anyone who does piecework, in which a certain number of items have to be completed before payment is given, is reinforced on a fixed ratio schedule. Some sandwich shops use a fixed ratio schedule of reinforcement with their customers by giving out punch cards that get punched one time for each sandwich purchased. When the card has 10 punches, for example, the customer might get a free sandwich.

VARIABLE RATIO SCHEDULE OF REINFORCEMENT

A **variable ratio schedule of reinforcement** is one in which the number of responses changes from one trial to the next. It is the *unpredictability* of the variable schedule that makes the responses more or less continuous—just as in a variable interval schedule. In human terms, people who shove money into the one-armed bandit, or slot machine, are being reinforced on a variable ratio schedule of reinforcement (they hope). They put their coins in (response), but they don’t know how many times they will have to do this before reinforcement (the jackpot) comes. People who do this tend to sit there until they either win or run out of money. They don’t dare stop because the “next one” might hit that jackpot. Buying lottery tickets is much the same thing, as is any kind of gambling. People don’t know how many tickets they will have to buy, and they’re afraid that if they don’t buy the next one, that will be the ticket that would have won, so they keep buying and buying.

Regardless of the schedule of reinforcement one uses, two additional factors contribute to making reinforcement of a behavior as effective as possible. The first factor is *timing*: In general, a reinforcer should be given as immediately as possible *after* the desired behavior. Delaying reinforcement tends not to work well, especially when dealing with animals and small children. (For older children and adults who can think about future reinforcements, such as saving up one’s money to buy a highly desired item, some delayed reinforcement can work—for them, just saving the money is reinforcing as they think about their future purchase.) The second factor in effective reinforcement is to reinforce *only* the desired behavior. This should be obvious, but we all slip up at times; for example, many parents make the mistake of giving a child who has not done some chore the promised treat anyway, which completely undermines the child’s learning of that chore or task. And who hasn’t given a treat to a pet that has not really done the trick?

So I think I get reinforcement now, but what about punishment? How does punishment fit into the big picture?

How does punishment differ from reinforcement?

People experience two kinds of things as consequences in the world: things they like (food, money, candy, sex, praise, and so on) and things they don’t like (spankings, being yelled at, and experiencing any kind of pain, to name a few). In addition, people experience these two kinds of consequences in one of two ways: Either people experience them directly (such as getting money for working or getting yelled at for misbehaving) or they don’t experience them, such as losing an allowance for misbehaving or avoiding a scolding by lying about misbehavior. Getting money for working is an example of *positive reinforcement*, the reinforcement of a response by the *addition*

or experience of a *pleasurable* consequence, as mentioned earlier. That one everyone understands. But what about avoiding a penalty by turning one's income tax return in on time? That is an example of *negative reinforcement*, the reinforcement of a response by the *removal or escape* from an *unpleasant* consequence. Because the behavior (submitting the return before the deadline) results in *avoiding* an unpleasant stimulus (a penalty), the likelihood that the person will behave that way again (turn it in on time in the future) is *increased*—just as positive reinforcement will increase a behavior's likelihood. Examples are the best way to figure out the difference between these two types of reinforcement, so try to figure out which of the following examples would be positive reinforcement and which would be negative reinforcement:

1. Sandeep's father nags him to wash his car. Sandeep hates being nagged, so he washes the car so his father will stop nagging.
2. Rahul learns that talking in a funny voice gets him lots of attention from his classmates, so now he talks that way often.
3. Sahil is a server at a restaurant and always tries to smile and be pleasant because that seems to lead to bigger tips.
4. Shweta turns her report in to her teacher on the day it is due because papers get marked down a letter grade for every day they are late.

Here are the answers:

1. Sandeep is being negatively reinforced for washing his car because the nagging (unpleasant stimulus) stops when he does so.
2. Rahul is getting positive reinforcement in the form of his classmates' attention.
3. Sahil smiling and pleasantness are positively reinforced by the customers' tips.
4. Shweta is avoiding an unpleasant stimulus (the marked-down grade) by turning in her paper on time, which is an example of negative reinforcement. I'm confused—I thought taking something away was a kind of punishment?

TWO KINDS OF PUNISHMENT: People get confused because “negative” sounds like it ought to be something bad, like a kind of punishment. **Punishment** is actually the opposite of reinforcement. It is any event or stimulus that, when following a response, causes that response to be less likely to happen again. Punishment *weakens* responses, whereas reinforcement (no matter whether it is positive or negative) *strengthens* responses. There are two ways in which punishment can happen, just as there are two ways in which reinforcement can happen.

Punishment by application occurs when something unpleasant (such as a spanking, scolding, or other unpleasant stimulus) is added to the situation or *applied*. This is the kind of punishment that most people think of when they hear the word *punishment*. This is also the kind of punishment that many child development specialists strongly recommend parents avoid using with their children because it can easily escalate into abuse (Dubowitz & Bennett, 2007; Saunders & Goddard, 1998; Straus, 2000; Straus & Stewart, 1999; Straus & Yodanis, 1994; Trocme et al., 2001). A spanking might be *physically* harmless if it is only two or three swats with a hand, but if done in anger or with a belt or other instrument, it becomes abuse, both physical and emotional.

Punishment by removal, on the other hand, is the kind of punishment most often confused with negative reinforcement. In this type of punishment, behavior is punished by the removal of something pleasurable or desired after the behavior occurs. “Grounding” a teenager is removing the freedom to do what the teenager wants to do and is an example of this kind of punishment. Other examples would be placing a child in time-out (removing the attention of the others in the room), fining someone for disobeying the law (removing money), and punishing aggressive

behavior by taking away television privileges. This type of punishment is typically far more acceptable to child development specialists because it involves no physical aggression and avoids many of the problems caused by more aggressive punishments.

The confusion over the difference between negative reinforcement and punishment by removal makes it worth examining the difference just a bit more. Negative reinforcement occurs when a response is followed by the *removal* of an *unpleasant* stimulus. If something unpleasant has just gone away as a consequence of that response, wouldn't that response tend to happen again and again? If the response increases, the consequence has to be a kind of *reinforcement*. The problem is that the name sounds like it should be some kind of punishment because of the word *negative*, and that's exactly the problem that many people experience when they are trying to understand negative reinforcement. Many people get negative reinforcement mixed up with punishment by removal, in which a *pleasant* thing is removed (like having your driver's license taken away because you caused a bad accident). Because something is removed (taken away) in both cases, it's easy to think that they will both have the effect of punishment, or weakening a response. The difference between them lies in *what* is taken away: In the case of negative reinforcement, it is an *unpleasant* thing; in the case of punishment by removal, it is a *pleasant* or desirable thing.