

K-9 COP ISSUE 41 **MAGAZINE™**

FOR POLICE AND MILITARY WORKING DOG HANDLERS • \$9.95

K-9 advice & opinion

Scenario-Based TRAINING & CERTIFICATION

Inside:

K-9 healthcare

The Poison-Proof Dog:
Part Two

K-9 spotlight

A Tribute to Police K-9 Jethro: Sculptor
Creates Life-Size Bronze Monument

K-9 tracking/trailing

Tracking After the 'Other' K-9

A Simple Lure-Reward Training System

By Jerry Bradshaw

T

raining a dog to do any task requires a system by which we can apply both the consequences of reinforcement and punishment (operant conditioning) and proper associations (classical conditioning). These concepts must be put into a simple and effective training program that can generally be applied to anything we want to teach our dogs, from obedience to tracking and detection to controlled aggression. One of the things I notice in many K-9 handlers is that when training requires teaching a brand new task, they have a hard time knowing where to start, and if they start, how to move from one phase in the training program to the next to arrive at a trained behavior. In other words, they do not have a grasp of the system of training, but rather pieces of the training process.

The simple training system I will propose is what we call a lure-reward training program. This program includes three phases: **Learning Phase**, **Correction Phase**, and **Proofing Phase**. Each phase will make use of both operant conditioning and classical conditioning. The flip side of teaching behaviors is “un-teaching,” or extinction of behaviors and we will discuss that as well.

The Learning (Acquisition) Phase

In the learning phase, our goal is to both develop behaviors (in obedience the physical positions of sit, down, through operant conditioning) and to start to teach the dog associations between verbal commands and these various body positions. That association of command to body action happens through classical conditioning.

In the old days, trainers would physically manipulate dogs into their various positions, through what was called “escape training” (also referred to as negative reinforcement). That is, the trainer would force dogs into sit position by pulling up on the choke collar and pushing the rear of the dog toward the ground. Doing this has the effect of choking the dog (unpleasant) and when the dog sits, the choking is released. The dog learns to sit to turn off the choking sensation, hence why it’s called escape training. The dog works to escape the sensation. This is not only unnecessary, but prolongs the learning process. It prolongs the learning process because when you use force to teach (negative reinforcement) you can only teach one thing at a time or you risk causing a lot of confusion. You have to finish one command before you begin another. If the dog has too many options to escape the unpleasantness, it will confuse the dog. Luckily, there is another way to achieve the desired result.

Consider the following challenge: You are hired by a trainer of large sea mammals and she puts you in a situation where you must train a killer whale to jump out of the water on command. How would you do it? Clearly, you cannot physically force this animal to do anything. Escape training is not an option. So you are left to use your imagination to figure out how to encourage this animal to not only do what you want, but put it on command.



To achieve the result, you must wait for the animal to show this behavior on her own, and then immediately mark the event (call attention to the event) in a way the whale can comprehend, and then reward the animal for performing the behavior. In behavior literature, this is known as *operant conditioning* through free shaping (no real influence from the trainer). The animal operates or is at the controls, so to speak; they learn to operate the reinforcement by offering natural behaviors. In dog training, we know that all of the commands we wish to teach are naturally occurring behaviors, as a dog will offer them on his own (sit, down, stay, recall, tracking, hunting, etc.). The job of the trainer in operant conditioning is to properly time the positive reinforcement when the animal offers the appropriate behavior. Then attach a cue to that behavior so we can call it up on command; this is the basis of the lure-reward system of training.

The *primary reinforcement* is the reward itself - the food or toy we offer. The *secondary reinforcement* is usually a sound of some kind that clearly marks or calls attention to the event of the appropriate behavior. The marker is followed by the primary reinforcement. For example, in training sea animals, usually trainers will use a whistle to mark the appropriate behavior before offering the food.¹ In training dogs, we use praise (“good” or “yes”) to mark the proper behavior. If you are a clicker trainer, the “click-click” sound is the marker or secondary reinforcement. If you heel with a tug under your left arm, you are luring the dog with the tug. If you put food on a track, you are luring the dog to the human scent with the food.

If all the behaviors we wish to teach our dogs are naturally occurring, then we can take advantage of this principle of operant conditioning. Actually, things are easier with dogs because we can help them offer these natural behaviors through a process called luring. We use food as a primary teaching tool for obedience, tracking, and articles for example, because it is a powerful motivator and because once the dog eats the food, the reward disappears and is not a constant distraction. The primary reinforcement can be changed to a toy later on to bring a bit more drive and power to the work if you prefer.

Schedules of Reinforcement

The next thing we need to understand is what is referred to as the schedule of reinforcement. That is, we need to decide the ratio of reward to the correct performance of the behavior. By reinforcement schedule, we mean how often the trainer will give primary reinforcement following a correct behavioral response. Five reinforcement schedules can be employed: (1) Fixed Interval (2) Variable Interval (3) Fixed Ratio (4) Variable ratio (5) Random.

1. In this type of training, we usually begin by pairing the marker or secondary reinforcement with the reward. For dogs you would sit down with a bunch of treats, click your clicker, and feed the dog a treat over and over again. With a sea mammal, you would blow the whistle, and feed the fish over and over again. This makes the clicker (for dogs) or the whistle (for the sea mammal) meaningful, thus when the animal performs a behavior correctly, the tone of the whistle or the sound of the clicker immediately alerts the animal they did something worthy of reward and increases the likelihood that behavior will be repeated.

Fixed interval means that the primary reinforcement will recur after a fixed amount of time. It could be every 10 seconds, 30 seconds, or every five minutes. The interval could be distance, so if you are laying a hard surface track, you put the food on the track at six inch intervals all the way down the track. Intervals can be relatively high frequency or relatively low frequency. In the tracking example, the spread of the reward (distance between food pieces) will be the frequency.

Variable interval means that the primary reinforcement will recur on a varying schedule, sometimes after 10 seconds, 51 seconds, maybe after three minutes, and again after eight seconds later. To continue with the tracking example, the interval at the start of the track could be every six inches, then a section of the track could increase to every two feet, then up to three feet for a short period, and then back to six inches toward the end.

Fixed ratio means that a behavior performed correctly “n” times, will be given one primary reinforcer on the “nth” time. So a 1:5 fixed ratio means that every fifth properly performed behavior will be given primary reinforcement. When we are luring the dog, putting food on his nose and luring him into a sit or down, or putting food on every footstep of a grass track, the ratio is 1:1. Every correct performance gets primary reinforcement. Suppose we try a fixed ratio of 3:1. How would that eventually be interpreted by the animal? He would learn the first two are never rewarded and lead to poor performance on the first two attempts at the command. The fixed ratio is therefore generally used most in the 1:1 form when teaching a dog to learn a new behavior.

Variable ratio means that primary reinforcement is given on an average number of correct responses. Thus, a variable ratio of 1:2 means that on average, one out of every two correct responses will be given primary reinforcement. It could be the first response or the second; this is what we also call a variable reward schedule. Technically we mean a variable ratio of reinforcement. In our training program, after luring the dog in a 1:1 fixed ratio, we then transition to variable reward. In making this transition, we begin with a high frequency (a high average ratio 1:3 perhaps), and as the dog progresses, we move to a lower frequency (a low average ratio of correct behaviors receiving primary reinforcement 1:10 perhaps) during any training sequence. Slot machines are an example of variable ratio reinforcement. If they never paid out, we wouldn't play them. But since we all have some experience with winning on a variable ratio, we try hard (spend a lot of money) to get the reinforcement. Las Vegas was built on variable reinforcement.

Random reinforcement is the final category and refers to there being no relationship between the behavior performed and the primary reinforcement given. Nothing is generally learned from random reinforcement.





Extinction

Here is a good time to talk a bit about extinction. The opposite of developing behaviors is getting them to go away or extinction. If a behavior that was heavily reinforced in the past no longer receives any reinforcement (not primary or secondary) the behavior ceases to be rewarded and will stop recurring. This is what we call behavioral extinction. The schedule of reinforcement matters, though, to how easily the behavior will become extinct. A variable ratio reinforcement schedule will tend to make the behavior less vulnerable to extinction because the elimination of the reward's frequency (from the dog's point of view) probably only means that the dog must work a few more times to receive the reinforcement. Thus, as we cease all primary reinforcement, we are likely to see improved effort for a while rather than no effort (this improved effort in the face of no reward is called an extinction burst).

A good example is going to a drink machine. The dollar bill acceptor is usually on a variable reinforcement schedule. After some unfolding and flattening out of the bill, it is normally accepted, so you don't expect it to work the first time. If the machine does not accept the bill the first time, you try again, and again, harder and harder (extinction burst). After a time, if it still doesn't work, you will locate another machine that is less temperamental.

Consider, however, a behavior that has been rewarded on a fixed ratio of 1:1, and the reinforcement stops. We are likely to get non-compliance due to immediate extinction. The dog isn't trained to work through not receiving a reward each time it performs correctly. Thus, behaviors that are reinforced on a 1:1 fixed ratio are easier to extinct than behaviors that are reinforced on a variable ratio. It becomes imperative to transition to reward your dog on a variable ratio to make the behavior reliable. Meaning, if you always heel with a tug under your arm or always reward your detection dog at every successful final response on target odor, your trained behavior is more vulnerable to extinction.

Detection Example:

Understanding reinforcement schedules shows us that to get a behavior trained, we must first start with a 1:1 fixed ratio. As the dog learns that the behavior he performs has a predictable consequence of reward, the dog will normally start offering these behaviors. For example, when we teach the final response in detection on reward from source boxes, each correct response in the beginning (sitting on target odor at source) results in a reward being offered. However, in detection, generally, we are both rewarding the dog for searching for an interval of time as well as finding target odor correctly.

If we use boxes with K-9 BSD™ reward from source devices, we start with one box, and the sit on target odor earns a Kong™ being released remotely for this correct behavior. Repeat this a number of times on the 1:1 fixed ratio. There is very little "searching" going on while we concentrate on teaching the final response. We then go to two boxes, which becomes our beginning on a variable interval for rewarding the searching behavior, but we remain with fixed interval (1:1) reward for a correct final response. Every time the dog finds the target odor and sits, he receives the reward. We are only changing one variable at a time by adding blank boxes into the search.

We continue adding in blank boxes, thus increasing the variable interval of search time between rewards at the source. Keep the actual reward for sitting at source on a fixed 1:1 ratio. As we get further into the detection training protocol, off the boxes and back to normal search areas, and the dog is solid on his final response in the 1:1 fixed ratio, this as well will become a variable ratio. We will often praise the dog off a correct final response at source and move to the next search area so that the power of the variable reward can solidify the final response behavior against unwanted extinction during deployments.



The Training Progression

The training progression we have outlined flows like this with the three phases at the top, the use of both motivation stages (lure, reward (1:1 fixed ratio), and variable reward), and correction stages (verbal reprimand which is negative punishment, to guiding corrections which is light positive punishment, and standard corrections which are positive punishment).

Figure 1: Training Progression

Phases:	Learning Phase (Acquisition)	Correction Phase (Fluency)	Proofing Phase (Generalization)
---------	---------------------------------	-------------------------------	------------------------------------

Motivation: Lure >>>>>>> Reward >>>>>>> Variable Reward >>>>>>>

Correction: Verbal Reprimand >>> Guiding Corrections >>> Standard Corrections

In the *learning phase* during luring, the food is clearly visible to the dog. We hold the food in our hands, show it to the dog, then lure him into position. For example, if we want to teach the dog to sit, we hold some food in our hands and put it right on his nose in our closed fist. Then we bring the food slowly up over the dog's head and as he follows the food with his nose, and when his rear-end hits the ground, we have a successful trial. As soon as she does this, the trainer must mark the event with proper praise such as "Good!" and immediately after that feed the food. This process of luring the dog into position is repeated over and over. This process of shaping the behavior is being done with operant conditioning. The ratio of reinforcement is fixed at 1:1. Remember also to give the appropriate command once the dog starts exhibiting the behavior with the lure just before you start steering the dog into position, so the dog will learn what word (command) cues the behavior that receives the reward. This is the classical conditioning part of training and associates the verbal cue (command) with the behavior we are shaping.

After the dog starts to respond each time by offering the proper behavior when prompted by the command and lure, we then can change the way in which the food reinforcement is offered. We want to get away from our fixed ratio of reinforcement which is most vulnerable to extinction to a variable ratio of reinforcement. Trainers who stay in the luring stage of training too long will create a dog that only responds to the sight or smell of food to perform. The constant use of this primary reinforcement never allows the dog to learn to operate without its presence.² Thus, if you always heel with a tug under your arm, then on certification day it goes away, the dog is working out of context, his visual focus on the reward he earns was never weaned away. The dog's performance will decline dramatically and possibly extinct quickly. Therefore, we must change the dog to reward with the tug or ball out of sight.

In order for the conditioned commands to take hold, we must remove the lure. We do this very gradually by going to the next stage called the reward stage. Hence why our system is called "Lure-Reward." In the lure stage of training, the dog can see the tug under the arm, or smell/see the food in your hand as you lure him in position. In the reward stage, we now take the food or toy out of sight of the dog.

One purpose of making the primary reinforcement available as an after-the-fact reward is to teach the dog to trust, which is an important concept. Without the stimulus of the sight and smell of the food in your hands, some dogs will choose not to participate in the work (after all the lure is gone). Moving the reward out of sight literally changes the game we are playing with the dog. We are now asking him to respond to our voice commands before he gets to see, smell, or taste the food, or see the toy. He must learn to trust that his reward is going to come but first must do what we are asking, a key transition point. Transition points in training are those moments where we change the rule set of the game on the dog and expect him to adapt to the new rule set.

I want to make clear that in the learning phase, both during the luring and reward stages, the only punishment for not performing is simply to mark the mistake with the verbal reprimand (negative marker like "No") then withhold the reward (negative punishment), start over, and try again. There is absolutely no collar correction (positive punishment) in this stage because the dog doesn't understand what we want him to do yet.

In our transition from luring to reward, now that we are changing the game, again we must look at things from the dog's perspective. We must show the dog that his trust is appropriate. If the dog refuses to comply with a command he has been performing well upon being lured, we simply verbally reprimand the dog, giving feedback to the dog over the non-compliance, and start over trying again to elicit the proper response.

2. Typically this is why many trainers object to using food reward in training. They say that the dog will always expect food to be produced before any obedience will happen. There is truth in this. If you don't change the reward system away from baiting, the dog will see the presence of food as the cue to do the obedience. This is nothing more than classical conditioning. The dog will expect to see food (either in your hand or in your bait pouch) before complying with your commands. This is obviously counter-productive.



Sometimes it helps to step back for an instant, lure him a couple of times in a row, then throw in a command on reward; this is what I call training momentum. If the dog gives us proper responses in quick succession through luring, the likelihood is that he will continue with these good responses. This behavior highlights the necessity of making your training sessions well organized and move quickly. This use of training momentum will increase the dog's attention span and keep him on task.

If you look at Figure 1, you will see that the reward stage spans the learning phase and the correction phase. This is critical. After the dog begins to work well in the reward stage, we now start to introduce guiding corrections. Corrections are how we physically communicate to our dogs that they are performing improperly. A correction is unemotional feedback to the dog that there is a consequence to non-compliance.

Correction (Fluency) Phase

Every leash correction has two components. It has force and direction. For example, a sit correction is a pop-and-release on the leash in the upward direction. We must always use proper direction in teaching our corrections, but we can vary the force. A guiding correction is simply a lighter version of a normal or standard correction. A guiding correction makes use of little force and relies on the trainer to use body language to assist the dog to be compliant.

Once the dog understands the guiding corrections, we then may vary the force by the circumstance. Remember, the purpose of the force in a correction is to cause the dog enough discomfort that he will work to avoid it - the essence of positive punishment. It is applying an unpleasant consequence to reduce the likelihood of a behavior. If the correction is always too light, we will nag the dog and will not get compliance, because the force used is not an aversive. However, we always follow the minimum force rule: when applying the consequence of positive punishment, use only the amount of force necessary to accomplish this task of changing behavior. This is punishment in the animal behavior sense, meaning the application of an unpleasant consequence to a behavior in order to reduce the likelihood of that behavior.³

To summarize: at this stage in training, the dog is asked to perform a command, and depending on his response, the trainer will administer either positive reinforcement (the dog complied) or positive punishment (usually a correction for doing anything else besides the command we wish the dog to perform). Once the dog complies, even after a correction, the positive reinforcement is then administered.

You can see that the compulsion is directed to "attach a cost" to any behavior other than the one we are expecting or encouraging. The goal of correction is to refocus the dog on the task at hand by making non-compliance have a consequence the dog finds unpleasant. There is always positive reinforcement given after the successful completion of the task. Meaning that even if a correction was administered to elicit the appropriate behavior, positive reinforcement is required. We are still giving the dog a 1:1 fixed ratio of reinforcement in the reward stage, the only difference is we have added corrections into the progression.



3. Corrections are not "punishment" in the usual sense that we understand: Johnny got a D on his midterm, so he gets grounded for a week. Punishment as "after the fact" correction is not appropriate. Punishment in this context relies on the ability to reason temporally and dogs can't reason in this way.



Now comes the time to begin weaning the dog off of the 1:1 fixed ratio and move to variable ratio reinforcement. First, we had to go from lure to reward. The reinforcement schedule remained the same at 1:1, but the manner of reinforcement went from lure to reward. The association of the marker “Good,” or the clicker before the positive reinforcement has put money in the bank for the positive marker. Thus, the dog will work for the marker. This process of replacing the primary reinforcement (food or toy) with secondary or conditioned reinforcement (“Good!”) is an example of classical or associative conditioning. Just as we replaced the primary stimulus, the sight and smell of the food in our hands, with the conditioned or secondary stimulus, the command word, we can replace the primary reinforcement with a secondary reinforcement. We continue to give the conditioned punisher (verbal reprimand) before administering a correction, so that we “put money in the bank” for the verbal reprimand as well.

The reinforcement schedule starts at a relatively high frequency, and over time moves to a lower frequency. Remember it is the association between the positive marker “Good,” and the tangible

reinforcement of his toy, or food reward that keeps the machine running. If we stop the variable reinforcement, we will at first get a very hard working dog giving us a lot of behaviors that will go unrewarded as an extinction burst, and then eventually the dog will stop performing altogether. Praise (or verbal marker, or positive marker, whatever you want to call it) is only as good as its association with the dog’s reward. Just as the negative marker “No,” is only as good as its variable association with actual corrections. If either the rewards go away or the physical corrections go away, the dog will learn that these conditioned markers have no real value.

The next time you think about training a behavior in your dog, think about how you can train it with a simple lure-reward system. Each stage of the lure-reward system must be layered properly as explained in Figure 1. You can think about it proceeding in the following five simple steps: (1) lure with only negative punishment for incorrect results, (2) move to reward from the lure, still 1:1, (3) guiding corrections with still 1:1 fixed ratio reinforcement, (4) standard corrections are next with still 1:1 fixed ratio reinforcement when the dog is correct and after the correction (5) variable reward is introduced (high frequency then moving to lower frequency) with standard corrections.

Proofing (Generalization) Phase

The positive punishment is marked with a negative marker, “No,” just as the positive reinforcement is marked with the positive marker “Good.” Using a complete lure-reward system like this over time, consistently, creates the habit of a trained dog. Your job as the trainer is to produce simple, understandable feedback to your dog with every interaction you have with him, in every area of training, from obedience to tracking to controlled aggression and detection. This feedback is what creates his habits, and once a dog is in the habit of performing, you have a trained dog. You still must always be ready to provide the necessary feedback in a dynamic environment like deploying a police dog. This takes time, and above all, consistent application of proper behavioral principles at all times. Don’t let your dog down or blame him for your inconsistency or expect him to know something that hasn’t been practiced on variable reward, with low-frequency reward and standard correction in an environment where we have proofed or generalized the behavior.



“...there are many levels of understanding depending on the place you are in the training progression. Understanding to a dog is a contextual experience.”



The continuous working of the dog in new and varying environments, adding in distractions, and changing training venues is critical for the dog to really know a concept inside and out. Many people see a few correct responses and believe that their dog actually understands the behavior. However, as we can see in this article, there are many levels of understanding depending on the place you are in the training progression. Understanding to a dog is a contextual experience. When the dog can perform in a generalized fashion, anywhere and everywhere the same, you have created sufficient habit to say you have a dog trained to that behavior. Don’t expect that he knows something without having progressed through all of the steps in the system with sufficient repetition that you are only rewarding correct responses and almost never correcting mistakes. ■