

Chapter 1

Perspectives on animal learning theory

Part 3

Principles of animal learning

Stewart Hilliard

Summary

The objective of this section is to summarise some of the main principles of animal learning, and to make these principles immediately relevant to the training of working dogs. Familiarity with the basics of animal learning will give the working dog trainer new names to apply to training procedures. More importantly, understanding of learning principles allows the trainer to adapt training to suit the individual dog, to understand why a given method is not working, and to understand how one lesson influences the dog's ability to learn the next. Above all, it is my hope that an appreciation of the psychological processes underlying training will move the reader away from the emphasis on physical force and punishment that has characterised training of working dogs throughout much of its history.

This theoretical perspective on working dog training emphasises the processes of habituation, classical and instrumental conditioning, and extinction.

Habituation is the gradual loss of responsiveness to a stimulus as a result of repeated exposure to that stimulus. In order to produce effective habituation (of, for example, a fear-eliciting stimulus like a loud noise), the trainer should present the stimulus in a weakened form until the dog exhibits little fear. Then the trainer should gradually increase the stimulus intensity until the dog exhibits little fear to the stimulus even at full strength. The stimulus presentations and training sessions should be well separated in time.

Classical conditioning is a process in which the dog learns to associate two environmental stimuli. One is a relatively weak stimulus such as a noise or a voice cue (called the conditioned stimulus, or CS). The other is a powerful stimulus (called the unconditioned stimulus, or US) such as food or a sharp collar correction. As a result of repeated pairings of the two stimuli, the CS elicits a response (the conditioned response, or CR) that was formerly elicited only by the US. Classical conditioning does not involve rewards or punishments,

but involves the dog developing an involuntary response to the predicting stimulus. Thus, a dog learns to salivate and become excited when it hears the ring of its food pan in the cupboard, or it learns to cringe and feel anxiety when it hears the word "No!". Classical conditioning principally involves learning reflexive and emotional reactions, such as likes and dislikes, attitudes towards the handler, work, and training. Forward conditioning, in which the CS precedes the US by a second or less, is the most effective procedure for classical conditioning.

Dog training seldom involves deliberately-staged classical conditioning procedures, but dogs are inevitably presented with CSs and USs as a part of their experiences and learn to associate current with future stimuli (such as tension in the handler's arms followed by a jerk on the collar). Classical associations can be exploited by the handler, for example by eliciting excitement using the cue "Ready to work?" (signaling a forthcoming ball reward). Classical associations can also interfere with training objectives, for example when the dog freezes instead of sitting in response to the "Sit!" command (because the command is often followed by a collar correction).

During *instrumental conditioning*, the dog learns to associate a voluntary behaviour (the instrumental response) with its result (the consequence). Reinforcers are consequences that encourage prior behaviour, while punishers are consequences that discourage prior behaviour. Behaviour is reinforced either by giving the dog something pleasant, or by withdrawing or omitting something unpleasant. For example, a sit will be followed by a piece of food (reward, or positive reinforcement), or releasing the ball on command will *prevent* a collar correction (negative reinforcement). Behaviour is punished either by giving the dog something unpleasant, or by withdrawing or omitting pleasant stimuli. For example, breaking the down-stay will be followed by a bump on the nose (punishment, or positive punishment), or jumping up on the handler will *not* be followed by petting and attention (omission, or negative punishment). Each of these relationships between a response and a consequence is called a response rule. Commands, such as "Sit!" and "Out!", are cues telling the dog what response rules are currently in force.

To reward an episodic behaviour (such as barking on command), a continuous reward schedule (CRS, the reward appears after every correct repetition) is used initially. Once the dog has learned the skill, it is trained to repeat the response a few times by using fixed ratio reward schedules (FRRS, the reward appears reliably after gradually increasing numbers of repetitions). Once the dog readily repeats the response, variable ratio reward schedules (VRRS, the reward appears variably or after unpredictable numbers of correct repetitions) are used to make performance persistent. To reward continuous behaviours (such as maintaining a down-stay or continuous searching for target odour), fixed interval reward schedules (FIRS, the reward appears at gradually increasing intervals of correct performance) are used initially. Once performance has been extended, variable interval reward schedules (VIRS, the reward appears at random or unpredictable intervals of correct performance) are used to make performance persistent. Random reward is the most effective procedure for creating steady and consistent responding, and it also renders that responding highly resistant to extinction.

Training using pleasant stimuli as consequences (i.e. rewards to encourage or omission to discourage behaviour) is inductive methodology, best for establishing a working relationship with a dog, for teaching the dog a positive attitude towards work, and for teaching most of the basic skills and

commands. Training using unpleasant stimuli as consequences (i.e. negative reinforcement to encourage or punishment to discourage behaviour) is compulsive methodology, used to produce more reliable or precise performance once behaviours have been established. For compulsive training to be effective, the dog must already have a good understanding of how to perform the target behaviour prior to being *forced* to perform it using compulsion. Negative reinforcement training normally proceeds through two distinct phases, escape and avoidance. During escape training (consisting of one or only a very few repetitions), the dog learns that it can “turn off” an unpleasant stimulus by engaging in a particular response. If the dog does not understand the escape behaviour or does not have the ability to carry it out, then further attempts at negative reinforcement training could be disastrous. At best, the dog will learn to dislike work and to distrust the handler; at worst, it could become fearful or aggressive. During avoidance training, the dog learns that by responding to a cue (such as the “Out!” command) it can avoid the unpleasant stimulation altogether.

Classically- and instrumentally-conditioned behaviours can be extinguished. Extinction of an undesirable classically-conditioned response (e.g. anxiety in response to a correction collar being placed on the neck) involves presenting the CS repeatedly until the CR disappears. To extinguish an undesirable instrumentally-conditioned response (e.g. jumping up on the handler for attention), the response is placed on an extinction schedule by allowing it to occur repeatedly without providing any reinforcement.

Classical conditioning does not involve rewards or punishments, but involves the dog developing an involuntarily response to the predicting stimulus. Thus, a dog learns to salivate and become excited when it hears the ring of its food pan in the cupboard, or it learns to cringe and feel anxiety when it hears the word “No!”. Classical conditioning principally involves learning reflexive and emotional reactions, such as likes and dislikes, attitudes towards the handler, work, and training. Forward conditioning, in which the CS precedes the US by a second or less, is the most effective procedure for classical conditioning.

Introduction

Fundamentally, the task of the detector dog trainer is to teach the animal two lessons. First, the animal must learn an association between a target odour and some highly rewarding stimulus, and second, it must learn to give a specific response in order to gain access to that reward. The association between the target odour and the reward induces the animal to search for the target substance, while the instrumental response provides an “indication” alerting the handler to the presence and location of the target. These two lessons are mediated by different learning processes that have been studied by experimental psychologists for about 100 years. These studies have revealed better ways to bring about learning in animals and human beings, but it has been only

relatively recently that scientific knowledge about animal learning has been applied to working dog training.

Animal learning

Learning is a more or less permanent change in the behaviour of an organism as a result of interaction with the environment. (The terms learning and conditioning are here used to mean the same thing.) This definition distinguishes learning-based behaviour change from short-term behaviour changes such as sensitisation, fatigue, and sensory adaptation. For the purpose of dog training, it is sufficient to discuss three types of learning — habituation, classical conditioning, and instrumental conditioning.

Habituation

Habituation is a gradual decrease in responsiveness to a stimulus as a result of repeated exposure to that stimulus. For example, the first time a dog hears a door slam it may startle. In all likelihood, when it hears this slam again and again, it will gradually startle less and less until, finally, it exhibits little response to the noise. Habituation is adaptive because it allows the dog to save its attention for important stimuli, such as the noise of the lid coming off the dog-food can. Habituation takes place continually during dog training, in ways that are both advantageous and disadvantageous to the dog trainer.

Advantageous habituation

In some situations, habituation may improve the dog's effectiveness. Dogs are normally to some extent frightened of, or interested in, things that are new to them. However, detector dogs are expected to carry out their duties in environments that feature very distracting and sometimes intense stimuli, such as taxiing airplanes and loud explosions. Through habituation, a working dog can learn to respond minimally to irrelevant stimuli and pay attention to the "job". For example, a detection dog trained entirely in quiet, vacant barracks may initially have difficulty concentrating on a search task when commanded to "clear" a building with people inside it. However, once the dog becomes habituated to an environment full of people, its ability to concentrate on the task will improve.

Habituation proceeds most effectively and rapidly with stimuli that are mild or moderate in intensity. It can be very difficult to produce habituation to intense or frightening stimuli. A fearful dog exposed repeatedly to a very intense stimulus, such as a running helicopter engine at close range, is likely to respond *more* intensely to this stimulus over time instead of less intensely.

In addition, habituation is most efficient when the stimulus exposures and training sessions are well separated in time. For example, a dog habituates more easily to gunshots when they are spaced out at intervals of 15 or 20 seconds, and when the training sessions are separated by 24 hours. The opposite procedure, exposing the dog to rapid series of gunshots several times a day, may lead to increased fear of the noise.

When habituation is conducted with mild stimuli spread through time, undesirable responses such as fear tend to weaken or even disappear.

Hierarchies of stimulus intensity. Although habituation proceeds most easily with stimuli that are very mild, mild stimuli are rarely a problem in dog training. Instead the stimuli that disrupt the dogs' performance by inducing fear and anxiety are often very intense, such as gunshots or running jet engines. However, habituation to even very intense or frightening stimuli can be achieved using a hierarchy of intensity — a scale on which stimulus intensity increases gradually from low to high. A practical way to decrease the intensity of noise stimuli is by exposing the dog to them from a great distance. For example, the dog is first presented with the sound of a running jet engine at a distance of 500 metres, where it exhibits only mild anxiety. Once this anxiety disappears through habituation, then the dog is moved to the next level in the hierarchy by walking to within perhaps 400 metres of the jet engine. Once any anxiety responses have habituated, then the handler moves the dog closer, and so forth. By introducing the dog to frightening stimuli at a level of intensity that is so low that it provokes little or no fear response, and by moving gradually from one stage on the hierarchy of intensity to the next, the trainer trains the dog to exhibit little fear in the presence of even very intense stimuli.

Counterconditioning. Habituation can be accelerated by pairing a strong pleasant stimulus with the fear-inducing stimulus. To again use the previous example, if the dog is frightened of a running jet engine, its fear can be offset by presenting the dog with a strong pleasant stimulus like food or a ball. This procedure is called counterconditioning, and should be combined with the use of hierarchies of intensity. This is because, if the dog is very afraid of the jet engine, it will ignore the food or a ball when it is close to the engine — it will be too afraid to eat or play. Therefore counterconditioning is conducted at distances progressively closer and closer to the engine, beginning at a distance great enough so that the dog's anxiety is easily offset by its pleasure at being presented with food or a ball. At each stage of the hierarchy (i.e. distance from the jet engine), it is essential that the trainer uses the dog's behaviour as the measure of when to proceed to the next stage, so that counterconditioning proceeds at a pace suited to the individual dog. Moving too quickly "up" the hierarchy (i.e. closer to the jet engine) will not produce reduction in fear, and could even countercondition the food or ball, reducing the dog's pleasurable response to these motivators. (*Note: Counterconditioning is actually a form of classical conditioning — to be explained below — but it is introduced here for the sake of continuity in discussion of methods of fear-reduction.*)

Spontaneous recovery. Fear responses are very durable and persistent and can re-emerge even after extensive habituation "therapy". This is because habituation includes certain short-term processes that "wear off" after a few minutes or hours, and it is normal for a habituated response to re-appear to some extent between training sessions. Thus, a dog may exhibit no fear of a stimulus by the end of one day's training session, yet show recovered fear at the beginning of the next session. This phenomenon is called spontaneous recovery. For example, a dog habituated to a loud noise like an air compressor (i.e. shows no fear), may display some fear the next time it encounters the compressor. It is important to understand that even when habituation and counterconditioning are correctly applied, the fear response will often re-emerge. However, from session to session there should be less and less spontaneous recovery of the fear response.

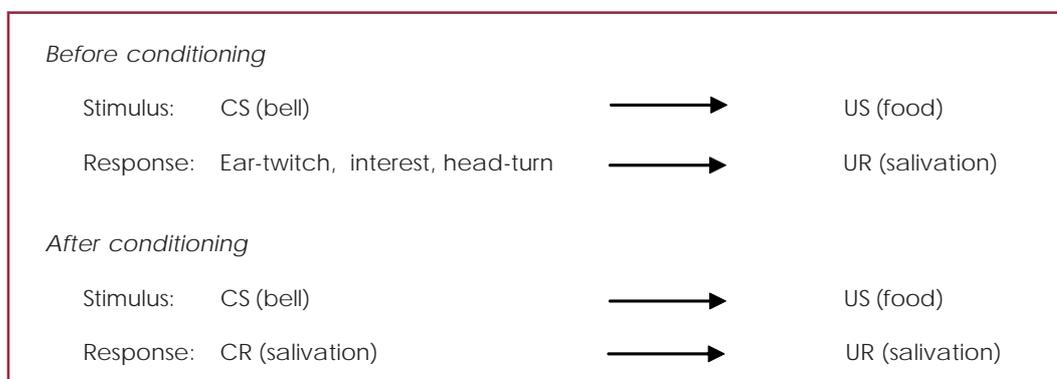
Disadvantageous habituation. Habituation may decrease rather than increase the dog's effectiveness because effective performance in a working dog depends upon a certain level of interest in, and responsiveness to, environmental stimuli. A dog that is relatively new to detection work may deliver very intensive and focused search behaviour because it is stimulated and excited by the new learning situation. But after more experience, the dog may become sluggish and appear "bored". The best weapon against this disadvantageous habituation of the dog's enjoyment of its work is to inject as much variety as possible into the dog's daily routine. For instance, a detector dog should be trained in as many different locations as possible, and a variety of different reward objects and games should be used as reinforcers. In addition, spontaneous recovery can be used to assist training — if the dog is not worked for a period of time there should be some recovery in its enjoyment of work.

Classical conditioning

Classical conditioning is the learning of emotional and reflexive responses through the formation of mental associations between stimuli. For example, a dog can learn fear of a veterinary clinic in the following way: If the dog has never before been in the clinic, the stimuli of the clinic (such as its look, smell, and sound) should be neutral or meaningless. However, after the dog is restrained by technicians and injected with a needle, it may associate the look, smell, and sound of the vet clinic with physical restraint and the pain of the injection, so that the next time it is taken in for a procedure the clinic stimuli are no longer neutral — they will elicit the same fear that physical restraint and injection do.

In classical conditioning (also called Pavlovian conditioning, Fig. 1) the dog learns a relationship between two events, or stimuli. One of these stimuli is a "neutral" or unimportant stimulus that a dog would normally pay little attention to. This stimulus is called the conditioned stimulus, or CS, because it can generate strong behaviour only as a result of conditioning. The other stimulus is a biologically important stimulus that a dog naturally pays a lot of attention to — like food. This stimulus is called the unconditioned stimulus, or US, because it can generate strong behavioural responses without any conditioning.

Figure 1
The structure of classical conditioning



In the classical example, the Russian scientist Ivan Pavlov taught dogs to salivate in response to the ringing of a bell. Pavlov did this by repeatedly pairing a bell (CS) and some food (US), presenting them close together in time. A naive dog normally responds to the ringing of a small bell by merely twitching its ears or looking towards the noise. However, a piece of food can cause the dog to show a great deal of strong behaviour like excitement, salivation, digging and pawing, chewing and eating. This very strong behaviour caused by exposure to a US is called the unconditioned response, or UR. Through classical conditioning, the CS and the US become associated in the dog's "mind", so that behaviour that is naturally triggered by the US (the UR) comes to be triggered to some degree by the CS also. When a CS develops the ability to trigger behaviour that is normally caused by a US, this learned response is called the conditioned response, or CR. Thus, eventually Pavlov's dog learned that the bell predicted food, and began to salivate in response to the bell (CR).

Applying classical conditioning

The most effective procedure for classical conditioning involves presenting the neutral stimulus/command (CS) immediately *before* the biologically important stimulus (US). Thus, if the handler wishes to train the dog to feel startled and anxious in response to the word "No!" then an effective method would be to wait until the dog engages in some misbehaviour like sniffing the trash. The handler then gives the "No!" cue, and throws a chain choke collar against the trash can so that it makes an unpleasant sound about half a second after the "No!". Originally the word "No!" (CS) means little to the dog and produces little change in behaviour. The unpleasant noise (US) is potent and causes a strong startle or freezing response (UR). Pairing the "No!" with the unpleasant noise teaches the dog to startle/freeze in response to "No!" (CR) within one or a very few CS-US pairings. Later, when the dog is engaged in misbehaviour, the handler can use the "No!" command, causing the dog to startle/freeze (which serves to interrupt the undesirable activity), and the handler can then recall the dog and praise it. The dog will soon learn to shy away from behaviours and objects when it hears the "No!" command (classical conditioning) and return to its handler for praise (instrumental conditioning — see below).

When the CS and the US are reversed, so that the US precedes the CS, this is called a *backward conditioning* procedure. Little or no learning takes place during backward conditioning. Thus, even many repetitions of a training trial in which the dog is startled with a loud noise and then hears a "No!" may not produce a startle response when "No!" is given by itself.

Extinction of classically-conditioned responses

Not everything a dog learns through classical conditioning is desirable. For instance, if a harsh correction is given the first time a chain choke collar is placed on the dog, the animal will probably show inhibition and anxiety the next time a chain collar is placed on its neck. Undesirable classically-conditioned responses can be weakened or even abolished by presenting the CS repeatedly without pairing it with the US, causing the CR to gradually decrease in strength. This process of *extinction* is just like habituation, except that in habituation an unlearned response is gradually abolished, whereas in extinction a learned response is gradually abolished.

For example, suppose a dog has been trained to become aggressive when it hears gunfire (by association of the sound of gunshots with the opportunity to bite), but in

its new job as a detection dog this response is undesirable. Here, a previously trained response has become a behaviour problem. The remedy is not simple, but a first step is to produce some extinction of the aggressive response by putting the dog in a neutral situation and firing a gun repeatedly until the animal's arousal response to the gun decreases.

Extinguishing a learned behaviour does not mean that it has been unlearned or "erased", even though the behaviour may no longer be expressed. Learning can produce permanent changes in the brain that are not reversed by extinction. Thus, a behaviour thought to be "extinct" may suddenly reappear, especially with a change in the context (environment) of training.

Instrumental conditioning

Instrumental conditioning is the learning process through which an animal's behaviour is changed by the consequences, or results, of that behaviour. For example, if a puppy approaches and sniffs a cat, and the cat spits and scratches the puppy's nose, the puppy is less likely to approach the next cat it sees. The behaviour, approaching cats, has been modified by its consequence, a startling and painful experience. This is an example of punishment. A contrasting example: if a dog chews at the door of its crate and the crate bursts open, then the dog is more likely to chew at the crate in the future. The behaviour, crate chewing, has been modified by its consequence, freedom. This is an example of reward, also called reinforcement.

In both examples of instrumental learning above, an association is formed between a behaviour and a stimulus, and this is a central distinction between instrumental conditioning and classical conditioning. Classical conditioning involves formation of an association between two stimuli, as in the CS-US association between the "No!" and an unpleasant stimulus like a collar correction, described above. Instrumental conditioning procedures emphasise formation of a response-stimulus association, for example between the act of sitting and some pleasant stimulus like a piece of food.

Instrumental conditioning model

Four elements are involved in an instrumental conditioning procedure:

- First, is some *response* on the part of the dog; usually some skeletal response (so-called because it employs the voluntary or "skeletal" muscles) such as sitting, lying down, or barking.
- Second, is the *consequence*, or result of that response. Examples include a scratch on the nose and freedom from the crate.
- Third, is the *response rule*, the law that links the dog's response to the consequence (also called a response contingency). For example, common response rules in dog training are the following: a sit is followed by praise; releasing a ball prevents a collar correction; breaking a down results in a collar correction. More than one response rule can, and usually does, apply to any given target behaviour. Thus, in traditional dog training methods, the down exercise often involves four response rules: (1) lying down results in a reward such as praise, and (2) the reinforcement of not being corrected; (3) not lying down results in the punishment of being subjected to a jerk on the collar, and (4) the punishment of not receiving praise.
- Fourth (even though in practice it usually comes first), is a cue or *command*. Commands are signals (also sometimes called discriminative stimuli) that tell

an animal when a particular response rule is in effect. The handler will not reward a sit anytime the dog sits, but only when a sit response is desired. The command “Sit!” tells the dog that one or more response contingencies are in force — for instance, a prompt sit will result in petting and praise and the omission of a collar correction, while refusing to sit will result in no petting or praise and the administration of a collar correction.

Consequences

There are two main categories of consequence in instrumental conditioning procedures— *reinforcement* and *punishment*.

Reinforcement is a consequence that encourages or strengthens prior behaviour. Examples of reinforcers are food, access to a toy, or a pat on the head. Any of these, when given to the dog after it sits, tends to strengthen sitting behaviour. Food, toys, and pats on the head are reinforcing because they are pleasant. These kinds of consequences are called “positive reinforcers”. However, unpleasant events can also act as reinforcers. For example, the handler can reinforce a behaviour by means of an unpleasant event like a jerk on the leash, *by withholding the jerk when the dog sits*. In this example, there is a negative response rule between sitting behaviour and a jerk on the collar — if the dog sits, there will be no jerk. Although the jerk itself is unpleasant, the absence of the jerk is a “satisfying state of affairs” and will, under proper circumstances, serve to reinforce sitting behaviour. This kind of reinforcer is called a “negative reinforcer”.

Punishment is a consequence that discourages or weakens prior behaviour. Examples of punishers are jerks on the collar (collar corrections) or bumps on the nose. Either of these, when administered to a dog after it misbehaves by, for example, breaking the down-stay position without permission, will tend to weaken down-stay-breaking behaviour. Collar corrections and bumps on the nose are punishing because they are unpleasant. These kinds of consequences are called, paradoxically, “positive punishers”. However, pleasant events can also act as punishers. The handler can punish an undesirable behaviour by withholding or taking away a pleasant stimulus like praise and petting. For example, if the dog tends to jump up on the handler when it is excited, jumping up can be punished by withholding praise and attention. In this example, there is a negative response rule between jumping-up behaviour and praise and petting — if the dog jumps up, it will not receive praise or petting. Because the praise and petting are pleasant, their absence is an “unsatisfying state of affairs” and will, under the proper circumstances, punish jumping-up behaviour. This is called “negative punishment”.

Use of the terms “positive” and “negative”. The terms positive and negative help define the four possible consequences of an instrumental behaviour — positive reinforcement, negative reinforcement, positive punishment, and negative punishment. However, “positive” and “negative” used in this sense have nothing to do with whether the procedure is pleasant or unpleasant for the dog. A commonly misused expression is “negative reinforcement”, because non-psychologists interpret the word “negative” as meaning bad or unpleasant. Thus negative reinforcement is commonly used as though it was synonymous with physical punishment, but it is not. The terms positive and negative in this context actually refer to the nature of the response rule (also referred to as a contingency) between the dog’s response and the consequence.

In relation to the notion of contingency:

- The adjectives “positive” or “negative” do not indicate whether the procedure involves applying discomfort to the dog.
- It is possible to reinforce a behaviour by means of either stimuli that the dog likes (by applying these stimuli to the dog immediately after the behaviour), or stimuli that the dog does not like (by removing or omitting these stimuli immediately after the behaviour).
- It is possible to punish a behaviour by means of either stimuli that the dog likes (by removing or omitting these stimuli immediately after the behaviour) or stimuli that the dog does not like (by applying these stimuli to the dog immediately after the behaviour).

The following simple language summarises the review above:

- *Reward (positive reinforcement)* encourages a response by giving the dog something it likes immediately after the response.
- *Punishment (positive punishment)* discourages a response by giving the dog something it dislikes immediately after the response.
- *Negative reinforcement* (no synonym) encourages a response by taking away from the dog something it dislikes immediately after the response.
- *Omission (negative punishment)* discourages a response by taking away from the dog something it likes immediately after the response.

Primary and secondary reinforcement and punishment

Many rewards and punishments are biologically powerful stimuli, such as the opportunity to eat or a painful jerk on the collar. In the language of classical conditioning, they are called unconditioned stimuli (USs). In the language of instrumental conditioning, they are called primary reinforcers or primary punishers. Dogs respond readily and strongly to these stimuli without having to be taught to do so. However, some rewards and punishments originally have little effect on a dog's behaviour. Called secondary reinforcers and punishers, they do not become effective until they have been associated with primary reinforcers or punishers.

Secondary reinforcers gain their ability to strengthen and encourage behaviour by being associated (through classical conditioning processes) with primary reinforcers. For instance, puppies probably do not instinctively enjoy being spoken to. They learn to like being spoken to in a happy voice because this voice is associated with physical petting and with the presentation of food. After enough of this conditioning, words like “Good!” spoken in a happy voice become pleasant stimuli. Subsequently, the word “Good!” has the power to reinforce prior behaviour (if the handler says “Good!” immediately after the dog executes the behaviour).

Secondary punishers gain their ability to weaken and discourage behaviour by being associated with primary punishers. For instance, the word “No!” (spoken in a neutral tone) means nothing to an untrained dog. The word becomes unpleasant because it is associated (through classical conditioning) with unpleasant primary punishing events like a jerk on the collar. After enough of this conditioning, the command “No!” becomes an unpleasant stimulus. Subsequently, the word “No!” has the power to punish prior behaviour (if the handler says “No!” immediately after the dog executes the behaviour).

Systems of instrumental training

So far, instrumental procedures have been classified in terms of whether they are reinforcing or punishing. Another useful way to classify instrumental procedures is in terms of whether they are inductive or compulsive; that is, whether they use pleasant or unpleasant consequences to motivate and teach the dog.

Inductive training

In inductive training the handler relies on the use of pleasant events and stimuli to obtain desirable behaviour from the dog. For instance, a dog can be taught to put its feet up on a wall by using a piece of food to entice it up onto the wall, and then letting it eat (reward). Or, a puppy can be taught not to cry in its crate by refusing to open the crate when the puppy is crying (omission). The root word “induce” means to persuade. Thus, inductive training involves the use of reward (encouraging desirable behaviour by administering pleasant stimuli) and omission (discouraging undesirable behaviour by withholding or withdrawing pleasant stimuli).

Compulsive training

Compulsion is a word that refers to forcing or coercing people or animals to do things. In dog training, the terms “correction” and physical compulsion are equivalent. For example, a dog can be forced to stay inside a kennel when the gate is open by shutting the gate sharply on its nose when it tries to walk out (punishment). Or, a dog can be forced to get down off a table by poking it roughly in the ribs until it jumps down to escape the poking (negative reinforcement). In each case, a dog trainer would commonly speak of “correcting” the dog. In compulsive dog training, the handler relies on unpleasant events to obtain desired behaviour from the dog. Thus, compulsive training involves the use of negative reinforcement (encouraging desirable behaviour by withdrawing or withholding unpleasant stimuli) and punishment (discouraging undesirable behaviour by administering unpleasant stimuli). Although the training of working dogs often involves the use of some compulsive methods, it is important to understand that:

1. Compulsive methods are effective and humane only under certain circumstances, specifically when the dog already understands the desired response and how to avoid compulsion;
2. Excessive reliance on compulsion will damage the dog’s rapport with the handler and decrease motivation for work;
3. Compulsion may stimulate defensive and aggressive responses in the dog, and can in many circumstances be counterproductive and even dangerous for the handler; and
4. Some phases of working dog training, most especially when training detection, are to an extent incompatible with compulsive techniques.

Inductive versus compulsive training. Inductive and compulsive training are complementary methods. Inductive methods are normally best for teaching the dog to understand commands and to enjoy working. Compulsive methods are often best for ensuring that the dog executes a commanded skill (especially obedience skills) reliably in any circumstance. Some compulsion is normally necessary in working dog training, especially in bitework and obedience. However, inductive methods are to be preferred whenever practical. In particular, inductive methods are normally better than

compulsive methods for the initial teaching of skills. In fact, the use of compulsive techniques for initial teaching can be counterproductive.

For example, to an untrained dog a command (e.g. "Sit!") means nothing. Therefore, if the handler gives the command "Sit!" and then administers a strong collar correction in the attempt to force the dog to sit (a classical compulsive method for teaching the sit), the dog will have no idea that it can avoid further unpleasantness by sitting. It may instead attempt to defend itself or, more commonly, seek to avoid the handler. More than anything else, such a method is a perfectly designed classical conditioning procedure that conditions fear and/or aggression to the command "Sit!" by pairing the command closely together in time with physical discomfort. However, if the dog is first taught to understand and respond to the sit command using inductive methods, then compulsion can be used constructively to hasten the dog's sit or to teach it to sit even in distracting circumstances. Thus, the proper role of inductive training is to teach the dog an understanding of skills, while the proper role of compulsive training is to enforce the performance of these skills, if necessary.

Another important application for inductive techniques is during the early stages of handling any dog, even a well-trained dog. The optimal way for a handler to build rapport and a good working relationship with a new dog is to perform inductive training exercises with the animal (such as food-rewarded obedience), even if the dog already knows these exercises.

When training detection dogs, the use of compulsive techniques must be minimised. Attempts to force dogs to search for long periods of time are normally unsuccessful. The detector dog should respond to cues like the command "Seek!" with arousal, excitement, and desire for a reward. Detector dog search behaviour is essentially a voluntary effort, with a reward as its objective.

The detector dog trainer must also be careful not to link compulsion with any stimuli or procedures that will commonly be part of detection scenarios, because these stimuli and procedures could take on secondary punishing properties; for example, if harsh corrections are used to teach the dog to sit, and subsequently the sit is trained as the dog's final indication response in substance detection. Similarly, punishing a dog for being distracted by food in the course of a search could turn food into a strong secondary punisher, and to the extent that search problems often contain edible items, the dog's work will be inhibited by anxiety and fear triggered by food odours.

Application of inductive training

In inductive training, the handler employs gentle means to persuade a dog to perform some target behaviour and then reinforces this behaviour. In the event that the dog does not execute the desired behaviour, or executes it incorrectly, the handler omits reinforcement. In the classical example, the handler teaches a dog to sit by drawing the dog's attention to a piece of food in the hand. Once the dog places its muzzle in contact with the handler's fist in the attempt to take the food, the handler then slowly raises the hand and moves it slightly backwards towards the dog's tail, simultaneously giving the command "Sit!". In following the movement with its head, the dog is very likely to sit. The handler then opens the hand to feed the dog and gives praise. If the dog fails to sit, the handler withholds reinforcement and praise and continues attempting to encourage the sit. In such a stress-free setting, a dog can learn very rapidly to sit on command, it enjoys the work, and it develops affection and trust for

the handler. The stimuli and events during the sit exercise all become secondary reinforcers. Similar methods can be used for teaching almost all obedience exercises.

Once basic skills have been taught to the dog in this inductive fashion, then moderate and humane compulsive techniques can be applied to “polish” the skills without fundamentally changing the dog’s basic liking for its work.

Successive approximation and shaping

Successive approximation is an inductive training procedure in which animals are taught new behaviours by rewarding responses that are progressively more and more like the desired target response. For instance, to teach a dog to sit through successive approximation, a handler gives the “Sit!” command, and then waits until a small approximation of a sit is seen, such as flexing of the hind legs, and then reinforces this movement. Once the dog is flexing its legs readily for reinforcement, then the handler should withhold reinforcement until the animal exhibits a flexing that is slightly greater than before and then rewards this behaviour, and so on. The entire process of extracting a trained response through successive approximation is called *behaviour shaping*. Successive approximation and shaping are of central importance in the training of animals such as killer whales and sea lions, but they play comparatively little role in dog training, for the simple reason that a good dog trainer can usually think of a way to get the dog to offer the complete behaviour, and then reward *that*. For example, teaching a dog to sit using a piece of food held in the hand (described above).

However, particularly in the case of very complex or difficult behaviours, or behaviours that the dog resists learning, a handler may reinforce the dog for a good “effort” in the direction of the desired target behaviour. For example, in the course of an initial sit-training session, if the dog only crouches without completing the sit, but is working hard to please the handler and obtain the food, it is usually wiser to reward the animal for the partial sit, encouraging the dog to continue trying to learn the lesson.

Reward schedules

A reward schedule is a rule that dictates how often a dog receives positive reinforcement when it correctly executes a skill. For example, apply the rule that the dog will be rewarded every time it downs promptly on command. Alternatively, the dog might be rewarded every third time it downs promptly on command. Such schedules produce different effects and are appropriate at different stages of training. There are six types of reward schedule.

1. **Extinction schedule.** An extinction schedule is used to eradicate or extinguish a learned behaviour. To extinguish an instrumental response the behaviour is allowed to occur repeatedly, without being rewarded. The behaviour will gradually decrease in strength and frequency until it disappears. Thus, to extinguish an undesirable behaviour (e.g. jumping-up), it is often sufficient to identify whatever is rewarding the behaviour (usually a reaction given by the handler when the dog jumps up) and then make sure that this reward never follows the problem behaviour. This process is called “putting jumping-up on an extinction schedule”. Some behaviours are “intrinsically” reinforcing — just doing them is rewarding for the dog. If a behaviour is intrinsically reinforcing, it will not extinguish even though we put it on an extinction

schedule. Thus, if an anxious dog releases tension by barking in its kennel, it may not ever stop barking, even if the handler never goes to it when it is barking.

2. Continuous reward schedule (CRS). A reward is given immediately every time the dog makes a correct (or sometimes a near-correct) response. Assisting the dog to assume the desired position or behaviour is permissible (i.e. in the case of the sit, gentle pressure on the rump to encourage the animal to sit), but it is preferable to encourage the dog into the behaviour by baiting it with food or some similar technique. Inducing the animal to perform the desired behaviour independently and then rewarding the behaviour will produce more rapid learning than “pushing” the animal into position and then rewarding it for allowing this to happen. CRS is the most effective reinforcement schedule for the initial training of a skill.

3. Fixed ratio reward schedule (FRRS). A reward is given to the dog after it makes two or more correct responses. It is most useful to think in terms of ratio schedules of reinforcement when teaching behaviours that are “episodic” like barks and scratches. At the beginning of a FRRS, every second response is rewarded. When the dog consistently makes two responses to obtain a reward, then the handler requires three responses. By increasing the number of responses, one at a time, and training the dog to perform at each level with 100 per cent proficiency, a high FRRS can be achieved. For example, a fixed ratio reward schedule might be used to train a dog to bark or scratch repeatedly at a door to indicate the presence of a target odour inside a room. During initial training, the handler will open the door and give the reward after one bark or scratch, then two barks or scratches will be required, then three, and so on.

4. Variable ratio reward schedule (VRRS). To use a VRRS, the handler selects a range of responses (e.g. 5 to 10 correct responses) and rewards the dog on a random basis within this range. For example, say the dog has already learned to bark 15 times in order to obtain a reward on a FRRS. In a new training procedure, the reward appears anywhere between 5 and 10 barks, so that the dog never knows whether it will have to bark 5, 6, 7, 8, 9, or 10 times for the reward.

5. Fixed interval reward schedule (FIRS). Reinforcement is given to the dog after it responds to a command for a given fixed period of time. This procedure is used when teaching behaviours that are “continuous”, such as a sit- or down-stay, or walk-at-heel. In initial training, the dog is required to stay for only a few seconds. Then short periods of time (e.g. 5 seconds) are added to the interval, while requiring the dog to attain 100 per cent proficiency at each interval. For heeling, a trainer initially rewards the dog for just a few moments of correct heeling. With time and practice the handler gradually extends the period of time that the dog must walk at heel.

Gradual extension of the period of time a dog works for the reward is crucial in detection training. Initially, the trainer arranges a search problem for a novice dog so that it can easily find the target odour and obtain the reward in a short time, say less than 30 seconds. An advanced dog is required to work for 5 or 10 minutes or more, prior to finding the target odour. The emphasis in these exercises is not to force the dog to work for longer and longer periods of time prior to finding the target. Instead, the aim is to help the dog to exhibit highly concentrated and focused search behaviour for increasing periods.

6. Variable interval reward schedule (VIRS). To use a VIRS, the handler selects a time range (e.g. 1 to 2 minutes) and rewards the dog on a random basis within this time

period. For example, if the dog has already learned to hold a down-stay for 3 minutes on a FIRS, then the reward is given somewhere between 1 and 2 minutes on a random basis. The dog will learn that it must hold the down for at least 1 minute and perhaps for up to 2 minutes in order to obtain the reward.

Application of reward schedules

Normally, in dog training it is not necessary or even desirable to consciously apply first fixed and then variable schedules of reward. It is usually sufficient to follow the following general rules:

- When teaching a dog to give an episodic response (e.g. barking), the reward is initially given every time it barks (CRS), then gradually the reward is provided after longer sequences of barking (FRRS). If the dog shows hesitation or confusion, the number of required barks is decreased until the animal regains proficiency, and then the number is increased again. Immediately the dog understands the idea of giving voice repeatedly (barking perhaps five or six times), then a random number of barks in the range of three to five is rewarded (VRRS). Finally, the maximum number of times that the dog must bark for reward is progressively increased. The required number of barks for reward must not be predictable, and the dog should occasionally be rewarded for just two or three barks.
- When teaching a dog to perform a continuous response (e.g. searching for target odour), performance is established by rewarding it consistently after a very short period of time. Then the period of search time is slowly extended (FIRS). The duration requirement should be decreased if the dog's performance deteriorates. Once the dog searches intensively for a meaningful period of time (say 20 or 30 seconds), then it is rewarded randomly for searches of various durations less than 30 seconds. Finally, the maximum duration of searches is gradually increased, although the dog still searches for randomly varying periods of time prior to reward.

Advantages of variable reward schedules

Using fixed reward schedules the dog can learn to bark many times in succession or stay for several minutes. However, fixed ratio schedules of reinforcement (FRRS and FIRS) with high requirements (i.e. many, many barks or very long search intervals) do not produce a steady level of effort and motivation. Animals under high fixed reward schedules tend to “scale” their efforts. For example, a dog required to bark many times for a reward tends to bark lazily and intermittently at first (or even take a long rest after each reward), and then to increase its level of effort as it nears the required number of barks that will bring reward. Similarly, a dog consistently required to search for very long periods of time prior to earning a reward will tend to search half-heartedly until its “internal clock” tells it that a reward is nearly due. To avoid these effects, the trainer should not apply fixed schedules of reinforcement for too long. That is, as soon as the dog has learned to bark several times for reward, or to search steadily and with a high level of intensity for 20 or 30 seconds, then the trainer should switch to random schedules. In the case of detection, longer search periods must still be accompanied by occasional short reward intervals. Otherwise the dog will invest comparatively little effort during the first few minutes of a search problem, only increasing its focus (and accuracy) once it is well into the problem.

One advantage of random schedules of reinforcement is the production of uninterrupted and high-intensity responding from the very start of an exercise.

Variable reward schedules also encourage persistence in the face of extinction. That is, variable schedules teach the dog to be persistent and stubborn in trying to obtain its reward through instrumental behaviour, even when the reward rate is low. Many studies have shown that variable reward schedules produce more persistent conditioned behaviour than fixed schedules. A simple explanation of the variable reinforcement phenomenon is this: when the dog never knows how many times (or how long) it will be required to perform before being rewarded, it “loses track” of how many or how long and concentrates on performing persistently, convinced that if it tries hard enough it will eventually get the reward. This is highly desirable behaviour in a detector dog.

Application of compulsive training

Just as it is important to understand certain basic principles (such as reward schedules) in order to perform effective inductive training, it is also important to understand certain basic principles in order to use compulsive training effectively.

Use of physical punishment

Punishment is used to teach a dog not to do something. For example, a dog can be punished for breaking the sit-stay by administering a sharp jerk on the choke collar at the instant that the animal lifts its hindquarters from the ground. The aim is not to train the dog to do nothing, but encourage it to do *something else*, such as maintain the sit. There are four major principles the trainer must understand in order to use physical punishment effectively and humanely:

- First, the dog must have the alternative response in its *behavioural repertoire*. That is, the animal must “know how to” perform the alternative behaviour, and be in a state in which it can do so. For example, if a dog is breaking the down-stay because it is frightened of a jet engine, the dog’s fear may render it unable to avoid punishment. If a trainer physically punishes a frightened dog for not staying, the punishment is likely to make the dog even more afraid and less capable of staying. This is neither humane nor effective dog training.
- Second, physical punishment should be administered at an intensity that is meaningful to that dog (but still humane and defensible) and sufficient to cause it to change its behaviour immediately. Dogs adapt quickly to physical punishment and can learn in a short period of time to endure very uncomfortable events without altering their behaviour. If a very soft correction is used, and then gradually increased, especially very excited dogs intent on working their way to a reward may adjust to the punishment and ignore it. This procedure can unintentionally create a “monster”, a dog that can absorb excessive amounts of physical abuse without ceasing to exhibit the problem behaviour.
- Third, physical punishment should not be continued if it is not working. For any number of reasons, the dog may be incapable of the alternative behaviour. For example, a history of bad training may have rendered humane and reasonable levels of physical punishment ineffective. It is inexcusable to continue to physically punish a dog without specific and achievable training objectives in mind. In addition, some working breeds (e.g. German shepherds, Malinois)

readily exhibit handler-aggression when subjected to painful corrections that they do not know how to avoid.

- Fourth, emotionality on the trainer's part compromises the use of physical punishment. That is, an angry or frustrated handler may lose the technical ability to control the timing and the intensity of inputs to the dog, and make mistakes as a consequence. Revenge or temper have absolutely no place in working dog training.

Use of negative reinforcement

Negative reinforcement involves rewarding behaviour by withholding compulsion. The classic example in working dog training is the “out”, in which the dog releases an agitator or a reward object (e.g. a rubber ball) on command. Although a handler uses any available positive reinforcement to reward the dog for releasing cleanly (e.g. praise, immediate re-bite, etc.), the “out” is normally taught and maintained principally through the administration of negative reinforcement. Thus, if the dog releases cleanly on command, it is NOT corrected with a jerk on the choke collar. All four of the principles stated above with reference to physical punishment also apply to negative reinforcement. In addition, the following terms and definitions should be understood.

Escape training. Escape is an initial stage of negative reinforcement training during which the dog learns to end or stop a state of discomfort by executing some target behaviour. For example, during the escape stage of training, the command “Out!” is meaningless. The dog does not yet understand that “Out!” means a collar correction will occur if it does not release. Therefore, on the first trial, when the handler gives the “Out!” command the dog will probably continue biting. The handler then applies a collar correction until the dog releases the bite, praising the dog once it has released. The next trial or two will proceed in much the same way, but the dog *is* learning. During this stage the dog learns to expect the correction when it hears the command “Out!”, and it also learns to “turn off” or terminate the correction by releasing the bite. This escape learning is very important. A dog that does not know precisely how it can “turn off” compulsion will be confused and stressed by corrections, and may engage in inappropriate behaviours to try and terminate discomfort, such as avoiding or biting the handler.

If a complex response like walking at heel or recalling to heel is trained using negative reinforcement, there must necessarily be a stage during which the handler teaches the dog to terminate collar corrections by placing itself at heel. If the animal does not know which behavioural response will terminate compulsion, then collar corrections will only make it move more and more strongly away from the handler. It is therefore essential to use reward-based learning to patiently teach the dog skills *before applying escape and negative reinforcement techniques* — to make sure that the dog knows how to perform the desired behaviours on command. The dog will then be ready to learn quickly how to terminate compulsion by executing a commanded behaviour with minimal stress or confusion. Quick learning under compulsion will minimise the amount and intensity of physical force required for training, and help to render work a pleasure for both handler and dog.

Avoidance training. Avoidance is the next stage of negative reinforcement training during which the dog learns that in addition to terminating compulsion by executing the target behaviour, it can completely avoid compulsion. In the example of the out, if the dog releases the bite quickly on command, the collar correction will never occur.

When avoidance is completely and cleanly taught, every time the dog releases on command it is rewarded by the absence of the correction.

Criterion avoidance. The end goal of negative reinforcement training is to secure a correct response to the command every time, without the need for compulsion. In working dog training, this goal has the additional dimension that the handler intends to eventually discard the means of compulsion (e.g. choke collar and leash) completely. For example, a fully-trained dog should release a ball cleanly on command when it is not wearing a choke collar or leash, and when the handler is 3 or 4 metres away. The handler will therefore not discard the means of compulsion until the dog has achieved a good avoidance *criterion*, such as consistent response to command over at least four or five training sessions. During these error-free training sessions the handler will be ready to correct the dog instantly, with all necessary equipment in place. If the response is not consistently offered without the need for compulsion, then the dog must continue to practice with the handler standing by ready to enforce obedience, until criterion avoidance is obtained and extensively practiced.

Misuse of escape training. Repeated and prolonged physical correction of a dog to force it to carry out exercises is a sure sign that positive learning is not taking place. For example, if the choke collar is used ten times in a row to force the dog to release the bite when the “Out!” command is given, then there is no progress and the dog will never learn to out this way. The escape phase of learning in any given exercise should be extremely brief — between one and, at most, four repetitions of the exercise with corrections applied. This is quite enough for the dog to learn how to “turn off” the correction. Then it should be moved smoothly and efficiently into avoidance training, where it releases on command (rather than when the correction is applied) in order to avoid the correction.

Supporting negative reinforcement with positive reinforcement. Although behaviour learned through negative reinforcement training can be very durable and reliable, negative reinforcement should be followed by positive reinforcement whenever possible. The trainer can give the dog pleasant rewards to work for *in addition to* the reward of not being corrected. For instance, after a clean, fast out from the ball, the dog is praised and allowed to re-bite the ball and play with it.

Relationship between classical conditioning and instrumental conditioning

Very little working dog training involves the deliberate creation of classically-conditioned associations, as in the example of the “No!” command above. Most of the “action” in dog training has to do with the use of reinforcers and punishers to produce instrumental conditioning. However, classical conditioning processes are relevant because they are in the background of most training experiences for the dog.

Classically-conditioned associations contribute to training in countless ways. For example, if a handler gives a detection dog a preparation cue (e.g. “Ready to work? Ready to work?”) prior to taking the animal into a search area and giving the search command, the dog will associate this cue with the opportunity to search for target odour and the excitement of receiving a ball. The dog will soon begin to exhibit excitement and arousal responses to the preparation cue, encouraging enthusiastic search behaviour. In this example of classical conditioning, the preparation cue is the CS, the opportunity to play with a ball is the US, and arousal responses are the CR. This classical stimulus-stimulus association (between preparation cue CS and ball

reward US) lies in the background of the instrumental response-consequence association (between search/detection behaviour and ball reward) used to teach the dog to search for target odour. Here, the ball serves as both the instrumental reward for a correct sit, and as a US that classically conditions stimuli (such as the preparation cue) associated with training.

However, classically-conditioned associations in the background of instrumental procedures may also hinder learning. For example, if a dog sits too slowly in response to command, the handler may decide to hasten the sit by applying physical force. The handler gives the command “Sit!”, followed by a strong jerk upwards on the choke chain when no sit response is given. The intention of demonstrating the unpleasant consequence of sitting slowly actually constructs a very effective classical conditioning trial — “Sit!” is immediately followed by a sharp jerk on the collar. Soon “Sit!” develops the power to trigger responses appropriate to a sharp collar correction. Such responses can range from pain-elicited aggression and biting, to avoidance and cowering. More commonly, they involve anxiety and a reflexive stiffening of the body’s muscles to defend against the sharp jerk on the collar. The physical stiffening hinders the dog’s ability to sit quickly, with the result that it receives yet another jerk on the neck, which makes it even more anxious and stiff when it hears “Sit!”, and so on. In this example, the classically-conditioned *stimulus-stimulus association* between “Sit!” (CS) and a collar correction (US) is superimposed on the instrumental *response-consequence association* (a quick sit will result in no collar correction) that the handler has in mind. Thus the “Sit!” command will begin to elicit involuntary responses from the dog related to the discomfort of a collar correction, and these involuntary responses will interfere with the dog’s ability to sit quickly and smoothly on command.

Generalisation of classical and instrumental conditioning

Generalisation is a process in which behaviour that is learned in response to one stimulus is expressed to some degree as a response to another stimulus. Generalisation takes place with both classically-conditioned and instrumentally-conditioned behaviours, and the more similarity there is between two cues or situations, the more generalisation there will be. Two examples of undesirable generalisation: a dog that has learned a strong startle response to the “No!” command may also startle when the handler says “Yo!” loudly to a friend; a dog that has learned to sit in response to one explosive odour, such as ammonia dynamite, may also sit in response to a similar non-explosive odour, such as ammonia-based house-cleaning liquids. Generalisation can also be useful, for example if a dog trained to sit in response to the odour from an explosive device made with one brand of smokeless gunpowder also sits in response to an explosive device made with a different brand of gunpowder.

Generalisation of behaviour learned in one context to another context

Trained behaviours are not just controlled by CSs (classical conditioning) and cues/commands (instrumental conditioning). They are also controlled by context, a word psychologists use to label all other environmental stimuli present in the conditioning

situation. Context participates in learning, and generalisation from one context to another is rarely perfect. For example, if it is obedience-trained only in the context of an isolated training field out in the country, a dog will probably perform poorly the first time it is “run through” an obedience routine in a busy city park with people crowded around. Likewise, a dog that learns search and detection behaviour in a warehouse may also perform when it is taken to an office building, but its performance is liable to be poor until it gains experience in the new context. The best way to ensure that trained behaviour is independent of context is to train in as many different places and situations as possible, once the initial learning phase has been completed.

Transfer of learning

Transfer of learning takes place when one learned skill or command affects learning of another skill or command. Transfer can be positive (favourable) or negative (unfavourable).

In negative transfer of learning, one learned response interferes with the attempt to train another. For example, if a dog has previously been encouraged to claw and scratch and bite through a barrier in order to reach a ball (a technique commonly used by trainers to build “drive”), this response may transfer to explosives detection training, making the dog more likely to claw and scratch at the source of a target odour.

In positive transfer of learning, one learned response helps the dog to learn another. For example, learning to sit in response to the “Sit!” command during obedience training transfers positively to detection training by helping the dog to learn to sit in response to odour.

A skilled handler minimises opportunity for negative transfer of training to take place, and maximises opportunity for positive transfer of learning to occur. Efficient dog training can be viewed as a sequential, cumulative process in which the learning of one skill assists the learning of the next skill in the sequence.