

workingdog

Issue 3 | May/June 2017 | \$9.95



**Creating
an Effective
K9 Unit**
Part 1

2017
workingdog
CONFERENCE

MEMPHIS

**FLSA and the
K9 Handler**
Canine Care &
Handler Compensation

Detection Training

WHY DIRECT REWARD METHODS ARE SUPERIOR TO MODERN INDIRECT REWARD METHODS

By Jerry Bradshaw & Aaron Kemp

After WWII and the successful use of canines in the war, training protocols for military dogs were developed and written. These protocols were largely based on compulsory principles, employing force, physical prompts, hindquarter manipulation, negative reinforcement, and handler guidance. In detection, the same principles were largely employed. The dog was brought to an apparatus, such as a box, and encouraged to inspect it through handler presentation. Once the dog sniffed target odor, the handler then manipulated the dog into a sit through a combination of verbal and physical means. Upon the completed sit, the dog was rewarded either directly at source or indirectly at the handler, depending on the trainer and program variation. Once the dog was sitting for the box with odor, additional blank boxes were added, and the location of the target box was varied. The dog learned to sit for the box that contained odor while passing “blank” boxes, possibly containing proofing items, such as distractors. From here, the search area was gradually expanded, placing the boxes within search areas and transitioning off them.

Over the course of decades, this training protocol has been modified and adjusted in countless ways. Most training programs today are based off it with varying degrees of physical prompts, markers, delivery methods, and apparatuses. Trainers may claim their program is “new” or “different,” and indeed it may appear so based on the look of the equipment. However, the essential elements of the program that use handler proximity during initial training and context-specific learning (boxes, tubes, and drug walls, for example) also lack hunting development early on. These methods teach odor recognition and final response at the outset and bring in hunting later in the back-chained style of an obedience sequence.





Photos by Dawn Rabinowitz
caninesonduty.com

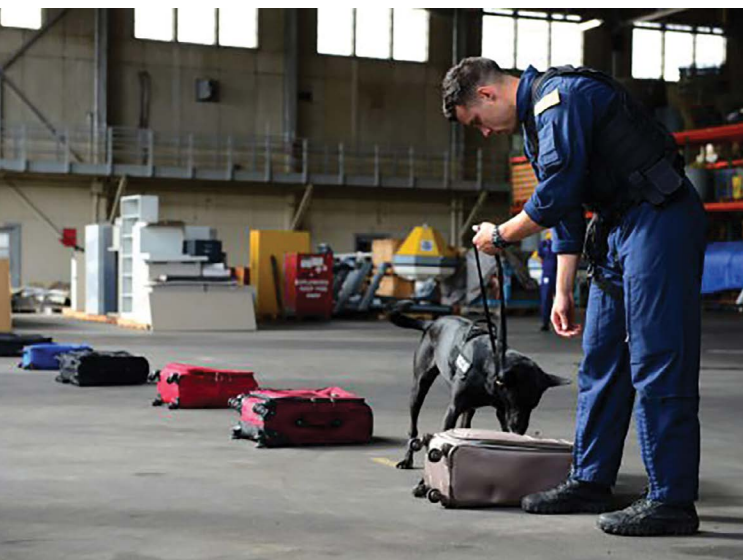
The old military style of training detection, commonly referred to as the reward not from source (RNFS) method, which uses handler guidance, physicality, and context-specific training is still employed. As the name implies, it's an **indirect reward** method of training, delivering the reward **at the handler** rather than **exclusively** at source while teaching odor recognition and a final response. Some would argue that variations of the method reward the dog "at source," but that is a bit misleading, as the handler is almost always present at source restricting or rewarding the dog, or an assistant rewards the dog, so for convenience here, we lump it all into indirect reward. As we will show, the problems associated with RNFS persist in the newer versions.

By teaching these two lessons (odor recognition and final response) simultaneously, we risk weakening the dog's natural hunting ability. The dog's associations with odor from the first encounter are in a low-drive state and in the context of a contrived and artificial apparatus meant to limit the hunting to, at most, four choices initially. It's essentially scent-associated obedience. Handler dependency and context-specific performance are instilled in the dog during the formative stages of training. As a result, many dogs never achieve an acceptable level of hunting and drive stamina at the conclusion of the training program. Starting with a dog that already hunts well and sticking him on boxes for odor recognition, final response, and proofing distractors dramatically changes the problem the dog must solve, and the context in which we want the powerful hunting can for some dogs be forever changed.

This is not an imagined deficiency. In fact, the U.S. military identified this problem of detection dogs failing to hunt adequately at the conclusion of training and changed their operations back in the mid-2000s with the adoption of clear signals training (CST) and deferred final response (DFR) methodologies (**reference, for example, AFMAN31-219 30 JUNE 2009**). They abandoned RNFS protocols because they observed the degradation of the hunting quality and wanted to address it to make better, more deployment-ready explosives and narcotics detection dogs. CST training adopted a defined use of marker training for obedience primarily, and DFR introduced a direct reward system for detection and deferred final response. This change engaged a more natural way of training detection by enhancing instinctive predatory behavior and channeling it into a development of hunting prowess and more powerful odor recognition. Today the quality of detection dogs in the U.S. military is second to none.

The belief is that other components of a successful detection canine, such as conditioning a level of drive for the task, hunting quality, task-specific stamina, and developing strong sourcing behavior, require development. The training setups foster distance from the handler in the formative stages of learning to develop independence and reduce handler-induced false responses. The final response is trained separately, away from detection, and progressed to teaching the dog to sit in heightened levels of drive. Once the dog is fluent with sitting in drive, that skill is brought into the context of detection when teaching the final response, allowing for the least amount of compulsion and handler influence around odor and minimizing the time the dog spends on apparatus, such as boxes or walls, to teach the final response at source. This makes it easier to fade out the hunt-drive-killing apparatus.

Many progressive training companies, such as my own, have used a direct reward method since we began offering trained detection dogs in the mid-1990s. We present the information in this article honestly having experimented with changing our system in 2009 to indirect reward to see if we could turn out bomb dogs more quickly. This experiment, while admittedly anecdotal for our company, showed us that the method of direct reward that we had employed turned out dogs that were much better hunters, stronger at odor source, and gave better,



“Hide placement teaches the dog to hunt with enthusiasm and precision. The handlers and assistants allow the dog to find the pipe through self-discovery.”

more readable changes in odor. It is a problem of perception that dogs who have odor recognition and sit on a box with contraband are “detection trained dogs.” They are not. In fact, the real power of the detection dog is that he can **take us** to the source, not simply sit on odor as an obedience exercise!

Further, we have not turned to the reinvented RNFS methods that now come with the newly fashionable addition of clickers or verbal markers. Our system, which differs somewhat from the military system, still uses a DFR, choosing to emphasize the hunting action pattern up front. We use odor in PVC pipe as a vehicle to deliver both reward and target odor simultaneously to the dog. The dog is trained first to retrieve the odor-filled pipe, then hunt it out in tall grass to reduce any visual cues, and hunt indoors in areas filled with luggage and boxes. This basically performs the same feature as tall grass but does so on slippery floors and in tight spaces, exposing the dog to environmental challenges. Then we use a system of initiating a high state of drive by holding the dog back on a flat collar, and a second trainer firmly taps the pipe on items around the search area (a room filled with furniture or a set of vehicles), exciting the dog with quick movement and the sound of the tapping. We stealthily hide the pipe as we are tapping around, continuing to pretend to put it in multiple cracks, crevices, and gaps in the furniture or seams of the doors and wheel wells of

the vehicles as the dog watches. This removes all chance of visual cueing and allows the dog to independently search out the odor-filled pipe. Hide placement teaches the dog to hunt with enthusiasm and precision. The handlers and assistants allow the dog to find the pipe through self-discovery. The activation by an assistant tapping is faded out in the training process so that the handler takes over the job of initiating the hunting sequence. There can be multiple people in the search area; the dog soon learns to ignore them. When he approaches, they pay no attention. The dog learns that only continued hunting and productive areas yield the reward. Height, depth, blank areas, and permeation time teach the dog how to locate the source of the odor, at which, his reward is present. There is no false responding initially because the reward exists only at the source, and final response is deferred until the dog is source focused. The premium in this method is on the hunting stamina, learning to hunt productive areas, functional independence from the handler and other people in the detection area, and to eliminate handler-induced errors (the dog takes the handler to odor source rather than the handler presents the dog’s way to odor source), as well as odor recognition. Variation of amounts and permeation times teaches the dog to hunt out different thresholds and get to the source, also virtually eliminating fringing behavior from the outset.

After a while, the pipes can be eliminated and the dog paid at source for locating the source by a secondary reward of “tossing” the reward over the dog’s head from out of line of sight to land at the source area as the dog stares at the hiding place or some light aggression at source. Both are easily eliminated and are a minor problem compared to handler cueing, fringing, and false responding behaviors.

It is critical that in this direct reward method of training, the dog have sufficient repetitions of hunting and the hunting areas be varied in size and type. The dog should go into the final response training with a well-developed action pattern derived from the hunting exercises and the clear association of the target odor with the reward item. The dog should have had sufficient repetitions to make him neutral to changes in flooring, room size, search area size, noises, activity (human and canines moving in and out of the detection area), lighting, and environment – indoors and outdoors. He should associate buildings, outdoor areas, and vehicles with one thought: to hunt with maximum effort and efficiency to locate the target odor. Direct reward then creates an intrinsically rewarding action pattern. The exciting thing is that these patterns are genetically present in the dogs we select for detection work, if only we use these action patterns to our advantage.

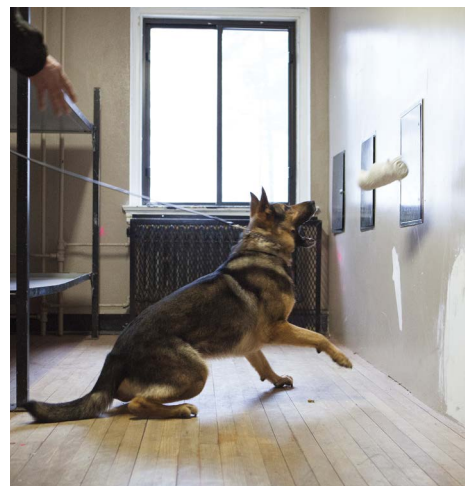
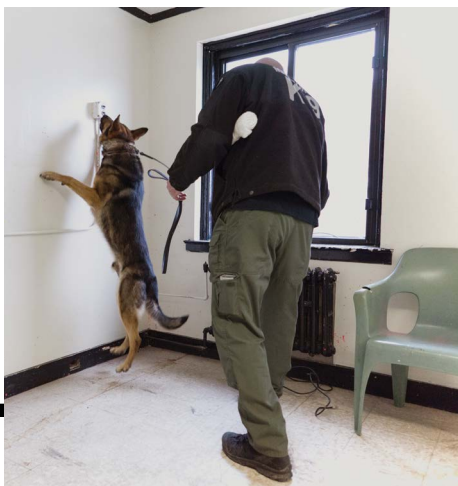
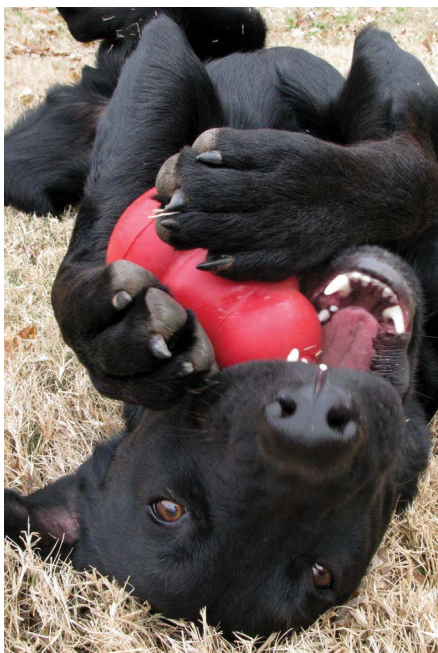
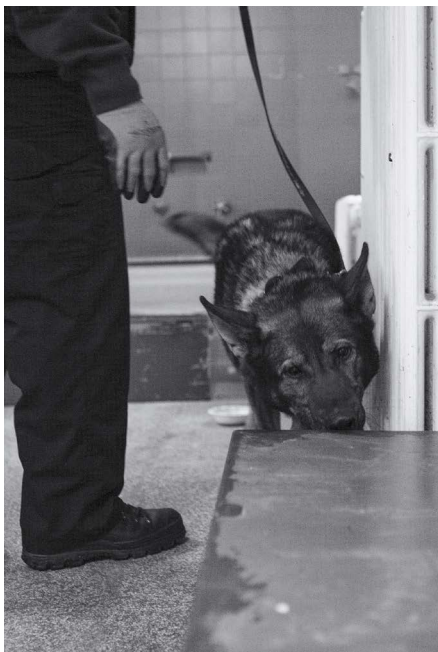
As trainers, we know that hunting behavior is a sequence of multiple behaviors interconnected in a chain-like fashion. These behavior chains are genetically hardwired, and the active rehearsal of these instinctive action patterns through engagement of predatory behavior awakens this within a dog. Repetition is a key factor in developing these action patterns, as we not only want to “bring it out” of the dog, but we also want the dog to learn the self-rewarding value of the action patterns.

A great example of intrinsically rewarding action patterns are dogs that chase cars, squirrels, cats, and other prey. The more repetition a dog gets while engaging in predatory behavior, the harder it is for a trainer to inhibit, as the act itself becomes increasingly enjoyable. The motivation to rehearse this behavior progressively outweighs the cost we can attach to it. However, if the behavior is caught soon, before the behavior can be rehearsed repeatedly in the context of a car, squirrel, or cat, applying aversive stimuli can change the contextual association the dog has with that visual stimulus. As with all forms of punishment, the trainer has to be present to apply it. The dog learns that the combination of the trainer, the cat, and himself is a bad mix and that punitive measures are possible and anticipated. A reduction in drive (a calming effect) will likely be noticed. This is why the benefits of creating intrinsically motivating action patterns don't take hold in traditional RNFS odor recognition and final response trials. The conditioned association with target odor is intertwined with the inhibitory effects of obedience and low-drive states and, sometimes, corrections and physical manipulation.

While instinctive behavior chains for hunting are fixed to a certain degree, proper training can increase the efficiency and duration of hunting. Not only will the dog gain immense pleasure from rehearsing the natural sequence, but each behavior in the sequence will become a conditioned stimulus triggering the next behavior in the chain. For an instinctive action pattern to be conditioned, we have to engage the dog in the correct drive state for the desired action sequence to be elicited. Canines are hunters by nature. Our job is to condition the dog to use his nose in prey drive (in the context of detection) to locate his quarry (prey object) by means of following his hunting instinct and then in the proximity of trained target odors to find the source. The direct reward sequence of hunting to find a toy conditions this sequence.

The belief here is that prey drive itself consists of complex species-typical sequences and routines. Understanding this and providing training setups conducive to the expression of the drive-related behavior, we actually uncover entire behavior sequences. The dog will show an outward expression of drive when we present a conditioned stimulus, such as a Kong or PVC pipe. The dog will orient himself to it, eye stalk it, chase it, hunt it, grab it, and bite it. These are the instinctive species-specific routines that we call the fixed action pattern, which a canid in the wild would do with prey it is hunting for food. The expression of these behaviors efficiently and purposefully is the result of being able to rehearse them. Wild canid puppies don't “learn to hunt” as much as the mothers and fathers provide opportunities for the pup to rehearse an inherent fixed action pattern to perfection. For our purposes in detection, we are simply taking a natural sequence of behavior and changing the environmental context in which these patterns of behavior are rewarded. Think about a dog you have seen who in the middle of hunting stops to relieve himself and once done just continues to hunt, or a dog who hunts obsessively and even after receiving his reward at the end returns to hunting after spitting out the reward, or one who at source is given his reward item, and he prefers staring at source. The action pattern of hunting is powerfully motivating for these dogs.

The curriculum of indirect reward, however, is focused on establishing behaviors required for detection rather than establishing an instinctive action pattern. Much like drive-based obedience, modern indirect reward systems teach behaviors motivationally with a high level of engagement and focus. The process then chains those behaviors together. These motivational shaping methods for indirect reward that start with teaching final response and odor recognition share a common thread with simple obedience exercises. For example, consider shaping a “send to place” in obedience. The objective is to create a meaningful association with an object, such as a dog bed, and shape an approach to the object and a sit at or on the object. This is all done through back-chaining. The dog is taught the last step in the chain first, sitting, and a meaningful association with the object by having the dog rewarded for getting on the dog bed and then sitting on the dog bed. The trainer then builds distance by sending the dog to the bed at



“Putting it bluntly, a K9 team has more important matters to attend to than marker timing and effective reward delivery.”

first from one step away, backing farther and farther from the dog bed, building up to as much distance as the trainer wants. The dog goes to the bed and sits. You will see the dog at first doing these sequences slowly, and as the dog rehearses and pieces together the patterns, speed and fluency increase. As the dog practices, he gets quicker. The trainer can then proof the dog to other objects. In a multi-dog household, the dog could be proofed off of other beds, locating the one with his scent.

In the case with odor detection, the dog must locate the object containing source among other identical “blank” objects. Once the dog knows what to find, in both scenarios the trainer now increases the distance from the object, visual and non-visual cues, and varies environments, generalizing the behavior while gradually increasing environmental distraction.

While this training progression works, it doesn’t produce the same level of quality detection as direct reward in our experience. When applied to detection, odor recognition and a final response are taught simultaneously as the last sequence in the behavior chain. The dog has to understand first that odor predicts reward and then what to do once he encounters it. In direct reward, the behaviors we extend from the terminal behavior (sitting on odor) are based on hunting duration in the dog’s mind rather than a classic obedience behavior chain that can be observed in a recall sequence or a “send to a dog bed” sequence. Real-life detection relies heavily on continuity throughout the behavior chain. If the behavior chain is broken due to a dramatic increase in environmental distraction, the continuity of the sequence is broken. The commonality between training a behavior chain through “artificial” means, such as back-chaining, and training it through instinctive drive is that each behavior elicited in the sequence must occur in the correct order and trigger the next appropriate behavior. Indirect reward shaping methods produce a behavior chain vulnerable to derailment through the presentation of not only startling stimuli but also salient competing motivators. This issue is further compounded by the program’s lack of environmental distraction incorporated into formative stages of training. When starting a dog in a direct reward format, however, the

dog is stimulated to a high level of focus and drive as it hunts, teaching it to ignore competing motivators, as well as stress inoculate to surface changes or other changes in the environment much more easily.

Direct reward systems create a self-rewarding behavioral sequence that isn’t based off of providing a single reward at the end of the sequence but rather by making hunting itself motivating. It creates a stronger behavioral sequence and dramatically increases the likelihood of the dog re-tasking himself should the continuity of the sequence be disrupted, as in the examples previously noted.

Direct reward also fosters easier maintenance training for new and advanced handlers alike. Unlike indirect reward shaping methods employing the use of a verbal or mechanical marker, the handler doesn’t have to attend to multiple “moving parts.” The handler starts the dog in the hunting sequence, and the dog runs pretty much on autopilot until he locates the target odor. The timing required to mark desired behavior effectively and the finesse required to maintain a solid stare and independence while maintaining a reward history from the handler is a tall order. Putting it bluntly, a K9 team has more important matters to attend to than marker timing and effective reward delivery. The simpler it is, the more effective the team will be. Direct reward systems have the downside that the dog is actually so engrossed in hunting behavior it can ignore the handler and become a little too independent, and the drive state is so high the dog prey-locks at source. However, if you have done the training well, the dog should effectively, at a minimum, stop on source and stare obsessively. He’s painting an easy-to-recognize picture at the conclusion of his hunting sequence. These, in our opinion, are easier problems to solve than the dog that is distracted in hunting, not hunting in the highest drive state for that dog, or lacking hunting stamina. This is especially an issue for explosives detection dogs that must hunt an extreme number of blank areas without losing interest and have the stamina to wait for their rewards in operational deployments. While shaping methods certainly provide the trainer with an elegant method of training, they cannot compete with direct reward on some of the most critical elements of detection quality.

Conclusion

When selecting detection dogs, most trainers, irrespective of method, test the dog similarly for prey drive intensity (how strongly he orients to the object), how quickly the dog runs to the object (chase speed), how forcefully he grabs the object (grab bite), and how possessive he is (kill bite). Further, the trainer tests the dog for hunting duration and intensity, how well he uses his nose to hunt relative to his eyes, and how he does this both outside and indoors. Notice that this almost universally accepted testing process is actually the exact framework of a direct reward method. In direct reward, we take dogs that are facile in these instincts and rehearse them to our purpose, deferring the final response to a later date.

Another set of issues to think about are the commercial viability of detection dogs. Many trainers are getting famous teaching puppies to respond on odor. However, these pups have yet to prove they have the quality action patterns genetically to hunt to standard, so why waste time training odor recognition and final response? We don't even know if these puppies are medically viable. This is not limited to puppies. Why train odor recognition and final response on a new detection dog first if he may wash out of the program due to hunting stamina and environmental soundness, which are the main reasons a dog is washed from detection aside from medical? If we concentrate on hunting up front, and the dog doesn't make it, we didn't waste so much time training final response and odor associations, which are way more time-intensive skills. If we do this up front rather than spend all this time teaching final response before ever expanding the dog's tiny, context-specific search area to real working areas, we are not wasting as much trainer time on potential washouts. The fact is, we never fail a dog out of a program for the final response because it's an obedience exercise. All we have to do is figure out what variables we have to change to teach the dog more effectively.



Training a superior detection dog relies on enhancing and building upon already present hardwired action patterns. This is not a sequence of behaviors that has to be built from scratch, quite the opposite, as performance actually degrades with an increase in trainer influence and increased handler responsibility. Training fads come and go, and the current fad is using a clicker in detection and indirect RNFS methods. While we are by no means against clicker training, as we both employ it for obedience, we are against the use of clickers and shaping as palatable terms in the active resurrection of an outdated and substandard training program. RNFS with the use of shaping is procedurally the same as the old-school military style of training, minus the use of compulsory methods at source. The multitude of similarities remaining between them have been noted in this article. Something old isn't always new again with the addition of window dressing.

Jerry Bradshaw is the Canine Training Director of Tarheel Canine Training, Inc. in Sanford, NC. Tarheel Canine's School for Dog Trainers holds police K9 instructor courses for police K9 trainers as well as civilians. Tarheel Canine trains dogs for police departments worldwide.

Email: jbradshaw@tarheelcanine.com

Aaron Kemp attended Tarheel Canine's Master Trainer's Course. Upon successful completion, he was invited back for an internship teaching students and training dogs for police detection and patrol application. He is now the founder and head trainer of Superior Canine Training Inc. in his home province of Abbotsford, British Columbia Canada. Aaron competes in two protection sports, PSA and IPO, and has close to a decade of experience training dogs in obedience, protection, and patrol for security application.