The Whole Truth About Variable Reinforcement Schedules

Christina Alligood, Michelle Skurski, and Angela Miller
Disney's Animal Kingdom®
Lake Buena Vista, Florida

Over the past few years, we have developed and conducted a series of workshops for animal care staff on topics at the intersection of the science of behavior and the practice of behavioral husbandry. Our hope is that, over time, the connection between these areas will become seamless.

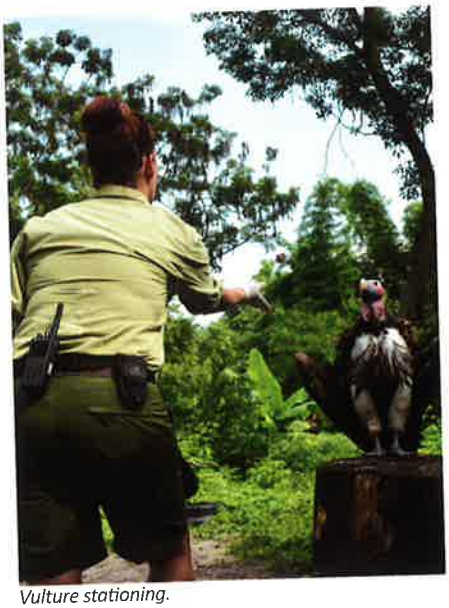
The first workshop we developed was on variable schedules of reinforcement. We became interested in this topic as it relates to animal training after seeing Ken Ramirez's presentation at the International Marine Animal Trainer's Association Annual Conference in 2014. Ramirez's careful explanation inspired us to broach this topic with our animal care staff in the hopes of disentangling some concepts that seem, often, to get jumbled together. The resulting workshop became the first in a series in which we discuss the science behind a training concept, its typical use in animal training, and how we might improve our practices. In this case, the take-home points that we shared with our animal care staff were:

Gorilla with infant

The "variable schedules" you are using may not actually be variable, or even intermittent. Even if you are using true intermittent schedules, you may want to reconsider when and why you use them.

Schedules of Reinforcement

From a scientific perspective, schedules of reinforcement fall into two broad categories: continuous and intermittent. In a continuous schedule, every correct response is reinforced. A correct response is a response that meets the trainer's criteria. We assume here that the trainer is only reinforcing correct responses. The criteria might change over time as a response is shaped, but any response that meets the current criteria is a correct response. In an intermittent schedule, some correct responses are reinforced, and others are not. Intermittent schedules can take several forms, and are typically identified using a two-word descriptor. The first word describes how the requirements for reinforcement are scheduled (e.g., fixed, variable, random), and the second word describes the type of requirement (e.g., time, ratio, interval, duration). For more information about



Vulture stationing.

some of the most common schedules, including examples, please see the AZA/AAZK Animal Training Terms and Definitions at https://tinyurl.com/trainingterms2016.

In this discussion we will focus on ratio schedules because they have been commonly used and commonly misunderstood in the animal training community. In a ratio schedule, access to the reinforcer depends on the number of correct responses. In a ratio schedule, a reinforcer might be delivered after one correct response, or three or five. When the number of correct responses required is held constant, the schedule is described as a fixed ratio (FR). An FR1 schedule is a continuous schedule of reinforcement. A fixed ratio schedule requiring more than one correct response per reinforcer (FR2, FR3, etc.) is an intermittent schedule. A schedule in which a variable number of correct responses is required is a variable ratio (VR) schedule. For example, in a VR5 schedule, a reinforcer might be delivered after as few as one correct response or as many as ten. The response requirement will vary between those two extremes in a carefully planned arrangement with an average response requirement of five. Over the years, trainers have sometimes

used the term "variable schedule" to describe techniques that do not fall into the definitions above. One example is the use of a variety of different reinforcers during a training session to keep the animal's motivation high. In many cases, trainers have also been able to reduce their use of food by offering different conditioned or secondary reinforcers after a correct behavior (Ramirez, 2014). In these cases trainers are using continuous reinforcement with a variety of reinforcers. Ramirez suggests that trainers use the term "reinforcer variety" to describe this practice, which is different from a variable ratio schedule. Another example of the misuse of the term "variable schedule" is when a trainer uses a bridging stimulus (clicker, whistle, etc.) after every correct response, but only "pays" (provides a backup reinforcer) after some correct responses. This is not a variable-ratio schedule and is not advisable. We will discuss this practice in more detail in a future column on conditioned reinforcement. When using a bridge, a variable-ratio schedule would involve delivering both the bridge and the backup reinforcer on some trials, and withholding both the bridge and the reinforcer on other trials in a carefully planned arrangement.

Free Operant vs. Discrete Trial Training

When trying to understand the general effects of variable ratio schedules, it may be useful to consider the type of procedure used to train a particular behavior. When discussing the typical effects of various schedules of reinforcement, we often talk about the rate of behavior. For instance, a variable ratio schedule typically produces steady, high rates of behavior and a fixed interval schedule typically produces a scalloped pattern where the rate of behavior increases just before a reinforcer becomes available and then declines sharply just after a reinforcer is delivered. Sometimes trainers attempt to use variable ratio schedules because the prospect of producing steady, high rates of behavior sounds attractive. But this effect is only possible in a certain type of training procedure.

In some training procedures, there is a period of time between the delivery of a reinforcer and the cue for the next behavior, during which the trainer does not offer reinforcement. If the animal offers a trained behavior during this period, the trainer does not reinforce it. This period is called an intertrial interval. In a free operant training procedure, there is no intertrial interval. After a reinforcer is delivered, more reinforcers are available if responses keep occurring. For example, think about a crow using a stick to extract grubs from a crevice in a log. The crow may or may not successfully obtain a grub every time the twig is inserted, so the schedule of reinforcement may be continuous or intermittent. But the crow does not have to wait before trying again; the stick can be inserted repeatedly as fast as the crow can move. This means that there is no intertrial interval and this is a free operant procedure. If grubs are available on a variable ratio schedule, the crow is likely to insert sticks at a steady, high rate. Free operant arrangements are very common in natural environments, and much of what we know about schedule effects comes from experiments using free operant procedures.

In a discrete trial training procedure, there is always a cue/discriminative stimulus. This is followed by a single response and an outcome (for example, a reinforcer). Then the next trial starts after an intertrial interval. For example, consider a husbandry training session in which the trainer asks a Komodo dragon to touch a target. The trainer gives a cue, the dragon performs the correct behavior, and the trainer delivers a reinforcer. Until the trainer gives the next cue, no further reinforcers are available for touching the target. This period between the delivery of a reinforcer and the delivery of the next cue is the intertrial interval. Even if the trainer uses an intermittent schedule of reinforcement (reinforces some correct responses and not others), the dragon's rate of responding is limited by the rate at which the trainer delivers cues. In this type of procedure, we can measure the percent of trials on which the animal makes a correct response. Rate of behavior will not be a useful measure in this procedure, because the trainer is controlling the rate of behavior via the intertrial interval. This is the type of procedure most often used in husbandry and medical training in zoological settings.

As we have worked with our teams on integrating science into our behavioral husbandry program, one thing that has emerged as critically important is the need to understand the reasons behind best practice recommendations. This gets us away from statements that include "always" or "never", and moves us toward customizing our techniques to individual needs and goals. Many animal trainers have heard that it is best to always use a continuous schedule of reinforcement when teaching a new behavior, and then to always transition to a variable ratio schedule to maintain the behavior. Often the reasons behind this recommendation get lost. When behavior scientists recommend transitioning from continuous to intermittent reinforcement, it is typically because the ultimate goal is for the behavior to occur without depending on the trainer asking for it and reinforcing it every time (Hanley & Tiger, 2011). For the majority of husbandry and medical training, that is not the case. When we train a tiger to open her mouth, a zebra to raise his hoof, or a gorilla to station for a blood draw, the ultimate goal is the successful performance of the behavior on cue to facilitate tooth brushing, a hoof trim, or sample collection. We do not expect or need the animals to demonstrate these behaviors without the trainer present to deliver reinforcers, so there is no clear advantage to using an intermittent schedule of reinforcement. To effectively maintain behavior, a variable ratio schedule needs to be carefully planned and implemented (Cooper, Heron, and Heward, 2007 p. 309). In practice, trainers intending to

Komodo Dragon target training.

use a variable ratio schedule typically implement a haphazard schedule instead. One serious risk of this practice is accidentally requiring more responses than the animal is prepared to offer for a reinforcer. This can result in unpredictable delays in responding, refusal to participate, and/or aggression (Cooper et al., 2007 p. 314). When there is no clear benefit to an intermittent schedule, we recommend keeping the training plan straightforward and sticking with a continuous reinforcement schedule. With this type of schedule, communication with the animal is clear and no special planning is needed.

There are some husbandry training situations in which intermittent schedules are useful. The benefit offered by a variable ratio schedule is that the animal is more likely to persist in performing the behavior when trainer-delivered reinforcers are not available. In some situations, this is an important benefit. For example, some female primates need assistance learning to tolerate infant nursing. While initially a trainer might work with the animal to train this behavior using a hand-delivered food reinforcer, the ultimate goal is for the animal to allow an infant to nurse without the trainer's involvement. To facilitate learning in this case, the trainer might gradually transition from a continuous schedule to an intermittent schedule of reinforcement. Each step in this transition would be carefully planned to ensure that the desired behavior is maintained. Ideally, as reinforcement delivered by the trainer becomes less frequent, natural reinforcers will become more frequent. Successful nursing interactions should produce an increase in oxytocin concentration for the mother. The intermittent reinforcement provided by the trainer will help the trained behavior to persist long enough to produce natural reinforcers that will keep the behavior going even without the trainer. Although there are certainly other situations in which careful use of intermittent schedules of reinforcement might be beneficial, in a typical discrete-trial husbandry or medical training procedure a continuous schedule of reinforcement is the most practical and effective strategy for clearly communicating with an animal and producing consistently accurate behavior. Sometimes trainers worry that using a continuous schedule of reinforcement will create an expectation that reinforcement will be delivered after every correct response, and if the trainer makes one mistake or is unable to deliver a reinforcer one time the animal will stop responding. It may be helpful to know that maintaining a behavior on a continuous schedule of reinforcement over a long period of time is another way to build persistence (Vollmer & Athens, 2011).

Conclusions

When discussing science-based behavioral husbandry, we do not advocate a hard-and-fast rule-based approach. There are very few absolute rules about what a trainer must *always* or *never* do to be successful. Instead we advocate continuing to learn about scientifically supported approaches and then tailoring behavior management plans to each individual animal's needs and challenges. For this particular topic, we conclude with a few brief suggestions.

Think about your ultimate goals for the behavior you are training. What is the purpose of the behavior? Under what circumstances will it need to occur? Will the animal need to perform this behavior when the trainer is not present?

Think about the type of procedure supporting the behavior. Think about the training procedure and the situation in which the animal will ultimately need to perform the behavior. For each of these, are reinforcers available in a free operant or discrete trial arrangement?

Try continuous reinforcement - it might work! If you are using a continuous schedule of reinforcement and your animal is disengaged, it may be worthwhile to re-evaluate some other aspects of your training before considering an alternative schedule. These could include but are not limited to the quality of your chosen reinforcer(s); timing and consistency of your cues, bridge, and reinforcer delivery; strength of your bridge.

Watch your animal's response. Be sure to define what success looks like in your training and watch your animal closely to monitor progress.

References

Cooper, J.O., Heron, T.E., and Heward, W.L. 2007. *Applied Behavior Analysis. Pearson Education, Inc. Upper Saddle River, NJ.*

Hanley, G.P. and Tiger, J.H. 2011. Differential Reinforcement Procedures. *In Fisher, W.W., Piazza, C.C., and Roane, H.S. (Eds). Handbook of Applied Behavior Analysis.* The Guilford Press, NY.

Ramirez, K. 2014. The paradox of the variable schedule of reinforcement: Aligning the science, the terminology, and the practical applications. Paper presented at the annual meeting of the International Marine Animal Trainer's Association. Orlando, FL.

Vollmer, T.R. and Athens, E. 2011. Developing Function-Based Extinction Procedures for Problem Behavior. *In Fisher, W.W., Piazza, C.C., and Roane, H.S. (Eds). Handbook of Applied Behavior Analysis.* The Guilford Press, NY.