Title: Dialogue not distraction: How to maintain long duration, voluntary participation with a dynamic reinforcement strategy.

Authors: Amy Schilz, Susan G. Friedman, Ph.D., and Rick Hester

Like all modern facilities, the animal care team at Cheyenne Mountain Zoo is in constant pursuit of effective, humane training strategies to benefit the animals in our care. One way we do this is by establishing a dialogue with them, whereby training decisions are made dynamically, moment by moment, based on what each animal's nuanced behavior, past and present, indicates they need to succeed. This contrasts with a monologue approach to training consisting of commands (demands, orders, and threats), and rigid adherence to schedules of reinforcement derived from the experimental analysis of behavior.

As the term dialogue is meant to convey, participants must have a shared system of communication to make their requests known (something wanted, needed or to end something undesirable), and we feel it is the obligation of progressive trainers to be responsive to those communications by changing what they do, while still meeting essential training goals. Trainers contribute to the dialogue with antecedent arrangements (setting events, motivating operations and discriminative stimuli) and consequences (mainly positive reinforcement) to shape new skills and maintain them once mastered. Animals communicate with trainers with approach, retreat, and delay (hesitance) behaviors, latency responding to cues, accepting reinforcers, and a myriad of small (and sometimes not so small) behaviors such as direction of gaze, hackles, tails, and shifts in weight distribution.

With a dynamic reinforcement strategy, the rate of reinforcement follows the animal's behavioral indicators of needing more or less encouragement to continue participating in a procedure that they show, or have shown, with their behavior is more or less demanding. In our experience, a dynamic reinforcement strategy is beneficial for mastered behaviors such as remaining on station during long or uncomfortable medical and husbandry procedures; however, when training the component behaviors that make up these behavior chains, discrete trial training (antecedent-behavior-consequence) is recommended.

For example, if we predict discomfort ahead (e.g., a knife will be used on a tender spot on a giraffe's hoof), we dynamically increase the rate of reinforcement before they remove their hoof from the block, making it worth their while to remain on station. When the discomfort ends, the rate of reinforcement goes back to their ordinary "pay-for-work." This approach is more nuanced than simply distracting our learners with a deluge of free food for at least three reasons. First, as mentioned above but it bears repeating, this strategy is most appropriate for fluent behaviors (and behavior chains) not for the initial shaping of component skills, which we recommend should be shaped with discrete trial training, sometimes called a "clean ABC" approach. Only after a husbandry and medical behavior chain is fluent do we then implement a dynamic reinforcement strategy, as needed. Second, the "just distraction" account falls short of the mark as it misses the subtle changes, or predicted changes, in the animals' behavior that a dynamic rate of reinforcement is tracking. Third, the dynamic reinforcement strategy allows trainers to complete critical-care procedures in a timely way.

This strategy has resulted in reliable participation in particularly long and/or uncomfortable husbandry and medical procedures such as voluntary stem cell treatments,

blood collection/banking, ultrasounds, and curative hoof care. In this paper, we discuss why, when, and how a dynamic reinforcement strategy can be used to maintain voluntary participation in fluent husbandry and medical procedures during those situations when participation is a big ask.

Prerequisite Skill Mastery Comes First

Of course, hoof care isn't one behavior. You can't shape hoof care per se. You shape (or capture) each of the component skills that comprise hoof care to mastery then assemble the hoof care behavior chain. At CMZ, the behavior "links" in the chain are discrete behaviors (aka movement cycles), trained with a discrete ABC approach and a continuous reinforcement schedule, i.e., (cue-behavior-bridge/backup reinforcer). The component skills we typically teach our giraffe include nose to target, back up, foot movement and placement, foot on block, curling over the fetlock, and stationing during the application of a variety of tool work on the foot.

For example, here is how targeting looks, as a discrete ABC contingency:

> **A**ntecedent: Trainer presents target,
> **B**ehavior: Giraffe touches nose to target,
> **C**onsequence: Trainer clicks and delivers cracker.
> **Prediction:** Giraffe will continue to touch nose to target for the cracker , given presentation of the target. (R+).

The benefits of continuous reinforcement (CRF) in discrete ABC trials are well known: Consistent reinforcement is the clearest way to communicate what behavior controls the reinforcement; CRF produces the most precise (least variable) topography (movement, stance, and placement); and CRF keeps our bridge strong by pairing it with a food reinforcer every time we click.

The component skill-building phase is also where the duration criterion is shaped, first discretely and also within the behavior chain. We continue using a CRF schedule when training duration with a changing criterion approach that bounces back and forth between longer and shorter criterion. For example, a giraffe may first be reinforced after touching nose-to-target for 1 second, with mouth and body still. Once this criterion is met without hesitation, the criterion would be increased by adding 2 or 3 second holds while still reinforcing 1 second holds periodically. When the 2 or 3 second hold is performed without hesitation, we stop reinforcing 1 second holds, and start mixing 4 second holds in with the 2 and 3 second holds. In this way, we slowly build duration over the repetitions, allowing the duration and correlated rate of reinforcement to change gradually. From the animal's point of view each repetition does not get predictably harder.

During the component skill training phase, we also teach the giraffes to hold still while we use various tools on their hoof. Each tool is named before it is applied so the giraffes can learn what tool is going to be used, further increasing the dialogue. The giraffe is reinforced (click/cracker) for holding still while the second trainer (called "the mechanic") uses the tool on the foot. Trainers are careful to only change one criterion at a time while we build duration of

tool use and appropriate pressure. For example, when using the rasp, the mechanic starts with light pressure on the hoof for a one second hold. Then we start building duration, changing only the length of time without changing the pressure, with the approach described above until we reach approximately 5-7 seconds duration without the giraffe showing signs of discomfort (tail swishing, head up, skin twitching). If we see those signs, we decrease the duration. Once we have 5-7 seconds of duration with light pressure, then we start increasing the pressure of the rasp for a shorter duration, and build up duration again with increased pressure. We can either build pressure, then duration, or vice versa, but not both at the same time. We trained holding for each tool (touch, brush, pick, knife, nipper, rasp, hoof boss) separately, with the goal of having still giraffe during the application of each tool on the foot.

It should be noted that hard pressure, digging out rocks, and working in sensitive areas of the hoof were reinforced with higher value or magnitude food items, even during the initial training. These repetitions were also never as long as the 5-7 second holds that we asked for, such as with light pressure brushing or touching. We watch our giraffe's behavior to dynamically guide our decisions regarding the duration and intensity of applying the aversive stimulus. If we see skin twitching, tail swishing, or some other behavior associated with hesitancy or discomfort we removed our hands and tools. We call these discomfort behaviors "tells." Tells indicate to trainers the that the next repetition of the tool needs to be either a shorter duration, done with less pressure, or moved to a different spot on the hoof. Anytime we see a tell, we remove the stimulus and neither bridge or deliver food. If the giraffe stays still during a long or uncomfortable segment we bridge, remove hands and reinforce with food. This allows us to develop a communication system with the giraffe. They learn two contingencies: 1) Staying still produces a food reinforcer; or, 2) skin twitching, tail swishing, etc. removes our hands/tool. The behaviors that communicate "Remove your hands and tool now!" are different for each giraffe. Our goal is to respond to the the smallest discernable behavior indicating hesitance or discomfort so that big behaviors are not necessary. Here are the two contingencies in ABC format:

Giraffe Tell:

Antecedent: Knife scrapes foot,
Behavior: Tail swishes,
Consequence: Knife comes off foot.
Prediction = Tail will swish more to escape the knife – Negative Reinforcement (R-)

Giraffe Stays Still:

Antecedent: Knife scrapes foot with less pressure in a different spot,
Behavior: Tail stays still,
Consequence: Keeper delivers click and cracker.
Prediction= Tail will continue to stay still to get crackers – Positive Reinforcement (R+)

We use the "tells" to provide the giraffes with an escape route during training. Every animal has a right to say "no, stop, I'm uncomfortable", but we get to help decide what that

behavior looks like. We would much rather remove our hands when we see an ear twitch or a tail swish than when we see a hoof kick. We can provide the tail swish behavior with the same function that a kick would have - to remove or escape the aversive stimulus (our hands, the knife, rasp, nippers, etc.). By catching the least discernable behaviors indicating hesitation or discomfort, we avoid escalation of "no" communications, resulting in a safer work environment for both the trainers and the animals.

Creating a Dialogue

As part of the dialogue, we also developed a system for the giraffe to let us know when they are ready to receive cues, i.e, "Yes! Let's do it!" If the giraffe's head is up, turned away, and not making eye contact, we do not cue any of the tool-on-hoof behaviors. Thus, the giraffes learned that if they brought their nose to the target, the opportunity to earn reinforcers from us would begin. In this way, they initiate the ABC trial by bringing their nose to the target stick, after which the trainer gives the hoof-care cue.

<insert "Proceed" video here. >

Using both the "yes, proceed" signal (nose on target,) and "no-tells" (aversive stimulus removed or delayed), we develop a clear dialogue with our giraffes. The dialogue entails the following communication system: A giraffe indicates when it's ready for cues; the trainer cues; the giraffe responds to cue, staying still until the duration and tool-pressure criterion is met; and, the trainer delivers the bridge (click) and back-up reinforcer (cracker). If the giraffe indicates hesitation or discomfort with a tell, the tools and mechanic's hands are immediately removed. If we see a tell, then we as trainers adjust our behavior and potentially the environment (e.g., move block in). We ask ourselves how we can modify what we are doing in order to get a "yes," or still behavior in these conditons, from the giraffe. What smaller approximation can we ask for? Shorter duration? Change the position of the block? Less pressure? Faster rate of reinforcement? Move to a different spot on the foot?

The nose-on-target with still behavior is a signal to move forward with our approximations. Any other behavior from the giraffe indicates to us that we should make a change in what we are doing or the setting. We're continuously looking for behaviors that tell us to proceed and behaviors that tell us to stop; or rather, behaviors that predict the animal will stay when we move forward with approximations or behaviors that predict the animal will leave. We try to get ahead of the latter by adjusting the environment to get them to stay.

Using Predictions During a Dynamic Training Session

During our training sessions, our goal is for the animal's behavior to continuously signal to us to proceed. While we want to succeed in completing husbandry and medical behaviors, the end goal of our sessions is to have an animal who is willing to participate as needed for the entirety of the procedure and who will come back enthusiastically for the next session. We are not necessarily looking for a perfectly trimmed hoof or a full vial of blood. We measure our success in training by the participation we see from the animal. A perfect looking foot is a secondary outcome of willing particpation.

We are, however, human. It is sometimes impossible to predict the unpredictable, but that is the goal. Each time we don't set the stage for the "proceed" signal and we receive information from the animal that they're uncomfortable, we log it into our ever-growing data set for that individual. Our history of past training sessions with each giraffe helps us to make better predictions during training sessions.

We, as trainers, are constantly taking in information from our environment and adjusting our training strategies accordingly. We might see rocks in the giraffe's feet, bacterial tracts we need to knife out, or bruising that would predict a giraffe offering a "tell" if we knife it out. We also track previous "tells" as information, so that we can adapt our ever-changing training strategies in the moment. Establishing a dialogue through proceed behaviors and tells is a critical prerequisite skillset for our trainers, as that's part of where our predictions come from that inform our dynamic reinforcement strategy. If we can predict a tell is likely, we can get ahead of it and aim for a "proceed".

<insert "Discrete ABC" video here. >

Dynamic Reinforcement Strategy

Once all the component tool-on-hoof behaviors are mastered to criteria (including location, pressure and duration), hoof work can be accomplished more efficiently by using a Dynamic Reinforcement Strategy (DRS) for the entire chain. During this phase in their training, we ask the giraffe to hold still with its foot on the block while we deliver reinforcers periodically throughout the procedure and adjust the rate of reinforcement based on what we predict the animal will do in response to a long or uncomfortable procedure. We adjust the rate based on influential variables, including things like the location of the tool on the foot (heels are sometimes more sensitive), the precense of bacterial tracts or rocks, the pressure of the tool, the time elapsed since the last reinforcer, and past history of places on the foot that predict "tells" for certain individuals. The schedule of reinforcement changes as dictated by the situation; it's what the animal does that influences rate, magnitude, and value of reinforcement, *not* rule governed adherence to ratio/interval or fixed/variable schedlues of reinforcement.

This training strategy doesn't neatly fit the within textbook definition of schedules of reinforcement. We suggest unhitching from the laboratory "schedule" wagon to instead watch the animal's behavior to determine our rate of reinforcement. The animal's meeting criteria determines the delivery of food, with feeding happening periodically at different intervals, based on observations and information from the two trainers and the giraffe.

The use-case for a DRS is when a procedure is likely to be inherently uncomfortable and we are asking the animal to remain in position while they're potentially experiencing some hesitance, pain or discomfort ***after*** all the prerequisite behaviors in the chain have been contingency shaped (cue-behavior-R+) with a continuous reinforcement schedule (CRF) and are then ready to be applied to an actual medical or husbandry session where we expect the unexpected. *DRS is used preemptively, just before discomfort or participation declining behaviors start.* The rate of reinforcement is increased when a procedure is likely to be

uncomfortable, based on information from the trainer best positioned to make that assessment based on history with the procedure, species and individual. It is meant to get ahead of the discomfort; we increase the rate right when the procedure begins. If discomfort behaviors (tells) occur, we stop the procedure and adjust the antecedent arrangement, reinforcement delivery, and/or behavior requested.

Medical and husbandry procedures are both inherently unpredictable and often painful or uncomfortable; we are usually doing things to the animals while they maintain a station/present behavior. If we could control all factors, as with an operant chamber, we could decide on and implement a schedule of reinforcement for these mastered behavior chains. Instead, we use the DRS with which we anticipate unpredictable forces (pain/discomfort from a sore spot on the hoof, additional duration, distractions from conspecifics, hardware malfunctions, etc.) and increase the rate of reinforcement before the tells are elicited/evoked. The animal becomes less likely to increase tells or decline engagement because the rate of reinforcement is increased before the animal indicates "no thanks". The DRS is based on the concept of dialogue training, i.e., the co-influence of two or more organisms as we work toward a targeted outcome. They work for positive reinforcement; we work for hoof care.

<insert "Dynamic Reinforcement Strategy" video here. Credit given to Daniel Ladner for video edits >

The Science Behind It

The Matching Law may account for the animal's choice to proceed with an uncomfortable procedure. By unhitching delivery of reinforcement from a predetermined schedule and instead following what the animal's behavior tells us, we can encourage them to hold their position during painful or uncomfortable parts of a procedure. We meet our medical and husbandry goals from which animals would otherwise retreat and increase the probability they will enthusiastically (short latency) return for the next medical or husbandry session.

With the matching law, there are at least two contingencies that the animal engages in, referred to as a "choice". When cued, the animal can choose to remain in station for positive reinforcement or leave the station for reinforcement elsewhere. By using the DRS, we increase the probability of the animal choosing to remain at station for the increased rate of highly valued reinforcement there.

Although this is a binary choice in this situation, it is appropriate to the task needing to be accomplished (medical or husbandry procedures). Across the lifestyle of the animal, we seek to provide more than two alternatives for which animals can use their behavior for a variety of reinforcers. Other options for completing medical and husbandry behavior might otherwise be more aversive, forceful, or coercive, or compromise welfare.

Aren't you just distracting them?

As with all constructed labels, we start the discussion by asking, how do we operationalize "distraction?" Oxford dictionary defines these three following terms:

> **Distraction**: A thing that prevents someone from giving full attention to something else.
> **Dynamic**: A force that stimulates change or progress within a system or process.
> **Encouragement**: Persuasion to do or to continue something.

How do we measure attention vs distraction? We are really describing a continuum of observable behaviors, for example, the orientation of an animal's gaze. Is it likely that an animal is ever completely unaware of what's happening to their bodies? Would the measure be a subsequent startle response, given a procedure? If we are concerned with a startle response, we consider what can be done reduce the startle effect. DRS is more akin to encouragement; we make it worth their while, given their behavior that indicates increased discomfort or effort during a session.

We can see, though, how this strategy might look unsystematic, indiscriminate, or similar to distraction. The reason it works so well is that it was built on top of established, component behaviors. We are not using this strategy to *teach* hoofcare, but rather implement it in real-time settings once the chain is mastered.

We have seen distraction training with individuals trying to *shape hoof care* with a constant feed strategy. We shape the links in the chain, then put them all together. An example of distraction training would be asking a giraffe to put its foot on the block, then offering a bucket for them to bury their head while you manipulate the feet. The DRS is a deliberate feeding strategy; it changes with the flow of the session. It should not be confused with indiscriminated free flow of reinforcers straight into their mouth.

Conclusion

The systematic effects of schedules of reinforcement (e.g., continuous and intermittent) were derived from the operant chambers in experimental settings with free operant behavior (i.e., the rats and pigeons were free to emit many responses, like lever pressing or key pecking, without deliberate cues from the experimenter). Although these known effects hold great interest for behavior professionals, most of our sessions are discrete trials, with obvious trainer-delivered cues and consequences. We train in highly variable settings and our outcomes are often demanding long and/or uncomfortable medical and husbandry behaviors. When we step away from force and coercion, what do we move toward? A dynamic reinforcement strategy produces a dialogue between trainer and learner characterized by careful observation and quick decision making within in the session, as we allow the animal's behavior to influence what we do. Behavior is never ignored in favor of following a predetermined or set schedule of reinforcement independent of the animal's experience.

In attempting to articulate the dynamic reinforcement strategy, it is clear to the authors that this is not a new or original strategy. Most of us have been training with this approach and we have heard many stories of trainers increasing rate of reinforcement dynamically based on the demands of the procedure, for example, a blood draw with a dolphin requiring a 30-minute station behavior. We hope that by describing the strategy, giving it a name – dynamic! – and a

rationale, we have distinguished the strategy from "just distraction" to "making it worth their while," and have further contributed to the discussion to make the approach more accessible. We look at it this way: If it were a kid needing to sit still for a nasal covid swab, we wouldn't stand by silently until the procedure was done. We would encourage them along the way. By analogy then, to encourage a giraffe, think cracker, browse, or lettuce!

In this paper, we focused mainly on one component of effective, humane training – a dynamic reinforcement strategy. But all aspects of training should dynamically follow the needs of each animal, consistent with our study-of-one philosophy. We should avoid dogmatism in our selection of training procedures, and instead watch the data – the animal's behavior – to custom fit what we do in light of what they need us to do, to increase their success living among humans. Here are some of the in-process, evaluative data we collect, and respond to dynamically:

1. Are the approximations moving toward the desired outcome or are we stalled at one approximation?
2. Is the target behavior increasing or decreasing?
3. Are "tells" of discomfort increasing or decreasing?
4. Is participation increasing or decreasing, including latency and duration criteria?

As the cool horse trainer Alexandra Kurland is known for saying, "Go to people for opinions and animals for answers! Then let their answers influence what we do."

Acknowledgements:
Special thanks go out to the rest of the Behavior Works Consulting Team (Megan Sanders and Bianca Papadopoulos), who offered multiple feedback sessions, insight, and guidance to the application and discussion around what it was that we were doing during training. Thank you, also, to Jason Bredahl, Diana Miller, Kayla Ringuette, Rachael Hahn, and Kelsey Newman, members of the Cheyenne Mountain Zoo giraffe team. They continue to innovative, adapt, and track reinforcers with a keen eye. We'd also like to acknowledge Dr. Liza Dadone, Brenda Cordova, and DeeAnn Wilfong from CMZ's veterinary team, who help guide the behaviors we train for. A final thank you goes out to Dr. Carl Cheney – without his Sunday dinner insights, we wouldn't have this paper.