

## Definition drift or Are we monopolizing marine mammal training by changing the definitions?

by Jaap van der Toorn

In the latest issue of Soundings, there was a paper in the Trainer's Forum section about training terminology (Beeler, 1996). In this paper, the drift from standard operant conditioning terminology towards a marine mammal specific jargon becomes apparent. This redefinition of terms is undesirable, since that will make communication with people from other disciplines (behaviorists, ethologists) more difficult. We should try to stick to the standard definitions as much as possible. Where the operant conditioning theory is not detailed enough for our purposes, we can of course add more specific terms, provided we also define them well, so that we can easily explain them to outsiders.

With this in mind let's have a closer look at some of the relevant terms and their definitions. I have taken information from several references, including Holland and Skinner, 1961, Pepper and Defran, 1975 and Barrows, 1995 and have also drawn on personal experience as outlined in Van der Toorn, 1987.

### Reinforcer vs. reinforcement

What is a reinforcer? A **reinforcer** can be defined as any stimulus that can be used to influence (operant) behavior. For instance, a piece of food can be a reinforcer. This is a **positive reinforcer**: the application of this reinforcer will make it more likely that the preceding behavior will be repeated. The application of a reinforcer (the presentation of that piece of food after a desired response) is called **reinforcement**, in this case more specifically, **positive reinforcement**. You often see the terms used interchangeably, but in fact they are two different things. The reinforcer is the (potential) reward, while the reinforcement is the presentation of that reward at the proper moment.

An aversive stimulus (for instance a disturbing light or noise) can also be used as a reinforcer. There are two different strategies for using aversive stimuli. Although marine mammal training is rooted in the use of positive reinforcement, I will mention these strategies for the sake of completeness, so that inadvertent use of these strategies can be recognized. **Negative reinforcement** is used when an aversive stimulus is removed after a desired response. The subject is rewarded by taking away something bad. It is also possible to apply an aversive stimulus after an undesired response. This is called **punishment**. In the following, I will assume that we are using a positive reinforcement strategy only.

### Primary reinforcers

A **primary (or unconditioned) reinforcer** is a stimulus, which has inherent rewarding properties. As such it can be used directly in reinforcement. Often food is used as a primary reinforcer in training situations. However, there can be other primary reinforcers. Also food does not always have a reinforcing value. So a statement like "Food is a primary reinforcer" are not always correct. Food can be used as a primary reinforcer only when an animal is interested in food.

Let's take a look at some cases. Assume you are training a male bottlenose dolphin named Delphi. Delphi has been in training for quite a while and he is comfortable with you. He is usually responsive in informal play sessions and likes to be rubbed and scratched. He is what we would call a very tactile animal. Now consider the following situation, which most trainers will recognize:

You have just started a training session and after a correctly performed behavior you give Delphi a fish. However, this morning Delphi does not feel like eating. He may have some discomfort in his stomach. What is the fish in this context?

- a. a primary (positive) reinforcer
- b. not a reinforcer at all
- c. an aversive stimulus

It should be obvious that, since Delphi does not want the fish, it is not a primary reinforcer, at least not a positive one. Both options b and c are possible at this point. Without more information we cannot say for sure which one is correct. If Delphi just acts disinterested in the fish and also drops the next fishes offered, the fish most likely has no reinforcing value whatsoever. If you give him more fish and he becomes agitated, the fish could even be an aversive stimulus. You should be aware of that situation, because applying an aversive stimulus is the definition of punishment: Delphi feels you are punishing him, while you are actually trying to do the opposite.

And now a variation on the previous example. You are training Delphi again and he drops the fish you offer him. At the same time he turns over on his back and obviously solicits a belly rub. You comply and rub him. Delphi then gets ready for the next behavior, performs, comes back and immediately turns over again and you comply again by rubbing him. Which of the following statements are true in this context?

- a. food is a primary reinforcer

- b. rubbing is a primary reinforcer
- c. there is no primary reinforcer involved

As we saw in the previous example, if Delphi is not interested in food, food is definitely not a primary reinforcer. In this situation, Delphi likes to be rubbed, so the tactile stimulation has a reinforcing value. Since it apparently satisfies a need in Delphi, the rubbing is a primary reinforcer. And because we now have a primary reinforcer, obviously statement c is false as well.

### Secondary reinforcers

A **secondary** or **conditioned reinforcer** is a stimulus that has acquired reinforcing properties through association with a primary reinforcer. Usually, a stimulus is used that was originally meaningless to the subject. We are all familiar with the **bridge** signal, often the blow of a whistle or the click of a clicker. These are examples of secondary reinforcers. (A bridge is a special case of a secondary reinforcer, chosen so that it can be applied at the exact moment the animal is doing what is desired). Since a secondary reinforcer is conditioned, it can be extinguished as well. If the stimulus is never again followed by the application of a primary reinforcer, it will lose its reinforcing value and will become meaningless. With this in mind let's have a closer look at another training situation with Delphi.

You ask Delphi to present his flukes. He performs the behavior correctly and you blow a whistle to signal that he has performed correctly. Delphi comes to you and you rub his back. Next you ask him to blow (for the collection of a sputum sample for instance) and he complies again. You blow the whistle and play his favorite game (tug-of-war) with him. No food is involved. In fact, you may not even have brought a bucket of fish with you. Now let's take a more analytical look at this session in the context of what we have discussed so far:

- a. When Delphi performed correctly, you blew a whistle. Delphi has been conditioned to the whistle and for him it is a signal that he has performed correctly. You have just bridged him, or in other words, you have used secondary reinforcement (the application of a secondary reinforcer, i.e. the whistle).
- b. When he popped to the surface, you rubbed his back, something you know from experience he likes. So you have now also applied a primary reinforcer, the rub, as well. By doing that, you have also strengthened the reinforcing value of the secondary reinforcer.
- c. When Delphi performed his blow behavior correctly, you again blew the whistle: you again used secondary reinforcement.
- d. After that you played his favorite game with him: you applied a primary reinforcer: the game.

### Is tactile stimulation a primary or secondary reinforcer?

Often, tactile stimulation and other non-food stimuli that are used as reinforcers are referred to as secondary reinforcers. It will be clear by now, that this classification is incorrect. Once established, these stimuli have an intrinsic reinforcing value, which cannot be extinguished by no longer following it with food reinforcement. At that point the stimulus has become a potential primary reinforcer. In fact, it has never been a secondary reinforcer in most cases. Let's look at the development of a backrub in a naïve dolphin (like Delphi, when he has just been introduced to a training program). At first, Delphi is wary of the trainer (and of people in general). He tries to keep a distance and does not allow any direct contact. Through targeted training, Delphi learns to tolerate a slight touch on his back. When he allows it, he is bridged and he is given some fish as a primary reinforcer. Over time, Delphi learns to tolerate a complete backrub and is bridged every time he stays in place for a backrub. At this point, the backrub is just another trained behavior for Delphi, just as a fluke presentation or a bow. He is reinforced for performing the backrub behavior on cue. Once Delphi has become more comfortable with his trainer he begins to like the physical contact and actively solicits backrubs during play sessions. This active soliciting of backrubs indicates that they have become rewarding for him: in other words, the backrub has acquired an intrinsic reinforcing value and can now be used as a primary reinforcer. Notice it was never a secondary reinforcer: it developed from a trained behavior into a (potential) primary reinforcer.

From discussions in Soundings and on conferences, it has become obvious that there is a need for a distinction between food as a reinforcer, which makes use of a very basic drive in every animal (hunger) and other types of reinforcers, like tactile stimulation, which can be used with animals which have been in training for a while and which are usually specific for a given individual animal (some like backrubs, others don't). I propose the following definitions:

- a. **basic (primary) reinforcer:** a stimulus that satisfies a basic need in an animal. Food and water fall into this category.

- b. **acquired (primary) reinforcer:** a stimulus that has acquired a reinforcing value through extensive contact with the trainers. Tactile stimulation and play, among others, fall into this category.

### Generalized reinforcers

Often, a secondary reinforcer, like a bridge, is conditioned using food as a primary reinforcer. This will establish the bridge as a signal that food will be given. This one-on-one link between the bridge and food is undesirable, because that means that the bridge will become ineffective, when the animal is not hungry. As we have seen before, food is not the only primary reinforcer available. When we let the bridge be followed by a variety of primary reinforcers, we broaden the possibilities of using the same bridge: it now indicates that a primary reinforcer is coming, but it is no longer dependent on an animal being hungry. As long as we have some form of primary reinforcement available, the bridge can be used effectively. In operant conditioning terms, the bridge has become a **generalized reinforcer**, which is a very powerful training tool.

### Reinforcement schedules

Basically, there are two ways of applying reinforcers:

1. After each correct response. This is called a **continuous reinforcement schedule**.
2. After a certain number of correct responses or after a certain time has passed. These are called **intermittent reinforcement schedules**. Four "flavors" of intermittent reinforcement schedules are defined:
  - a. **Fixed ratio schedule:** every Nth correct response is reinforced, where N is a fixed number for the session. For example, you ask Delphi to touch a target a number of times. Every 3rd time he touches the target, you use the bridge (apply a secondary reinforcer) and also supply some primary reinforcement. Delphi is reinforced on a 1:3 fixed ratio schedule.
  - b. **Fixed interval schedule:** a correct response is reinforced every Nth minute or second, where N is a fixed number for the session. Again, Delphi is asked to touch a target a number of times. You don't count the number of times he touches the target, but instead you reinforce him every 15 seconds, immediately after he touches the target. When he touches the target somewhere in the 15 second interval, he is not reinforced. He is now reinforced on a 15 second fixed interval schedule. Note that in this schedule, Delphi is not asked to hold the target for 15 seconds; the duration of the response is irrelevant in this context.
  - c. **Variable ratio schedule:** every Nth correct response is reinforced, where N varies randomly around a predetermined average number for the session. If we reinforce Delphi on a variable ratio schedule with an average of 1:3, he gets reinforced after 3, 5, 4, 2, 3, 5, 1, 4, 2 etc. correct responses. This yields an average of 1:3, but the actual ratio varies.
  - d. **Variable interval schedule:** a correct response is reinforced after a certain amount of time has elapsed. The interval between the availability of reinforcements varies randomly within the session around a predetermined average. If you reinforce Delphi on a variable reinforcement schedule with an average of 15 seconds, you may for instance reinforce a correct response after 1, 8, 21, 13, 11, 25, 5, 26, 16 etc. seconds.

From the point of view of the animal, there is no obvious difference between variable interval and variable ratio schedules. Outside a more or less controlled laboratory-type environment, interval-based schedules are often hard to use. Also, a real random distribution of reinforcement as is needed for a real variable ratio schedule needs preparation and a constant monitoring. Often, trainers vary the ratio of reinforcement by "gut-feeling", which yields at best a pseudo-random distribution of reinforcement, which can be predictable.

Within the marine mammal community, there also is a special flavor of a variable ratio reinforcement schedule: **V.R.R.V.** or **Variable Ratio with Reinforcement Variety**. V.R.R.V. is a variable ratio schedule, as described above. It also needs the use of a variety of primary reinforcers, so it relies on the use of a generalized secondary reinforcer (or bridge).

A reinforcement strategy, that is often seen, but that has no official name (yet) is the following: each correct response is bridged (secondary reinforcement is applied) and this is followed by primary reinforcement on a variable ratio schedule. A variety of primary reinforcers may be involved as well. So not every bridge is followed by primary reinforcement. Since every correct response is reinforced, this is strictly speaking a continuous reinforcement schedule. You might call this **Continuous Reinforcement with Variable Ratio Primary Reinforcement Follow-up** or **C.R.V.R.P.R.F.** (yes, an even longer abbreviation than the V.R.R.V. You can forget this name if you want and propose a better one, as long as you remember the concept). This approach is very useful when you are still in the process of

developing a behavior. In such a situation, you want good control over the behavior and want to reinforce every response that meets or exceeds your criterion for the session. Once the behavior is performed reliably, you can switch to an intermittent reinforcement schedule, like V.R.R.V.

## Conclusion

In this paper I have presented the original definitions of operant condition terms. Since what we as marine mammal trainers are doing still fits within the description of operant conditioning, I think we should stick to these definitions, so that we can communicate better with people from other fields. Marine mammal training is a dynamic field and we see new developments emerging regularly. To describe the new developments, the existing operant conditioning terminology may be too restrictive. When that is the case, we should introduce new terms and clearly define them, as it has been done with the introduction of V.R.R.V. This will make it instantly obvious to outsiders that this is something new. We should not change the definitions of existing terms, because then we run the risk of isolating ourselves from experts in related fields, like ethologists and behaviorists. I hope this paper has contributed to putting things back in their proper place.

## References

- Barrows, E.M.** (1995)  
Animal behavior desk reference. CRC Press, Inc., Boca Raton
- Beeler, M.C.** (1996)  
Basic marine mammal training terminology and techniques defined and explained. *Soundings* 21(1): 22-25
- Holland, J.G. and B.F. Skinner** (1961)  
The analysis of behavior. McGraw Hill Book Company, Inc., New York
- Pepper, R.L. and R.H. Defran** (1975)  
Dolphin Trainer's Handbook. Part 1. Basic Training. *Naval Undersea Center, Kailua, Hawaii* 96734: NUC TP 432, April 1975
- van der Toorn, J.D.** (1987)  
The importance of variable reinforcement training for dolphin husbandry. Summary of oral presentation in **E. F. Krajniak and M. S. Bryant** (Eds.), *Proceedings of the Annual Conference of the International Marine Animal Trainers Association* (1987) pp. 133-134. International Marine Animal Trainers Association.

## Paper history

This paper was submitted for publication in *Soundings* as a response to Beeler (1996) but it was not published in its current form. Part of this paper has been included in:

- Ramirez, K.** (1999)  
Animal training: Successful animal management through positive reinforcement  
John G. Shedd Aquarium  
ISBN 0-9611074-9-9