

Why Reinforcement is Illusory

Duane M. Rumbaugh, Great Ape Trust of Iowa; also, the Language Research Center,

Georgia State University

David A. Washburn, Language Research Center and Psychology, Georgia State

University

James E. King, Department of Psychology, University of Arizona; also, the

Language Research Center

Michael J. Beran, Language Research Center, Georgia State University

Now the tradition in psychology has long been a search for the property of the stimulus which by itself determines the ensuing response, at any given stage in learning. This approach...is no longer satisfactory as theory (Hebb, 1949, p. 4).

Although we fully agree that the immediate correlates and consequences of behavior are fundamental to the learning process and acknowledge the fact that the concept of reinforcement has played a very significant role in the history of learning and behavior, it is time that we jettison the term. The reasons for so doing are as follows: 1. the term carries excess meaning in that it encourages the belief that reinforcers actually strengthen associations between stimuli and specific behaviors and that all behaviors have reinforcement histories. 2. Its definition has been inherently circular. 3. It emphasizes behavior inordinately in its relation to specific stimuli and does not encourage consideration of whatever the subject might bring as a sentient and knowledgeable agent-of-action to the determination of its behavior. 4. It cultivates a

sense of inappropriate controlling power in those who apply the term to their studies of animal and human behaviors. 5. It inordinately emphasizes fixedness in behaviors and detracts from our likelihood of observing emergent behaviors – those that are creative, new, efficient, and insightful. 6. By jettisoning the term and using reward in its stead, we still acknowledge the importance of consequences upon behavior and recognize that what we have called reinforcement is really a resource of value to our subjects and frequently is equivalent to “pay for work done.” These so-called reinforcers may be critical to understanding an animal’s motivation and performance, but are only indirectly and not causally related to the processes of animal’s learning.

In its most basic traditional definition, reinforcement is posited as a theoretical process that strengthens an association a stimulus and a response. In its most basic definition as an operation, a reinforcer is something that happens or that is given or vended to subjects contingent upon their behavior. If that procedure increases the likelihood of the behavior, then the event following the behavior is said to be a reinforcer.

In psychology’s historic ambition to become as hard a science as possible, reinforcement has had its attractions because it posits nothing by way of sentience, awareness, understanding, or cognition of/by the subject. It is to be understood in the history of psychology as useful in the formulation of an empty-organism, even empty-headed psychology of learning and behavior.

But now that we have solid evidence of animals’ (including of course, humans’) complex cognitive skills and potentials, the concept of reinforcement is no longer very appropriate. There is far more evidence that animals’ learning either is or becomes more cognitive as the animals interact with a complex worlds and are no longer driven only by

simple stimulus-response associations, bonds, or habits. One may *reward* cognition; one may *reward* responses. *But cognitive operations are not at all subject to reinforcement as though they were muscles being strengthened by exercise.*

Yet, it would be folly to assert that the contingencies – or aftermaths – of behavior have no effect. We are not saying that. What we are arguing for is that the concept of reinforcement should be supplanted with one or more other terms -- reward, pay for work done, resource garnered by foraging or taking, or benefit. We prefer to keep it simple and to use the single term, *reward*.

Rewards can be positive; rewards can be negative. To the degree the terms reinforcement and reinforcer are carried forward in our language, we believe that they should be used strictly in reference to an operation established by a researcher that avails generally a beneficial contingency to a subject if it responds in some specific way.

The manifestation of learning by the subject is expected to be the specific behavior that is reinforced. But we know, for instance, that a rhesus monkey trained to use its foot exclusively to control a cursor via manipulation of a joystick learned something far more comprehensive. After the task had been mastered and the monkey was given the first opportunity ever to use either its hand or foot, it used its hand – and it was better with its hand than it had ever been with its foot. Though reinforced in training to use its foot, it learned comprehensively about the task, namely the relationship between movement of the joystick and the movement of the cursor on the video screen!

Although reinforcement sounds like a really solid scientific term, its definition has been more circular and problematic than not. And to say that a reinforcer is anything that increases the probability of response or behavior is a clear loop. To hold that it is the

more probable of two behaviors requires *a priori* measurement of individual response probabilities.

At this point, the listener might retort, “But conditioning procedures, both in classical and operant conditioning, are intended to result in responses that are predictable, relatively stereotyped responses! They are conditional responses!” True, but that is not where the learning ends. We have strong evidence that within the constraints of time, neural system, and the developmental stage of the learner, relational processes come into play. These relational processes are the natural operations of the brain, constructively biased by the species to which it belongs. Relational processes, natural to species’ various brains, process experiences of life, including conditioning and learning trials, into what we will call a knowledge base. The subject can be expected to learn far more than a specific response or response pattern, but rather that it about the relationships among events and how resources/rewards can be acquired either through specific work or through creative solutions so as to enhance the efficiency of effort and to obtain a comfortable life.

As noted earlier these relational processes are not in any sense subject to reinforcement as though they are muscle responses – which, indeed, they are not. Neither are they even accessible for being reinforced. Rather, it’s the brain’s business per species to make sense out of the experiences of passing time. The brain is a very expensive organ to maintain, and it cannot justify its costs by being passive. Indeed, it not passive but rather a marvelous system whereby knowledge is generated from predictable events and eventualities of perceiving and behaving. We posit that the design of brains serves to bias the selective perception of events in accord with their salience – natural or acquired

– and, also, to organize or inter-relate them in accord with the ecological resources and needs of the subject so as to survive.

What is salience?

There are several *natural sources* of salience – natural sign-stimuli, species-defined; strong stimuli (e.g., energies with high db and/or lumen and/or pressure, etc.) that are within the sensory boundaries of a given species; and biologically predicated need states, as for moisture, nutrients, and an ambient temperature range that varies widely across life forms. These sources of salience are not dependent upon experience. They are givens.

Other sources of salience are accrued, not through learning but rather by reason of what we view as a natural principle – Things that occur reliably in about the same time and space reliably tend to couple, to mix. This is true of most liquids, fumes, and even metals. It clearly is the case in the production of colors and odors. The celebrated neuroscientist, Professor Edelman, observes that neurons and neural circuits that fire together get wired together. It's a natural law or principle, echoing the work of Hebb and others, though we suspect, not without its exceptions.

Our framework holds that *the basic principle of learning is based on the reliable temporal or spatial contiguity of events*. Stimuli, events, and/or behaviors that co-occur reliably become at least metaphorically coupled. In their coupling, stimuli, events, and/or behaviors acquired a unique dimension of salience. Some events are coupled naturally, such as lightening and thunder; whereas others are produced by action on the environment, such as lactation being induced by suckling. Generally, the co-occurrence

itself is the coupling, and its recognition of this co-occurrence that is the learned percept. (Our position is no exception to historical recognition of the fact that temporal contiguity is fundamental to learning.)

A second principle of our framework is *that in the coupling of stimulus events there will be a sharing of their response-eliciting properties as some positive function of the vigor of the responses produced by each stimulus and the relative strengths of their responses when they co-occur in time and/or space*. Thus, in classical conditioning, both the conditional and unconditional stimuli bidirectionally share their response-eliciting properties. It is only because the unconditional stimuli are the stronger of the two that their response-eliciting properties are more strongly manifested with the presentation of the conditional stimulus rather than vice versa. In other words, the high strength unconditional stimulus will cause a large change in the response to the conditional stimulus whereas the low strength conditional stimulus will have little or no effect on the response to the unconditional stimulus. Notwithstanding, we believe that the unconditional stimulus accrues an approximation of the relatively minimal response-eliciting property that inheres in the conditional stimulus. (See Note No. 3 below.)

A third principle of our framework is *that species' brains are uniquely designed to process coupled stimulus events, to somehow file and process them to the end that emergent behaviors and emergent capacities might be formed to service adaptation in both familiar and novel challenges*. Organisms detect coupled events; that is, animals recognize reliable and predictable patterns in the stimuli they experience. The couplings are “out there” to be perceived or they are reliable sequelae of behaviors. They are the regularities and invariants of experience across time. The more complex the pattern, the

more complex the cognitive system (and the brain) must be to recognize it; notwithstanding, all animals learn by detecting predictable temporal or spatial relations.

Among the several sterling contributions made by primatologists is the uncontested principle that rearing and environmental factors, variables, and experiences can have long lasting sculpting effects upon organisms' intelligence, emotions, interests, personalities, and morphology.

In the area of primate behavioral development, Mason (2002) sees *emergence* as a concept that is fundamental to the understanding of behavioral development and that requires new descriptive categories and measurement. The comprehension of human speech and the meanings of various word-lexigrams by apes without formal training are prime examples. Just how emergents are generated by the normal operations of the brain from the experiences that all of life offers is at present imperfectly understood, to say the least. When asked how they are formed, one eminent neuroscientist replied, "God only knows." To come to understand the parameters of emergents at any level will take decades of research at all levels, but to understand them better along the way will be reward sufficient to the task. In the meanwhile, primatologists can make invaluable contributions in defining the antecedent and subject parameters of emergents and how emergents impact upon subsequent behavior, including learning and all other basic processes.

Summary

1. Choices must be made by organisms as to what they will attend to in their worlds. The role of salience in determining attention and behavior is very significant. Genetics determine whether some stimuli are or are not salient, whereas experience declares that for others.

2. We hold that the basic process of sensory preconditioning likely prevails in both respondent and operant conditioning situations. Sensory preconditioning enhances the salience of basically neutral stimuli simply by pairing them together temporally with more salient stimuli. The previously neutral stimuli gain in salience and become functionally equivalent to the more salient stimuli in their roles. If one member of a pair becomes a conditional stimulus, then quite likely the other one will prove effective in that role despite the fact that it served as a conditioned stimulus in a conditioning procedure.

3. The near-contiguous pairing of the CS with the UCS in respondent (Pavlovian) conditioning allows for the CS (initially an innocuous stimulus) to acquire, in part, the role of the UCS. Whereas the UCS likely accrues the role function of the CS, the stand-alone role of the CS is so subtle that it likely is not detectable.

4. The phenomena of autoshaping and the topography of the conditioned response of pigeons acquired therein provide strong support for the frame of reference here advanced – that the functional role of the reinforcer is shared (e.g., become elicited by) with a visual target and/or a discriminative stimulus (S^D) temporally associated with the reinforcer. If the reinforcer is grain, the bird pecks at the target as though it were food; if it is water, the bird “pecks” as though it is drinking water.

5. From our frame of reference, a conditional response essentially is a manifestation of the partial functional equivalence of the CS for the UCS. (The response, once conditioned, perhaps never duplicates the response elicited by the UCS.)

6. Reinforcement works, according to our frame of reference, because it shares its salience with other current stimuli and behaviors.

7. In our view, reinforcement does not serve to strengthen response probabilities directly to a stimulus or to a context. Rather, it defines a resource of value to the organism because of that organism's biological and acquired needs. In conditioning contexts, the organism learns about resources that can be obtained and, also, about how to obtain them. That said, the organism is likely to learn primarily about the task and how to get the valued resource in a number of ways.

8. Thus, we recommend that the term "reward" be used instead of reinforcer, due to the history that declares that the reinforcer directly strengthens a specific response or behavior. Rewards can be used to bring focus to fixed behaviors because rewards basically are "pay for work done." In other words, it gives the organism a reason to care about the predictable patterning of stimuli that is available to be learned.

9. Animals and we acquire information about our environments, with their resources and risks, not by S-R Reinforcement of specific behaviors but through interactions with the environment and thereby discern reliable patterns that become information when processed by the brain and, hence, can then service behavioral efficiency and creativity.

Notes

1. The framework here advanced is that of rational behaviorism advanced by Rumbaugh and Washburn in their book, The Intelligence of Apes and Other Rational Beings (Yale University Press, 2003). Preparation of this paper and much of the research referenced in it were supported by the National Institute of Child Health and Human Development grants HD-06016 and HD-38051. Additional support was extended to the first author by the Great Ape Trust of Iowa (www.GreatApeTrust.org). Email – drumbaugh@aol.com drumbaugh@greatapetrust.com

2. Conditioning procedures, both Pavlovian and Skinnerian, alter the response-eliciting properties of stimuli – termed conditional or discriminative stimuli. Furthermore, previously neutral stimuli associated with stimuli referenced above acquire reward properties! To a degree, they even acquire the response-eliciting properties of those stimuli. Thus, in autoshaping, for example, the pigeon is not required to peck at a light that comes on when food or water are presented, with the pigeon being hungry or thirsty at the time. But they gradually begin to peck at the light -- as though it were food and, on the other hand, peck as though drinking. Clearly, the response-eliciting properties of food and water become shared with the light panel – that fundamentally was irrelevant to the situation. This is but one of several examples to be drawn from the literature.

3. Regarding future research that bears upon our view, we anticipate that renewed research on the parameters of sensory preconditioning should be highly informative regarding the consequences of selecting stimuli, for pairing, from the same natural class (e.g., photos of two fruits, vegetables, beverages, persons, conspecifics, locations, etc.)

and from different natural classes (e.g., one photo of a fruit paired with one of a location, etc.). Variations in relative size and brightness of the stimuli to be paired might reveal that the impact of pairing is some positive function of the strength of the individual stimulus members (e.g., stronger stimuli, that is the larger and brighter ones, should more readily impart their subsequent roles to the smaller, dimmer ones, etc.). Relationships between the stimulus members paired and the reward to be subsequently given in conditioning paradigms, where one member is used as a CS, might tell us of the role of relatedness between stimuli and their role-sharing potential in relation to or synergized by the reward to be obtained in learning situations. Will the sensory preconditioning of stimuli result in quite altered findings in studies of blocking, overshadowing, disinhibition, and so on, contexts in which a stimulus is added after a procedure has been implemented? Will a stimulus that has been sensory preconditioned with another disinhibit extinction established where the other stimulus having served as the CS in learning?

The frame of reference here offered and in our book serves to unify behaviors at several levels and suggests many lines of empirical research that will help to assess the validity of the ideas and hypotheses here advanced. Data are always the best way to assess ideas and theories. With that we are in full accord. That said, we know of no area that cannot be embraced by the one advanced here. We are confident that studies of the temporal associations of stimuli in accord with the nature and levels of salience will in the future tell us a great deal about basic learning processes from sign stimuli through mediational learning and creativity.