

[manuscript of commentary which appeared in , *Behavioural and Brain Sciences*, 17, 154-155.]

How general is a general theory of reinforcement?

Stephen F. Walker

Department of Psychology, Birkbeck College, Malet Street, London WC1E 7HX, England
Electronic Mail: UBJTA25@cu.bbk.ac.uk

Both the title of the target article and its contents invite comparison with the mathematically formulated principles put forward by Hull (1943, 1952) which were intended to apply to all behaviour but which were stated precisely enough for their lack of generality to be eventually demonstrated. The first paragraph refers to the “wide-ranging implications” of the principles presented, and section 5 is headed “A General Theory of Reinforcement.” There is a proviso in section 9.3.1 that the present paper “addresses only distance along the dimension of homogeneous operant responses”, but it is implied that this is an example which will be capable of extension. One of the strengths of the theory presented is that it contains parameters which are good candidates for explanations of differences between responses categories and between species. I shall therefore comment first on questions of the relevance of the theory to data outside its base of homogeneous operant responding, and second on whether the theory is sufficiently powerful even within this base.

The most immediate difficulty for Hullian response- reinforcement theory was dealing with spatially directed behaviour. A very simple example described by Hull (1934) was that of a rat trained to run a fixed distance past a closed door, and then to come back to it, the door then being open to allow the animal to proceed to a food reinforcement. Hull noted that under these conditions highly trained rats would waste time trying to scratch through the closed door as they passed it, and some would go through the door on the first pass if it was left open. This, and many other examples provided by Tolman (1932, 1948) could not be directly predicted from Hull’s first principles which were entirely frequency sensitive and response based, and he developed elaborate ramifications of his theory (“habit-family hierarchies”) to account for them. A much more direct option appears to be open to Killeen, since he suggests that processes of incentive coupling may occur on any dimension of “the organism’s psychological space” (9.3.1). If representations of geographical space are major component’s of most species’ psychological space (Gallistel, 1990), then the coupling of incentives to locations, rather than only to the responses currently required to reach them, would seem to be an important area of a general theory of animal learning. An enormous amount of experimental evidence, for instance from the radial maze (Olton, 1979) and memory for hoarded food (Shettleworth & Krebs, 1982) is available for the testing of principles in this area, and yet Killeen’s discussion of foraging (9.3.4) refers only to memory indexed by responses. Good evidence is presented that response indexing occurs under operant schedules of reinforcement, whose contingencies require it, but other circumstances (including many examples of Pavlovian conditioning and operant discrimination learning) may induce coupling of incentives to “the stimulus as coded” or “the location as coded”.

A second area of difficulty for Hullian theory was the rapidity of behaviour change after reinforcement manipulations (latent learning and reward-devaluation), and this resulted in the separation of habit formation from incentive learning. Killeen appears to have brought them back together again, which may be constructive, but how in this case can one account for evidence which distinguishes between stimulus-response associations, and response-reinforcer associations? Can the treatment of motivation by means of the activation function (which bears an informal resemblance to “Incentive Motivation”: Hull, 1952) account for recent evidence on reward devaluation effects? (E.g. Dickinson, 1985; Rescorla, 1990).

The focus of the target article is on schedule effects in conventional operant conditioning, and it particularly provides an alternative to theories which propose direct effects of reinforcement on IRT's. An acknowledged area of vagueness is in the structuring of “response-units” (section 5.4; last para) which have to be inferred in some cases, but are especially obvious where there is sensitivity to response number, either in FR schedules, or in explicit “counting” schedules (Mechner, 1958; Davis & Pérusse, 1988). In these cases it is arguable that “number” or “run-length” is directly reinforced. Limitations on generality within the domain of repeated operant responses occur in so far as there is no treatment of negative reinforcement, but one would expect that straightforward modifications to the mathematics (or even redefining what constitutes a positive incentive) would allow an interesting extension of the present principles. The issue of “IRT reinforcement” in free-operant avoidance learning was raised by Sidman (1954), but later studies (Sidman, 1962; Herrnstein, 1970) suggest that explanations which do not require reinforcement of specific IRT's should be preferred. However, there are likely to be difficulties in identifying the terminal response and the point of reinforcement: “contiguity” in free operant avoidance, especially in probabilistic versions (Herrnstein & Hineline, 1966) is by definition less visible than in schedules of positive reinforcement.

An apparent gap in the theory, which applies to punishment of operant responses (Boe & Church, 1967) but also to aspects of extinction, is that there is no separate mechanism of response inhibition. In some cases the weighted moving average might be expected to suffice, but there is compelling evidence from many areas of animal learning, including behavioural contrast effects in multiple schedules, that a theory which includes only excitation and its absence is only half a theory.

Clearly the target article is intended to be narrow but thorough, rather than all-embracing. But part of its appeal is its potential generality, and this can only be realized by testing the theory against a wider range of phenomena than has so far been attempted.

References

- Boe, E.E. & Church, R.M. (1967) Permanent effects of punishment during extinction. *Journal of Comparative and Physiological Psychology* 63: 486-92
- Davis, H. & Pérusse, R. (1988) Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioural and Brain Sciences* 11: 561-79.
- Dickinson, A. (1985) Actions and habits: the development of behavioural autonomy. *Proceedings of the Royal Society, B* 308: 67-78.

- Gallistel, C.R. (1990) *The Organization of Learning*. Cambridge, Mass.: MIT Press
- Herrnstein, R.J. (1969) Method and theory in the study of avoidance. *Psychological Review* 76: 46-69.
- Herrnstein, R.J. & Hineline, P.N. (1966) Negative reinforcement as shock frequency reduction. *Journal of the Experimental Analysis of Behaviour* 9: 421-30.
- Hull, C.L. (1934) The concept of habit-family hierarchy and maze learning. *Psychological Review* 41: Part I 33-54, Part II, 134-152.
- Hull, C.L. (1943) *Principles of Behaviour*. Appleton- Century-Crofts: New York.
- Mechner, F. (1958) Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behaviour* 1: 109-121.
- Olton, D.S. (1979) Mazes, maps, and memory. *American Psychologist* 34: 583-96.
- Olton, D.S. & Samuelson, R.J. (1976) Remembrance of places passed: spatial memory in rats. *Journal of Experimental Psychology: Animal Behaviour Processes* 2: 97-116.
- Rescorla, R.A. (1990) The role of information about the response-outcome relation in instrumental discrimination learning. *Journal of Experimental Psychology: Animal Behaviour Processes* 16: 262-270.
- Shettleworth, S.J. & Krebs, J.R. (1982) How marsh tits find their hoards: the roles of site preference and spatial memory. *Journal of Experimental Psychology: Animal Behaviour Processes* 8: 354-75.
- Sidman, M. (1954) The temporal distribution of avoidance responses. *Journal of Comparative and Physiological Psychology* 47: 399-402.
- Sidman, M. (1962) An adjusting avoidance schedule. *Journal of the Experimental Analysis of Behaviour* 5: 271- 277.
- Tolman, E.C. (1932) *Purposive Behaviour in Animals and Men*. Century: New York.
- Tolman, E.C. (1948) Cognitive maps in rats and men. *Psychological Review* 55: 189-208.