PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Actions and Habits: The Development of Behavioural Autonomy

A. Dickinson

Phil. Trans. R. Soc. Lond. B 1985 **308**, 67-78
doi: 10.1098/rstb.1985.0010

Phil. Trans. R. Soc. Lond. B **308**, 67–78 (1985) [67]
Printed in Great Britain

Actions and habits: the development of behavioural autonomy

By A. Dickinson

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K.

The study of animal behaviour has been dominated by two general models. According to the mechanistic stimulus–response model, a particular behaviour is either an innate or an acquired habit which is simply triggered by the appropriate stimulus. By contrast, the teleological model argues that, at least, some activities are purposive actions controlled by the current value of their goals through knowledge about the instrumental relations between the actions and their consequences. The type of control over any particular behaviour can be determined by a goal revaluation procedure. If the animal's performance changes appropriately following an alteration in the value of the goal or reward without further experience of the instrumental relationship, the behaviour should be regarded as a purposive action. On the other hand, the stimulus–response model is more appropriate for an activity whose performance is autonomous of the current value of the goal.

By using this assay, we have found that a simple food-rewarded activity is sensitive to reward devaluation in rats following limited but not extended training. The development of this behavioural autonomy with extended training appears to depend not upon the amount of training *per se*, but rather upon the fact that the animal no longer experiences the correlation between variations in performance and variations in the associated consequences during overtraining. In agreement with this idea, limited exposure to an instrumental relationship that arranges a low correlation between performance and reward rates also favours the development of behavioural autonomy. Thus, the same activity can be either an action or a habit depending upon the type of training it has received.

1. Teleology and intelligence

The study of the cause of behaviour has been dominated by two models of the animal. The mechanistic model assumes that at some level of causal analysis the occurrence of a particular activity is to be explained simply by appealing to the presence of an eliciting, releasing or triggering stimulus. Such a model underlies the physiologist's concept of a reflex, the ethologist's notion of a fixed action pattern and the psychologist's conditional and unconditional responses; indeed, so pervasive is the basic assumption of this model that it is common to refer to any behaviour as a 'response' and thus by implication, and often without any evidence, assume that there must be an eliciting stimulus.

This stimulus–response model of the animal is typically contrasted with a teleological account, according to which a particular behaviour occurs, not because it is elicited by a releasing or triggering stimulus, but rather because it is controlled at the time of performance by the animal's knowledge about the consequences of this activity. In other words, the teleological model claims that, at least, some behaviour is truly purposeful and goal-directed, and to distinguish such activities from 'responses', we refer to them as 'actions'. Of course, there is some truth in both these accounts: much behaviour is response-like in character, whereas there is little doubt that we, at least, are capable of goal-directed actions.

The relevance of this distinction between responses and actions for the general question of animal intelligence arises from the fact that a teleological system provides an animal with a much more flexible form of behavioural control. More specifically, a teleological system allows an animal to adjust its behaviour immediately and appropriately to changes in the value of its goals brought about either by an alteration in the animal's motivational state or by the acquisition of new knowledge about the value of the goals. I shall illustrate this important point by considering simple examples of both these types of goal revaluation.

In the case of motivationally mediated revaluation, let us assume that an animal has learned the routes to two water sources while thirsty, one of which has a much higher saline content than the other. The question of interest is whether, having had this experience, the animal could select the route to the saline goal when salt-deficient. Clearly the teleogical model potentially allows for such a selection; having knowledge about the consequences of taking each route, the animal could choose the one leading to saline when the value of this goal is enhanced. By contrast, a simple stimulus–response account would not permit appropriate selection. The training received while thirsty should simply have strengthened the capacity of the various stimuli along each route to elicit approach without providing the animal with any knowledge of the goals.

An analogous example can be constructed in the case of knowledge-based goal revaluation. In this case the animal might have learned two different routes to a particular food source. If subsequently the animal found out by following one route that the food source had become contaminated, a teleological, but not a stimulus–response process would allow the animal to avoid any further selection of the other route as well. It is clear from these examples that teleological processes greatly enhance the flexibility and power of an animal's cognition and the consequent control of behaviour, and that the possession of such processes gives a quantum jump in general intelligence above that exhibited by simple stimulus–response systems. It is, therefore, an important question in the field of animal intelligence to decide whether creatures other than ourselves are capable of true purposive actions. Answering this question is not the easy matter it might at first appear.

2. Goal revaluation

The teleological status of a particular activity can be determined rarely, if ever, by simple observation. The ethological and natural history literature has many examples of apparently purposeful and goal-directed behaviour that on simple experimental analysis turn out to be elicited responses. As a result, many years ago animal psychologists developed a variety of assays, that generally fall under the rubric of 'latent learning', for determining whether or not behaviour is controlled by knowledge about the goal. And I am sure that many psychologists would regard the question of whether mammals, at least as represented by the laboratory rat, are capable of true actions is an issue that has been resolved for a quarter of a century or more. As an anonymous referee of a recent paper on this problem observed, 'were he alive, Tolman would be perplexed that someone would still consider the issue in need of further experimentation'. But a closer inspection of the classic studies in this area reveals that the question is far from resolved.

This point can be illustrated by considering one of perhaps the most compelling demonstrations of apparent animal teleology in operation in an ingenious irrelevant-incentive study conducted

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

by Krieckhaus & Wolf (1968). In effect, they implemented the assay for the teleological status of a behaviour suggested by the example of a motivationally induced change in the goal value that we have already considered. Two groups of thirsty rats were trained to press a lever, one for the sodium solution (Na) and the other for a potassium solution (K). Subsequently, the animals were sated for water and a sodium appetite induced before they were given an extinction test in which the rate of lever pressing in the absence of any reward was measured. Of course, if the animals trained with the sodium reward learn that lever pressing produces sodium, we should expect these animals to press more in the extinction test than those trained with the potassium reward.

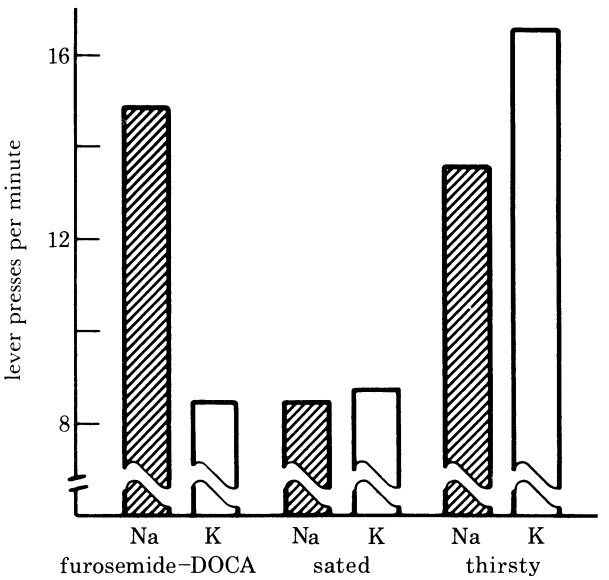

FIGURE 1. Mean lever press rates during the extinction test for groups trained with either the sodium (Na) or potassium (K) reward. Different groups were tested following the induction of a sodium appetite (furosemide-DOCA), sated for water and water-deprived.

That this prediction of the teleological account is fulfilled is shown in figure 1 which illustrates the results of a replication and extension of the Krieckhaus & Wolf (1968) study conducted in the Cambridge laboratory (Dickinson & Nicholas 1983). The rats trained with sodium pressed more in the extinction test than those rewarded with potassium when a sodium appetite was induced by the combined injection of the diuretic, furosemide, and desoxycorticosterone acetate (DOCA). This difference was not because the saline acted simply as a more effective reward. When other groups of rats were tested while they were either sated or thirsty, those trained with potassium performed at least as vigorously as those rewarded with sodium.

Plausible as the teleological account of this effect may seem, there is, in fact, a missing link in the empirical support. Krieckhaus & Wolf (1968) provided no evidence that the effect depends upon the animals having the opportunity to learn about the instrumental contingency or, in other words, the fact that lever-pressing causes saline delivery. If the same effect is seen whether or not the animals have the opportunity to learn about this contingency, the teleological account would be in trouble. To investigate the effect of the instrumental contingency, we trained two groups of rats to lever-press, one for the sodium solution (Na-W) and the other for, in this case, water (W-Na). In addition, however, both groups also received

the other solution, but non-contingently or independently of lever-pressing. In fact, strictly speaking, these presentations were not non-contingent because the schedule was designed to minimize the possibility that the animals might believe that lever-pressing caused saline presentations owing to chance pairings (for details see Dickinson & Nicholas 1983). This means that both groups received the same number of saline presentations during training but only for one group (Na-W) did this solution act as a reward for lever pressing.

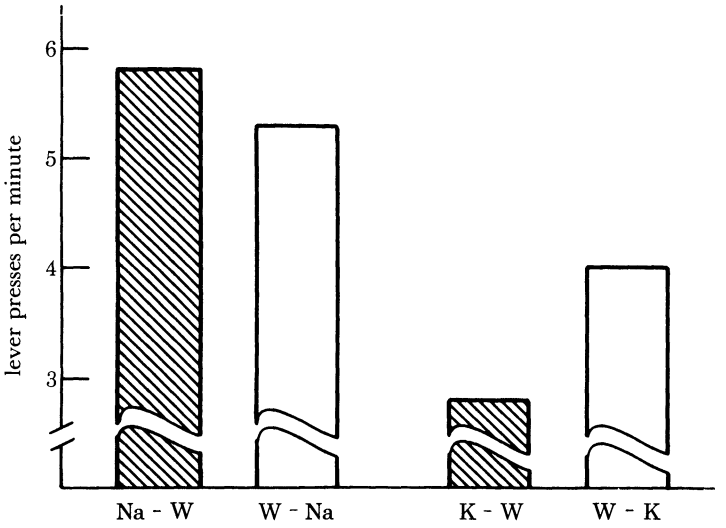

FIGURE 2. Mean lever press rates during the extinction test. Half of the groups were trained with either contingent sodium (Na-W) or potassium (K-W) and non-contingent water. The remaining groups received contingent water and either non-contingent sodium (W-Na) or potassium (W-K).

Given this training, the teleological account would anticipate essentially the same result as in the first study with animals that had the opportunity to learn that lever-pressing produced the saline (Na-W) performing more rapidly on test than those with no such opportunity (W-Na). In fact, we have consistently failed to detect such a difference when the amount of exposure to the sodium solution is equated as in the present study. Although group W-Na pressed at a slightly lower rate than group Na-W (see figure 2), this difference did not approach significance in the present study nor in other replications. Our failure to detect an effect of the instrumental contingency is not simply because the basic Krieckhaus and Wolf effect cannot be demonstrated with behaviour acquired on our complex schedule. Two further groups, groups K-W and W-K, were trained under exactly the same conditions as groups Na-W and W-Na, respectively, except that the potassium solution replaced the saline. When tested under a sodium appetite, groups K-W and W-K pressed significantly more slowly than the animals that received sodium during training. Any prior experience of saline, whether or not it is under the animal's control, elevates subsequent performance in the presence of a sodium appetite.

So what appeared to be a classic example of teleological control, on further experimental analysis turns out to be nothing of the sort. And, as far as I know, all the other previous demonstrations of latent learning either employ a behaviour, such as runway or maze performance, whose instrumental status is ambiguous (Mackintosh 1983), or fail to show that the effect depends upon the instrumental contingency. This is not the place to develop an explanation of the Krieckhaus and Wolf effect, although it is worth noting that our results would

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

have been anticipated by a sophisticated development of the stimulus–response model, two-factor theory (Rescorla & Solomon 1967; Trapold & Overmier 1972).

Our failure to detect teleological control in response to a motivationally induced change in the value of the training goal may imply that the rat's behaviour is not under such control. Alternatively, lever-pressing may be an action, but one based upon associative knowledge that does not encode the aspect of the goal, in this case its saltiness, that is changed by motivational manipulation. I favour the latter interpretation for there is good evidence that this type of behaviour can be brought under teleological control using the second technique for changing the value of a goal, namely the acquisition of new knowledge.

Some years ago Christopher Adams and I (Adams & Dickinson 1981*a*) trained hungry rats to lever press for one type of food, either sugar or mixed diet food pellets, while they received the other type non-contingently on the schedule employed in the previous study. We then devalued (D) the contingent or goal food in one group of rats, (D-N), while maintaining the value (N) of the non-contingent food. By contrast, the non-contingent food was devalued in the other group (N-D). This was done by giving the animals access to meals of the contingent and non-contingent food in the absence of the lever on alternate days after the lever press training had been completed. Animals in group D-N were averted from the contingent food by making them mildly ill immediately after they had consumed each meal of this food. The illness was induced by an injection of lithium chloride. Group N-D was similarly averted from the non-contingent food.

If lever pressing is an action, we should expect animals for which the contingent food is devalued to perform less vigorously when given the opportunity to lever-press again than those averted from the non-contingent food. In line with this prediction, figure 3 shows that group D-N pressed at a lower rate in the extinction test than group N-D. This difference was not because the aversion procedure had been more successful for the contingent than for the

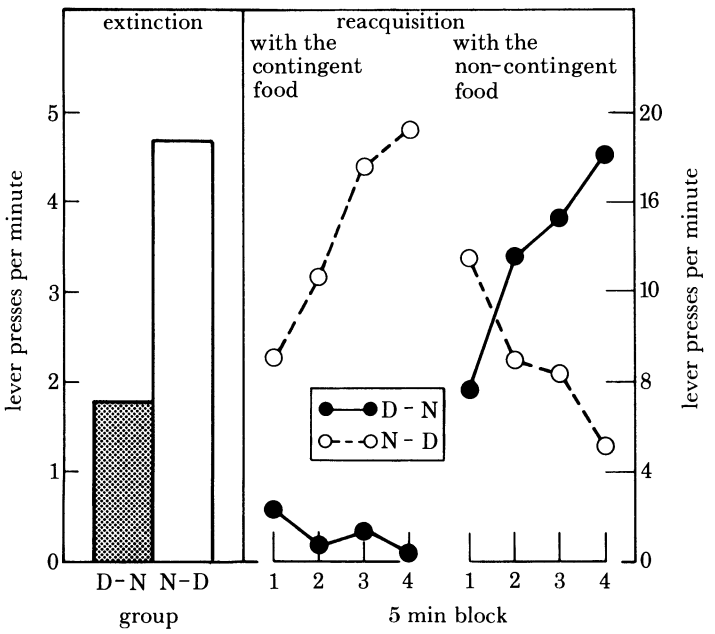

FIGURE 3. Mean lever-press rates during the extinction (left-hand panel) and reacquisition tests (right-hand panel) following the devaluation of either the contingent (group D-N) or non-contingent food (group N-D).

non-contingent food, which was in any case counterbalanced for type. Reacquisition tests demonstrated that both the contingent food in group D-N and the non-contingent food in group N-D had lost their capacity to act as rewards. The right-hand panel of figure 3 shows that during the first 20 min reacquisition test when lever-pressing produced the previously contingent food, performance was not re-established in group D-N, the group averted from this food. When a switch was made to the non-contingent food in the second reacquisition session, the decline in lever pressing by group N-D indicates that the non-contingent food had also lost its rewarding property for these animals.

I think that there can be little doubt, given this evidence, that the laboratory rat fits the teleological model; performance of this particular instrumental behaviour really does seem to be controlled by knowledge about the relation between the action and the goal. This conclusion will surprise few people. What might be more surprising is that the issue was not properly settled long ago.

3. Habits and the development of behavioural autonomy

The fact that an animal is capable of goal-directed actions does not imply that all its behaviour is under teleological control. Even the most purposive animals, such as ourselves, exhibit response-like behaviour. What then determines whether a particular behaviour is an action or a response? One might suppose that it is an immutable property of a particular behaviour to be either a response or an action, but a moment's reflection shows that this is probably not so. Perhaps the most obvious case in which apparently the same activity can be both an action and a response at different times is that of habit formation. The popular account of habit formation is that an instrumental behaviour, which starts out as an action controlled by knowledge about its relation to the goal, with repeated practice becomes a response, autonomous of the current value of the goal and simply triggered by the stimuli in whose presence it has been repeatedly performed. The response-like character of much of our own well-practised behaviour is revealed by what have been called 'slips of action' (for example, Norman 1981) that occur as a result of life's goal devaluation. Norman (1981), for example, quotes William James' (1890) well-known claim that 'very absent-minded persons in going to their bedroom to dress for dinner have been known to take off one garment after another and finally to get into bed, merely because that was the habitual issue of the first few movements when performed at a late hour'. For the middle class in the 19th century, the stimuli of the bedroom at that late hour tended to trigger a going-to-bed response even though its goal was inappropriate on that occasion or, in other words, devalued.

In spite of the strong anecdotal evidence that habit formation results from repeated practice, to my knowledge there is little or no experimental evidence in either humans or other animals that extended training renders behavioural control autonomous of the current value of the goal. Recently, however, Adams (1982) investigated the process of habit formation in the rat. By using our standard procedure, he trained two groups of rats to lever-press for sucrose pellets. One group was given a small amount of training, being allowed to perform only 100 rewarded lever-presses, whereas 500 presses were rewarded for the other, over-trained group. In both cases training was conducted at the rate of 50 rewards per session. We anticipated that pressing might have become a habit for the 500-press group through extended training and thus would be unaffected by goal devaluation. By contrast, this activity should have remained an action under the control of the current value of the goal for the 100-press group.

For half of the animals in the two training groups the reward was devalued (D) by pairing

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

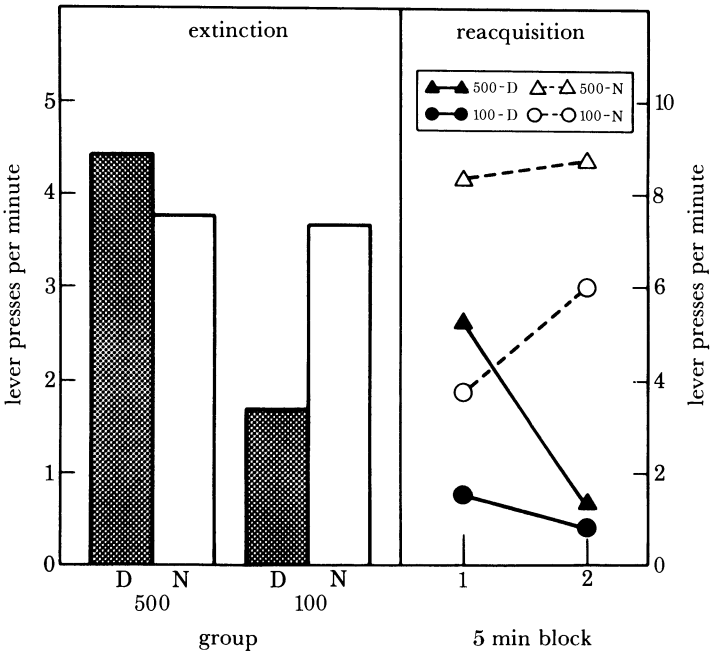

FIGURE 4. Mean lever-press rates during the extinction (left-hand panel) and reacquisition tests (right-hand panel) following either the devaluation (D) or non-devaluation (N) of the reward. The groups were allowed to make either 100 (groups 100-D and 100-N) or 500 rewarded lever presses (groups 500-D and 500-N) during training.

the consumption of sucrose pellets with the induction of illness, as in the previous experiment. The remaining rats acted as control animals for whom the reward was not devalued (N), although they received the same number of both meals of sucrose pellets and lithium chloride injections as the experimental animals, but on separate days. The left-hand panel of figure 4 shows that in a subsequent extinction test the group for whom the reward was devalued following 100 presses (100-D) performed less vigorously than the appropriate control group (100-N). Clearly, after a moderate amount of training, the lever-press remains an action under the control of the current value of the goal. By contrast, the outcome of the extinction test was completely different after over-training. After 500 rewards the rate of pressing by rats averted from the sucrose pellets (500-D), if anything, was greater than that of rats for whom the reward was not devalued (500-N).

This apparent difference in sensitivity to goal devaluation was not due to the ineffectiveness of the aversion procedure for the over-trained animals. The first 5 min block of a reacquisition test, not surprisingly, reproduced the pattern of result seen in the extinction test. Performance during the second block, however, shows that the sucrose pellets would no longer sustain lever pressing in group 500-D, indicating that this food had completely lost its rewarding property for these animals. Over-training can transform an action into a simple habit that is relatively autonomous of the current value of its original goal.

4. BEHAVIOUR–REWARD CORRELATION

The results of this study suggest that there might be some truth to the popular supposition that repetition produces habits. But it is not at all obvious that the failure of the animals allowed to make 500 presses during training to change their behaviour following goal devaluation was simply because they had pressed the lever so often. In the study comparing the effects of

devaluing contingent and non-contingent rewards (page 71), the animals performed at least 500 presses and yet the behaviour remained sensitive to reward devaluation. This observation led me to consider the other factors that are changed by overtraining.

I approached this problem by attempting to track the animal's changing experience as training is extended. Figure 5 shows the acquisition function for the animals trained for 100

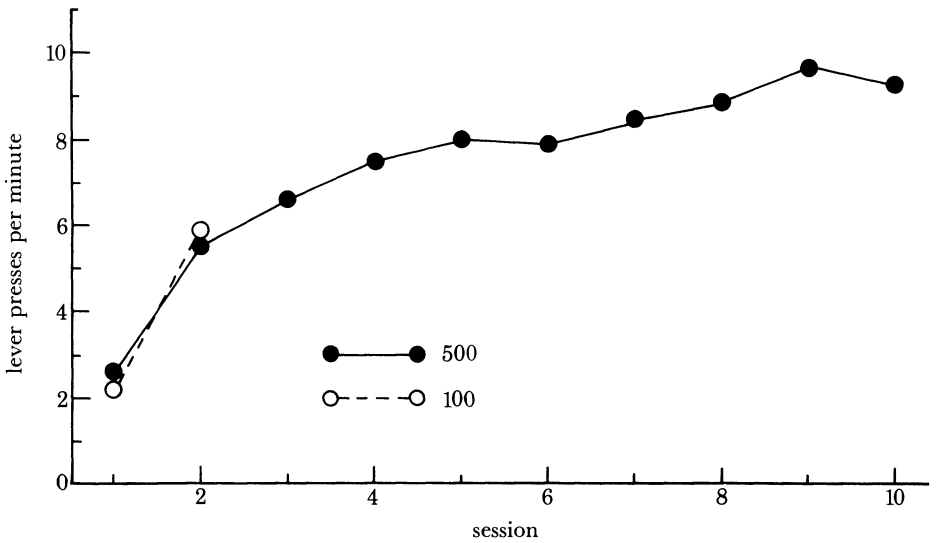

FIGURE 5. Mean lever press rates during acquisition for the groups allowed to make either 100 or 500 rewarded lever presses.

and 500 rewards. Extended training produced the typical negatively accelerated acquisition function in that the animals showed relatively large changes in the rate of performance across the initial sessions, but only small changes during the later sessions. This means that the animals experienced the relation or correlation between the rate of pressing and the reward rate over a large range of values during the first few sessions, but over only a restricted range later in training.

This point is illustrated more directly in figure 6 which portrays the relation between behaviour rate and reward rate. For the type of reward schedule used in the previous study in which each press produces a reward, a ratio schedule, there is, of course, a perfect correlation between behaviour and reward rates as represented by the linear function in figure 6; the faster that the animal performs the higher the rate at which the rewards occur. The points on this schedule function represent the relative average behaviour and reward rates on each session derived from the acquisition functions shown in figure 5. As can be seen, the animals in the over-training groups experienced the correlation between behaviour and reward rates over a wide range of values during the first five sessions, whereas their experience was limited to a much narrower range in the second five sessions. In fact, the variation in the animals' performance is so restricted during the later stages of training that they make little contact with the instrumental contingency which, of course, is defined by the way in which the occurrence of the reward varies with behaviour.

If we assume that an animal's performance is controlled by knowledge that directly reflects its current experience, then the development of behavioural autonomy follows directly from

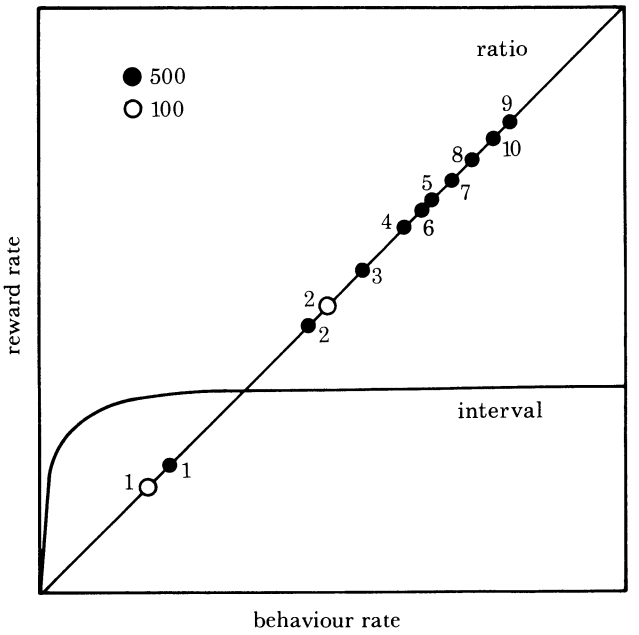

FIGURE 6. Reward rate as a function of the behaviour rate for ratio and interval schedules. The points on the ratio function represent the relative performance in each session of training for the groups allowed to make either 100 or 500 rewarded lever presses.

this analysis. When the animal's performance is varying during the early stages of training, it experiences, and thus stores knowledge about, a strong relation between behaviour and reward rates. Consequently, devaluing the reward depresses performance via this knowledge. By contrast, when the animals are over-trained, they no longer experience the behaviour–reward correlation with the result that their performance is no longer controlled by knowledge about this relation. In the absence of such knowledge, reward devaluation can have no effect and a habit has been established.

A clear prediction follows from this account of habit formation. If we found some other way of minimizing the animal's experience of a strong behaviour–reward correlation while maintaining performance, behavioural autonomy should result. A second type of reward schedule, an interval as opposed to a ratio schedule, provides such a procedure. Whereas on a random ratio schedule there is a fixed probability that each press will be rewarded, a random interval schedule arranges that a reward will become available with a constant probability in each time unit, for instance in each second. This reward then remains available until the next lever press collects it. Unlike a ratio schedule, which arranges a linear behaviour–reward rate function, this relation varies with the behaviour rate under an interval schedule. As figure 6 shows, at low levels of performance there is a strong positive relation between the behaviour and reward rates. As the behaviour rate rises, however, this relation weakens rapidly so that at modest levels of performance the reward rate is relatively unaffected by variations in the behaviour rate. This means that with minimal training on a random interval schedule the animal should no longer experience nor encode a strong behaviour–reward relation. Thus, an activity should be established as a habit much more rapidly on an interval schedule than on a ratio schedule.

To test this prediction, we (Dickinson *et al.* 1983) gave a group of rats limited training on

random interval schedules. During the final training session these animals pressed on average at a rate of 12.4 per minute, a rate well within the range across which variation in performance has little effect on reward rate for this schedule. When the computer procedure that actually controlled the random interval schedule used during the final training session was driven at different rates by a mechanism that produced artificial lever-presses randomly in time, it was found that a fivefold increase in performance from 5 to 25 presses per minute only increased the reward rate, on average, from 0.92 to 1.06 pellets per minute. This 15% increase contrasts with a 400% rise in reward rate that would be produced by the same increment in performance on a ratio schedule. Consequently, we should expect lever pressing to be more readily established as a habit following this interval training than following an equivalent amount of training under a ratio contingency.

To check that any behavioural autonomy observed in these animals was owing to the lack of a behaviour–reward rate correlation rather than some other feature of the training, another group of rats were given the equivalent amount of training in terms of the number of rewards received on a ratio schedule. Furthermore, an attempt was made to match other features of the ratio and interval training, namely the probability that a press would be rewarded and the reward rate, by using a yoking procedure (for details see Dickinson *et al.* 1983). On the whole the matching was successful in that there was no significant difference between the reward probability in the two conditions, although the reward rate was, on average, 28% higher during the ratio training. I think, however, that we can be fairly confident in attributing any difference in behavioural autonomy following ratio and interval training to the behaviour–reward rate correlation arranged by the two schedules rather than some other feature of the contingencies.

The pattern of performance in an extinction test following reward devaluation by the food-aversion procedure confirmed our expectation. The left-hand panel of figure 7 shows that we replicated the basic reward devaluation effect following ratio training (R) in that animals (R-D) for whom the reward had been devalued (D) pressed at a lower rate than the control rats (R-N) for whom the reward was not devalued (N). More importantly, figure 7 also shows that performance following interval training (I) was impervious to reward devaluation; there was no detectable difference in the rate of pressing by animals for whom the reward had (I-D) and had not been devalued (I-N). The fact that the control animals trained on the ratio schedule (R-N) pressed more vigorously than those trained on the interval contingency (I-N) is a well-documented finding, although its explanation is currently a matter of dispute. The results of the reacquisition test again show that the extinction performance cannot be explained in terms of the differential effectiveness of the reward devaluation procedure. For both animals trained on the ratio (R-D) and interval schedules (I-D) the devalued sugar pellets would no longer act as an effective reward (see right-hand panel of figure 7).

These results suggest that, contrary to popular belief, habit formation is not a simple consequence of over-training or practice. Rather it appears to arise because over-training typically tends to reduce the variation in behaviour and thus the animal's experience of the relation that controls actions, namely the behaviour–goal correlation. Other ways of preventing the animal experiencing this relation once performance is established, such as training on an interval schedule, have the same effect. In the absence of the relevant experience, it is perhaps not so surprising that performance is no longer controlled by knowledge about the behaviour–goal correlation.

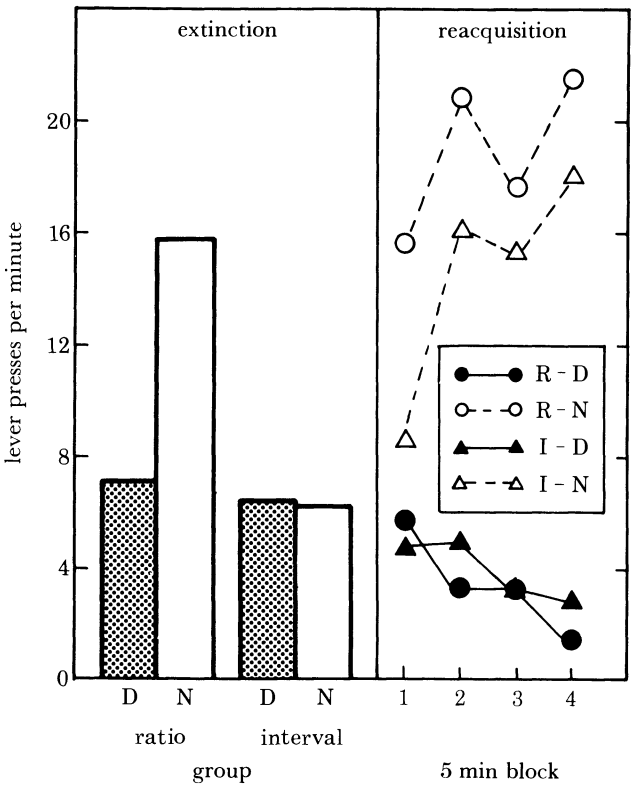

FIGURE 7. Mean lever press rates during the extinction (left-hand panel) and reacquisition tests (right-hand panel) following either the devaluation (D) or non-devaluation (N) of the reward. The groups were trained on either a ratio (groups R-D and R-N) or an interval schedule (groups I-D and I-N).

5. CONCLUSIONS

The conclusion drawn from the classic studies of latent learning was correct even if the outcome of these experiments did not necessarily justify this claim. The selective effect of devaluing the contingent reward on performance demonstrates that the laboratory rat, at least, is capable of purposive goal-directed actions. Moreover, our findings show that instrumental behaviour does not represent a homogeneous category. By varying the training conditions appropriately, the same behaviour in the same species can both show behavioural autonomy and be brought under teleological control. This conclusion has implications for the study of comparative intelligence. For instance, we do not yet know whether the two types of control can exist independently of each other or whether they are both based upon a common learning process. If the former possibility is true, comparative psychology is not faced just with determining whether the species of interest is capable of instrumental conditioning, but also with the problem of identifying the type of instrumental control.

From my point of view, however, the main implication of our results concerns the nature of the cognitive processes controlling instrumental behaviour. My colleagues and I (Adams & Dickinson 1981*b*; Dickinson 1980; Mackintosh & Dickinson 1979) have argued that teleological control of instrumental behaviour cannot be explained, at least at the psychological level, in terms of internal associations which have just excitatory or inhibitory properties. Rather, we

argue that the knowledge about the action–goal relation must be encoded in a propositional-like form so that it can be operated on by a practical inference process to generate the instrumental performance. In this sense actions are inherently rational in a way that responses can never be.

The work reported in this paper was supported by a grant from the S.E.R.C. I would like to thank N. J. Mackintosh, D. Charnock and C. Preston for their comments on an earlier draft of this manuscript.

References

Adams, C. D. 1982 Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. Jl exp. Psychol.* **34** B, 77–98.

Adams, C. D. & Dickinson, A. 1981*a* Instrumental responding following reinforcer devaluation. *Q. Jl exp. Psychol.* **33** B, 109–112.

Adams, C. D. & Dickinson, A. 1981*b* Actions and habits: Variations in associative representations during instrumental learning. In *Information processing in animals: memory mechanisms* (ed. N. E. Spear & R. R. Miller), pp. 143–165. Hillsdale, New Jersey: Erlbaum.

Dickinson, A. 1980 *Contemporary animal learning theory*. Cambridge: Cambridge University Press.

Dickinson, A. & Nicholas, D. J. 1983 Irrelevant incentive learning during instrumental conditioning: the role of the drive-reinforcer and response-reinforcer relationships. *Q. Jl exp. Psychol.* **35** B, 249–263.

Dickinson, A., Nicholas, D. J. & Adams, C. D. 1983 The effect of the instrumental contingency on susceptibility to reinforcer devaluation. *Q. Jl exp. Psychol.* **35** B, 35–51.

James, W. 1890 *The principles of psychology*. New York: Holt.

Krieckhaus, E. E. & Wolf, G. 1968 Acquisition of sodium by rats: Interaction of innate mechanisms and latent learning. *J. comp. physiol. Psychol.* **65**, 197–201.

Mackintosh, N. J. 1983 *Conditioning and associative learning*. Oxford: Oxford University Press.

Mackintosh, N. J. & Dickinson, A. 1979 Instrumental (type II) conditioning. In *Mechanisms of learning and motivation* (ed. A. Dickinson & R. A. Boakes), pp. 143–169. Hillsdale, New Jersey: Erlbaum.

Norman, D. A. 1981 Categorization of action slips. *Psychol. Rev.* **88**, 1–15.

Rescorla, R. A. & Solomon, R. L. 1967 Two-process learning theory: relationship between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* **74**, 151–182.

Trapold, M. A. & Overmier, J. B. 1972 The second learning process in instrumental learning. In *Classical conditioning*. Vol. II. *Current research and theory* (ed. A. H. Black & W. F. Prokasy), pp. 427–452. New York: Appleton-Century-Crofts.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES