

4

Operant conditioning: reward and positive reinforcement

Operant conditioning has to do with reward and punishment, with achieving goals and avoiding disasters. In Skinner's terminology, goals, rewards and incentives may all be referred to as *positive reinforcers*; achieving the goal or receiving the reward is *positive reinforcement*. Escaping from unpleasant or dangerous situations is classified as *negative reinforcement*. Reinforcement is thus always the occasion for things becoming better than they were, but is divided up according to whether it is some new good thing which happens, or something bad' which goes away. Punishment is distinguished from reinforcement because it is a moment when things get definitely worse, either through loss of positive reinforcers, as in fines or confiscations, or through the onset of an aversive state of affairs such as physical pain or social rejection.

The lynch-pin of the Skinnerian system (see Ch. 1) is positive reinforcement, which allows behaviour to be changed by the influence of attractive consequences. Both negative reinforcement and punishment involve some degree of *aversive control*, which is the use of unpleasant stimuli to modify behaviour, and they will be left to the next chapter.

Positive reinforcement is of interest for two separate reasons; first as one of the most powerful techniques we have at our disposal for directing or motivating the actions of other people or animals, either in the laboratory or in the outside world. The second reason is almost a philosophical one; the versatility of the *concept* of reinforcement as an explanation of

46

behaviour. The answer to the question 'why do people behave as they do?' can often be given in the form 'because they are reinforced for it'. People can be said to work for the reinforcers for working, and play because of the reinforcers for playing. The main advantage of giving a preliminary answer in this way is that it prompts the further question 'what *are* the reinforcers?' in any particular case, and this may be something which can be determined by the traditional scientific means of observation and experiment. It remains to be seen how far this

explanatory use of the positive reinforcement concept can be justified, but Skinner (1953) has given analyses of almost every area of human activity. The reinforcement idea has recently been taken up in both clinical and social psychology (F3, B1) but the original examples of carefully measured operant conditioning came from the animal laboratory.

Operant learning in various forms

Whenever rewards are given or behaviours change according to their usefulness, the essential features of positive reinforcement are present. Operant learning therefore takes in a very broad sweep of circumstances in which separate areas have their own special characteristics. A brief survey of several kinds of training or experience in which positive reinforcement plays a part is given below.

Shaping operant responses by successive approximation Figure 4.1 illustrates the arrangement used by Ferster and Skinner (1957) in their extensive research on positive reinforcement (see Ch. 6). The reinforcer is food, given to a hungry animal. This is by far the most frequently employed incentive in animal experiments since it is convenient, harmless and very effective. However there is nowadays a greater interest in studying a variety of reinforcers to see which works best for a particular response (Hinde and Stevenson-Hinde, 1973). A number of rewards have been explored, including access to the opposite sex, opportunities for exercise and the delivery of bits of paper to make nests with (see Ch. 7). Returning to the apparatus in Figure 4.1, the all-important reinforcer is in this case the availability of grain for a few seconds

47

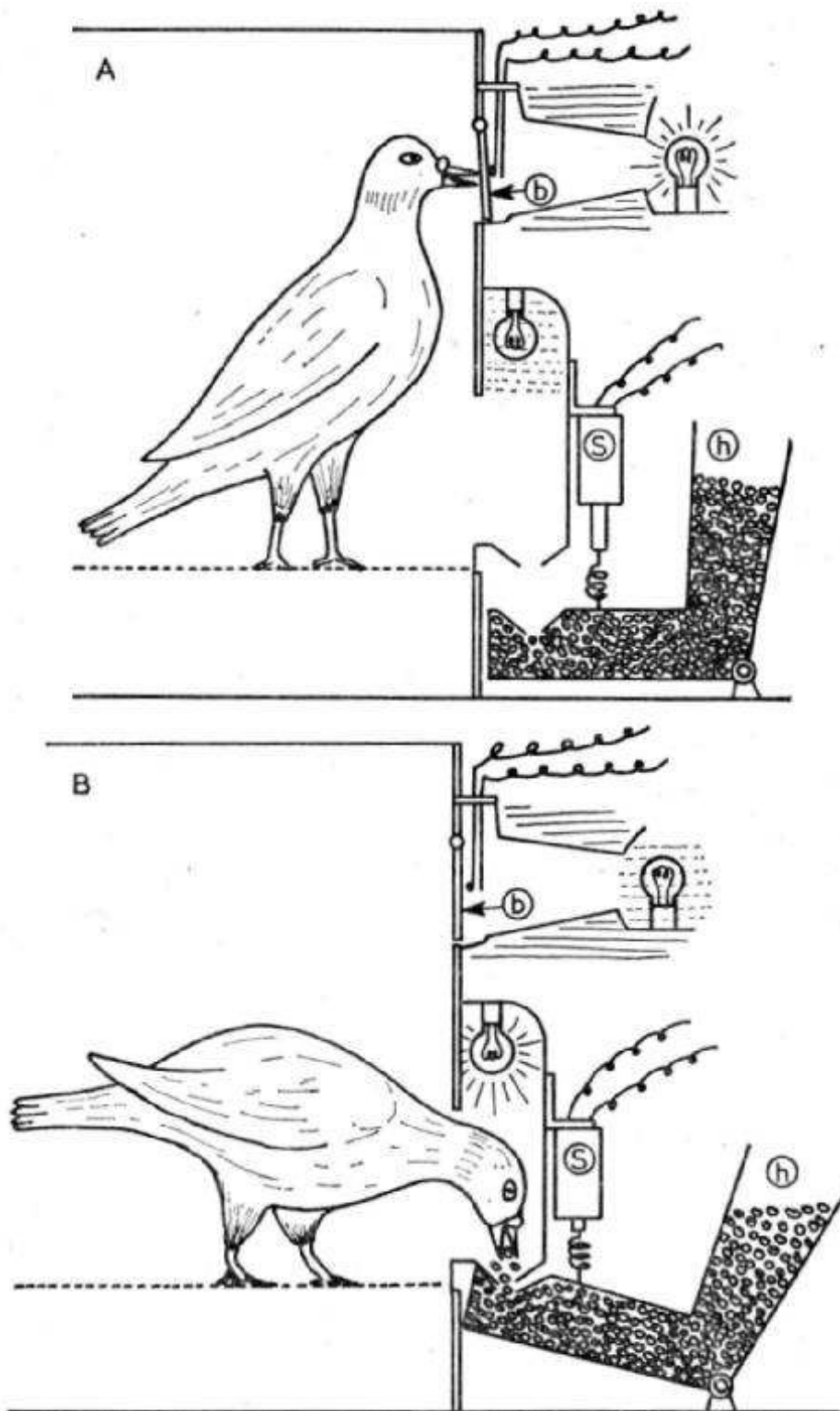


Fig. 4.1 *Operant conditioning apparatus for pigeons.* In A the bird has just made contact with the pecking button (b). In B this peck is reinforced: current is supplied to a solenoid (s) which lifts up the grain hopper (h) for a few seconds. (After Ferster and Skinner, 1957)

48

at a time, when a food-hopper is brought within a pigeon's reach.

As a prerequisite for any response shaping the subject must be adapted to the experimental situation and become accustomed to eating out of the food magazine (*magazine training*). A hungry bird soon becomes adept at recovering grain as fast as possible while it is presented. By this point, if food is available for only three seconds out of every minute, the pigeon will waste no time getting down to eat the food as soon as food is signalled by the sound of the

mechanism and the lighting up of the hopper. The sound and light are said to have become *discriminative stimuli* for getting to the hopper and eating. Now the very strong influence of food delivery on the pigeon's behaviour patterns can be demonstrated. The most commonly studied response of the bird in this context is that of directing a strong peck towards a flat recessed button on the wall. This records the response automatically and delivers the reinforcer. To persuade the pigeon to 'make this response by successive approximation an experimenter must watch the bird carefully, and make a series of decisions about what the bird must do to earn a few seconds access to food. To start with, the experimenter might deliver food (by pushing his own button to activate the mechanism) if the bird raised its head to within a few inches of the pecking button. This usually has a quite dramatic effect, the bird quickly returning to the posture required by the time four or five reinforcements have been given. Now the experimenter might wait until the pigeon makes a movement towards the pecking button before delivering food. This results in the bird repeating the movement, and then the criterion for reward can be made closer and closer to a real peck. Sooner or later the bird succeeds in operating the pecking button itself, and it can then be left to feed itself automatically, if it gains a small amount of food every time it pecks the button.

Although, as outlined in the next section, any form of pecking is an easy response to shape in birds, the value of the response-shaping method is that it can be used to induce a wide variety of behaviour patterns, provided the reinforcer is powerful enough, and progress is made gradually. Shaping is a major tool in the training of handicapped or retarded people by operant methods, especially in combination with some form of prompting (see below). It is especially appropriate as a method of teaching when other forms of communication are

49

impossible. I recently saw shaping being used to teach a severely retarded blind child to operate his wheel chair. A spoon of icecream was held just in front of his mouth as he sat in the chair, so that a slight movement of his hands to turn the wheels of his chair forward was reinforced by contact with the icecream. After starting off with the criterion that just the placing of his hands on the wheels was rewarded by giving icecream, greater and greater success in moving the chair was needed as the icecream was held further and further away. This is an example of how positive reinforcement can provide both information and incentive at the same time. Getting icecream and social approval can be an encouragement for the task, and at the same time the prompt delivery of positive reinforcers supplies information about the correctness of target responses in a way that resembles *feedback* for response skills.

Autoshaping, prompting and guidance The gradual-shaping procedure can sometimes work quickly, but often requires a good deal of patience and skill on the part of the shaper. He must wait for the subject to make the appropriate response, but must also make sure that enough rewards are given to maintain interest. Measures to get the correct response without prolonged shaping have always

been sought; rats may be attracted by cheese smeared on a lever and there are several short-cuts for training pigeons (Ferster and Skinner, 1957).

Autoshaping. Another method of ensuring that animals come up with a response has been devised by Brown and Jenkins (1968). The general idea is to attract attention to a stimulus source by using it as a signal for food, and to wait until the subject makes some kind of response *directed at* the stimulus source. With pigeons, it is sufficient to light-up the pecking button for a few seconds before the grain hopper operates. As we would expect from classical conditioning experiments (see p. 36), the birds cannot remain indifferent to a stimulus that signals food, and after about fifty trials (fifty pairings of the light signal with the reinforcer, with an average of a minute or so between pairings) they peck at the light sufficiently strongly to operate the button. If operating the button pays off with immediate food delivery this will strengthen the tendency to respond, but even if there is no pay-off for responding, the

50

tendency to peck at the food signal is pronounced. This *autoshaping* procedure is a mixture of stimulus-learning (classical conditioning) and response shaping (operant conditioning) which makes use of the investigative responses of the subject (Jenkins, 1973). Thus subjects can be lured to a particular location by a light source that signals reward, and this acts as a form of response shaping. Pigeons peck at the signal source, rats and dogs poke at it with their nose or paw, and monkeys or people may grab it with their hands.

Prompting and guidance. Autoshaping can be viewed as a form of *prompting*, which is a term used liberally for a number of additional techniques of facilitating a certain activity including even physically pushing and pulling the subject through the required movements (which can also be referred to as *guidance* or *putting through*). Many operant training procedures are combinations of shaping by approximation with a variety of prompts. For instance a combination was used to inculcate the social behaviour of greetings in some children in a home for the retarded (Stokes, Baer and Jackson, 1974). The positive reinforcers were candies, potato crisps and the social rewards of a smile and a pat on the head. The goal was to get some rather withdrawn boys to smile, wave and if possible say hello when they met someone; improvements in such basic social skills often reap additional benefits in social adjustment and interaction. Initially, a very low response criterion and physical prompts were necessary. The experimenter greeted the boy, and in the absence of any spontaneous reaction, gently pulled the boy's arm back and forth in a crude waving motion, as a *physical prompt*, before giving some crisps or sweets. After some training like this, a *visual prompt* was added to assist the learning of a freer 'wave': sweets were waved to and fro slowly, followed by the subject's hand, before they were delivered, along with social approval and encouragement. Gradually more realistic greeting responses were required, and the final stage of training was to employ a different training person, with the emphasis on social reinforcers, so that the

greeting behaviour generalized to normal meetings with other people in day-to-day interactions.

Prompts which require some degree of comprehension by the subject include imitative prompts, or 'showing how', and instructions of the form 'do this' or 'do that' (see Ch. 9). If

51

ordinary instructions and explanations are sufficient to determine future behaviour patterns, there is of course no need for special training procedures. But for most skills practice is necessary even if detailed verbal advice is available (you can be 'told how' to ride a bicycle without much benefit). And teachers and parents soon find out that instructions which sound easy enough may require some motivational emphasis if they are to be followed. Combinations of instructions and positive reinforcement are often called for because of these factors, but are always prominent when the normal course of educational or social inducements has proved inadequate. Allyn and Azrin (1964) found a case where neither instructions nor a simple increase in positive reinforcement worked to bring about a change in behaviour. Mental patients had lost the habit of picking-up their cutlery in a ward dining-room run on the cafeteria system. Offering 'extras' (additional cigarettes, cups of coffee etc) when patients remembered their cutlery, without explanation, did not make any difference. Reminding the patients ('please pick up your knife, fork and spoon') at each meal helped at first, but the effect was short-lived. However, giving verbal reminders plus rewards for successful behaviour led to almost complete recovery, in all twenty patients involved, of the expedient of collecting their cutlery. Since they used the cutlery when they were in possession of it, but tried to manage without cutlery if they had forgotten it, the extra training improved general dining behaviour considerably.

Often when prompts are used in initial training, the final goal will require the subject to act without help from this source. Leaving out prompts so that the subject responds on his own, or *fading*, is one aspect of the gradual shaping method. It has to be done with care because too much prompting may produce an awkward dependence on the prompts, and fast removal of the prompts may make the task too difficult. For instance, children may be helped to use spoons by an adult holding the spoon as well, and doing most of the work to start with. Especially with handicapped children, there may have to be very gradual fading of this prompt, so that the child does eventually learn to feed himself, but does not give up along the way.

Schedules of reinforcement A tremendous amount of work has been done on the effects of

52

schedules of reinforcement on animal behaviour. The experiments are usually controlled automatically (nowadays by a computer) and the schedule of reinforcement is the automatic rule about when positive reinforcers are delivered. Obviously simple rules or *schedules* have been studied most. The simplest schedule of all is that where a single behaviour is measured, and every response gets a reinforcer. In Skinner's original experiment, a food pellet was dropped into a bowl for a rat to eat every time the rat pressed down

a bar. This is called fixed ratio 1 (FR 1) or continuous reinforcement (CRF). For fixed ratio 2 every second response brings down the food pellet, for fixed ratio 3 every third response, and so on. All these are *fixed ratio schedules*, in which there is an exact relationship between the number of responses made and the number of rewards given. Spacing out the reinforcers without requiring very much behaviour can be done with *fixed interval schedules*. Here only one response is actually needed to get the reinforcer, but responses don't work until a certain length of time has passed since the last reward. Usually what happens is that animals do not wait until reward is obtainable, but respond during the interval until the reinforcer is delivered. On fixed interval (FI) schedules the interval is the same every time, and animals learn to respond more vigorously as reward becomes due. To produce a steadier rate of response the intervals can be made of unpredictable length so that animals learn to bash away at their response very regularly, because the reinforcer might become available at any moment. With these unpredictable intervals, the procedure is called a *variable interval schedule*. It is often employed to ensure a stable behavioural *baseline*. Less often used, because training with it is more difficult, is the variable version of the fixed ratio, the *variable ratio schedule*.

Note that in these intermittent schedules of reinforcement reinforcers do not have to be given for every response.

Response skills It is often possible in everyday situations to distinguish between the motives for attempting a task, and the factors which allow for mastery of it. In many competitive sports for instance, the reinforcers for engaging in them may have to do with the excitement of the competition, the joys of winning, or the social fringe benefits, whereas the learning and preparation for taking part may involve many hours of tedious, painful or lonely

53

practice. The reinforcers for acquiring a skill need not be the same as those obtained by exercising it. Rewards given for success may be sufficient for inducing further effort without supplying a noticeable increase in proficiency : no amount of celebration surrounding the achievement by a golfer of a hole-in-one is likely to improve his swing, although it may encourage him to spend much more time playing. On the other hand the swing may be improved by certain prompts, such as criticism of style, changes of grip and so on, which are not much fun in themselves.

In some cases, however, the immediate effect of responses is the single most important factor in the further development of skill, including cases where the main result of a response is to produce a positive reinforcer. *Response differentiation* by selective reinforcement is an example of this.

Response differentiation. The lever-pressing apparatus for rats provides a method for studying the learning of precision movements. A rat may press down the lever with one or other of its paws, or its mouth, from various angles or positions. These are called variations in response topography. Easier to measure are the quantitative aspects of lever movement, such as the exact force exerted or distance moved which are normally pretty unpredictable.

Both response topography, and quantitative features of lever pressing, may be refined by selective reinforcement. Very fine response differentiations can be recorded when quantitative methods are used. Rats may be conditioned to press down the lever to a certain angle, or to press with a certain force, or for a certain duration.

Biofeedback. It is generally true that selective reinforcement will encourage behaviour in the required range and that these tasks are made much easier by appropriate external feedback. The enhancement of behaviour by external feedback is nowhere more evident than in control of biological functions by the method known as *biofeedback*, which seems to enable a degree of voluntary control over the activity of internal organs. It is not normally sufficient merely to try to bring about a particular rate of heart beat, or a particular state of the brain which produces the 'alpha rhythm' form of brainwaves. This is partly because it is difficult to know when success has been

54

achieved, and there can be no reinforcement or feedback from good responses. A remedy has been found in the electrical measurement of the target biological function, and the provision of a clear external signal for success and failure. Such 'bio-feedback machines' are now commercially available as shortcuts to meditation or relaxation. Of course connecting up someone to one of these machines will not produce a state of meditation unless he makes an effort to produce the correct signal, that is unless the correct signal acts as a reinforcer. After being connected up one must follow the instructions to try and 'make the needle stay on the right' or 'try and produce the low pitched tones' and this takes some practice.

When trying to relax, it may help to 'make the mind go blank' or concentrate on a peaceful image, while lying or sitting in calming positions. When trying to increase arousal it is possible to change breathing patterns, tense muscles, imagine nightmarish situations.

Would biofeedback work without any of these stratagems? Attempts were made to answer this question by doing experiments on rats temporarily paralysed by a curare-like drug, and artificially respired, in the hope of finding results not due to muscle movements made by the animals. Such internal activities as intestinal contractions and heart-rate increases or decreases were selectively reinforced by electrical brain stimulation (see Ch. 7). The original results from such experiments suggested a high degree of direct control of such functions by electrical brain stimulation, which was assumed to have combined feedback and reinforcing functions. But further experimentation has modified the initial conclusions somewhat, since the effect is not as powerful as was first thought. The effect is indirect because it depends on previous experiences of the animals before they were paralysed (Miller and Dworkin, 1974). Current opinion supports the conclusion that biofeedback works indirectly, by allowing for reinforcement of any response strategem or internal activity which helps produce the target behaviour. This type of reinforcement can assist in behavioural therapy for physical symptoms, as in the re-training of abnormal heart rhythms in cardiac patients (Brener, 1973). Patients

are trained to produce normal heart rhythms with external feedback and the feedback is then faded out (in the same way that prompts are gradually removed) so that more normal heart functioning is maintained outside the training laboratory.

55

Response timing. Correct timing of responses is an important part of most skills and the direct reinforcement of responses made at certain intervals has been much investigated in the operant conditioning apparatus. A procedure called the *differential reinforcement of low rates* (DRL) ensures that responses are spaced apart in time. If responses are made too frequently no rewards are given, but any response made at the correct time after the previous one (fifteen seconds for example) delivers the reinforcer. Very elaborate sequences of accurately-timed responses can be observed. It is possible that internal timing of responses is the explanation for many of the effects of intermittent reinforcement schedules (Blackman, 1974, and Ch. 6).

Spatial learning: finding food and finding out

Testing the ability of animals to find their way through mazes is a well-tryed form of psychological experiment. In their natural environment many animals exhibit astounding navigational abilities (like the return of salmon to their native streams and the 'homing' of pigeons). Most laboratory feats are much less impressive, but provide the basis for quantitative study.

Alley running. Although rats should be good at remembering where food is, it actually takes them a considerable amount of time to reach a final level of performance in a task as simple as running from one end of a straight alley to another to get to food. It might take twenty or thirty trials before the rat runs down the alley at its maximum speed (three or four feet per second), although the most rapid change in behaviour would take place over the first five or ten trials. If food is left out of the goal box (*extinction*: Ch. 6) there would again be a fairly rapid change in the first ten experiences; this time a slowing down, which would continue for many more trials before reaching a stable level. One concept clearly illustrated by alleyrunning is that of *incentive*, which corresponds roughly to the degree of enthusiasm attached to reinforced behaviours. In alleyrunning this is expressed as speed. In subjective terms this eagerness might be a product of an expectation of or hope for the reward, but in the practicalities of the experiment it simply reflects the tastiness or size of the rewards previously given. Rats given minute amounts of food in the goal box will increase their speed of running towards the goal very gradually, and will never bother to run very fast. Rats given large tasty portions

56

at the end of their run react a great deal more energetically. In Hullian theory (see Ch. 1) this effect of *amount* of reinforcement was referred to as *incentive motivation*; nowadays any effect on the vigour or energy of reinforced behaviour due to the quality or quantity of reinforcers is conveniently put in the incentive category. This view of incentive as an emotional anticipation of reinforcement is supported by rapid changes in mood apparently produced by

changing the amount of reward given. If training takes place with a large reward, switching to a medium reward has a depressing effect, whereas with some training routines, switching to a medium reward after training with a very small reinforcer produces an extra degree of incentive (Crespi, 1942). Another factor which moderates enthusiasm in alley-running is the *delay of reward*. If the animal has to wait at the end of the run for some time before food is dropped in, its speed of running diminishes. In the response-shaping technique it is usually valuable to give reinforcers as soon as possible after the target response, in case some other behaviour intervenes and the reward is wasted. But even if reward is contingent on the correct response it is obvious that having to wait for the reinforcer will decrease its incentive value.

Choice and cognition in mazes. The element of choice can be included in the running task by releasing a laboratory animal from the bottom end of a T-shaped maze with food at only one of the ends of the cross bar. If sufficient geographical information is available in the form of landmarks - recognizable objects and constant sources of light or smells - rats will learn the food is 'over there' rather than remembering any particular route to get to the food, although details of route are important in more complex mazes of the Hampton Court type. That the form rather than the substance of a maze problem can be remembered is shown by the ability of rats to swim through a maze which they have learned by walking. In the early stages of learning rats are hesitant about making a decision at a *choicepoint*, where they can go either left or right. They have a marked tendency to make one or two steps in one direction, give a couple of sniffs, and then withdraw back to their original position. Tolman dubbed this behaviour *vicarious trial and error* as it seems to consist of a small scale testing out of the choices of turning left or right. It emphasizes the fact that positive reinforcement may produce dilemmas and that choice between

57

two activities that have equal incentive may be difficult. Rats have more difficulty learning a T-maze when there are four pellets of food on the left and three pellets on the right, ; than when the choice is between four and zero. Thus, although 'knowing where we are is mainly a function of the familiarity of the ground, and 'cognitive maps' of geographical information can be built up in the absence of specific goals, it is the importance, or incentive value of specific goals which most influences , decisions about which way to turn next. '

Creative responses

One of the limitations of the shaping procedure as an educational technique is that the learner does only what he is trained ;to do, which might not inculcate the valuable qualities of imagination and initiative. It is often held that positive reinforce- •ment is an inherently narrowing influence on behaviour, and ;cannot assist in developing intelligent or original responses.This narrowing aspect of operant learning is almost certainlya matter of the uses to which positive reinforcement is normallyput, rather than an inherent limitation. There is certainly noreason why prizes cannot be given

for originality to give incentive for creativity, or why teachers and parents should not give approval for novel activities. If originality does not result, this may be because of the inherent difficulty of the task, rather than the ineffectiveness of the encouragement. Provided the demands made are not too great, it is possible to use reinforcement procedures to train a subject to make a *different* response, rather than to make the *same* response. Not much inventiveness is usually expected of rats, but it is easy enough for them to learn to press a different lever than the last, instead of the same lever as the last, when they have the choice of two (Foster *et al.*, 1970).

Actual invention of new acrobatic tricks seems to have been achieved by porpoises trained to produce novel responses (Pryor *et al.*, 1969). With one porpoise, the plan was not to induce originality, but simply to shape-up a new trick (selected by the trainer) every day for the purpose of public demonstration. A reward of fish signalled by a whistle was given for successive approximations to the desired trick. However, after several days of this the animal pre-empted shaping by coming up with 'an unprecedented range of behaviours' off her own bat. The effect was repeated for thirty-two sessions with a second porpoise,

58

until the novel patterns of aerial movement achieved by the subject became too complex for objective description.

The result suggests that, at least for porpoises, 'novelty' may be reinforced. Another way of putting this is to say that a *class* of behaviour may be subjected to the training procedure. The class that is specified by stating that behaviour each day must be novel is probably very difficult initially, because the nonreinforcement of previously 'good' behaviour is frustrating. But once the response class has been established, or the principle learned, this problem is not so important. *Generative response classes and response generalization* A frequent objection to the idea that positive reinforcement applies to human behaviours has been that external rewards and punishments do not seem to explain those many areas of human accomplishment where *rules* appear to be of greater importance than specific items of conduct. It is obvious that we are able to learn response classes such as 'being polite' and 'being aggressive', as well as 'being original', which may reveal themselves in unaccustomed circumstances. More fuss has been made about the business of rules from the point of view of linguistic analyses of speech than in other areas (see Ch. 9) and the term *generative response class* has arisen from this. -If the reinforcement of a limited number of specific responses appears to be sufficient to establish the pattern of making similar responses in different, but appropriate, situations, a generative response class has been learned. This is distinct from having 'learned the rule', in that we (as well as animals) may be able to perform generative response classes without being able to say what the rule is. Conversely, we may 'know the rules' without being able to play, in the sense of being able to recite the principles of how to land a lunar module, or make the perfect tennis backhand, with no personal proficiency at these performances.

Several experiments have made it clear that the formation of

generative response classes in verbal behaviour may be assisted by positive reinforcement. External reinforcers may not be critical during language development in infants (see Ch. 9) but can be used for experimental or remedial purposes. A major yardstick of fluency is the number of words strung together; babies start off with one-word utterances and take some time to develop long sentences. Grammatical rules or

59

generative response classes include getting words in the right order, and adjusting the endings of words to fit in with the rest of what is being said (see Ch. 9). One of the easiest rules about word-endings, in English, is putting an 's' on the end of plural nouns. Most children get used to this very soon, and will apply the rule to new words, sometimes incorrectly, as in 'mouses'. However, a child who does not say plurals properly can be helped to learn this response class by extra rewards for correct individual cases. After being praised, and given bites of food just after saying 'horses', 'cars' and 'shoes' when shown collections of these objects, a child may be better able to give the correct response when shown buses, dogs or hats, provided of course that he knows the correct singular term.

This kind of training for an eight-year-old autistic boy was described recently by Stevens-Long and Rasmussen (1974). The boy was given food and praise for using plurals correctly when he was asked to describe pictures. Imitative prompts (giving examples of proper kinds of phrase) and disapproval for errors, were also incorporated in the training programme and the same programme was continued to promote longer descriptions in the form of compound sentences. Similar teaching methods were effective with retarded children and normal toddlers who had not developed their speech to the target level when Lutzker and Shermna (1974) studied the learning of certain kinds of subject-verb agreement with these kinds of pupil. Rewards produced a very rapid improvement in the ability of the toddlers, as well as the retarded children, for giving descriptions of the type 'boats are sailing', instead of, for instance, 'boats is sailing'. The generative aspect of this response class lies in the facility for giving the correct kind of phrase to *new* pictures after the reward training. In other words, having been rewarded for saying 'boats are sailing' correctly puts the child on the right lines when it comes to a new description like 'girls are riding', even if no reinforcement has actually been given for phrases about girls or about riding. Often, however, a wide range of examples has to be used before the child being taught 'catches on' to the response class.

For adults as well as children positive reinforcement, especially in the form of social approval, biases people towards adopting or discarding complex styles of speech or attitudes of mind. These cannot easily be described or measured, but

60

they can count as *response classes* as long as we can tell which kind of behaviour belongs in each response category. Many of these classes of response, like 'taking care' or 'keeping a stiff upper lip' may be learned on the basis of limited experience, to be brought into action later in totally different situations. At least that is the assumption behind the belief that characterbuilding 'on the playing fields of Eton', or in other forms of schooling, has a lasting effect. A

somewhat similar attempt to cultivate broad classes of response takes place in a form of therapy known as *assertion training*. In this, shy or nervous people become more confident after they have been prompted and encouraged to practise acting assertively, initially perhaps by playing make-believe roles and 'acting' in the theatrical sense in improvised scenes. Obviously if a shy person suddenly becomes very assertive we might talk of a 'personality change' but 'altered response-classes' more accurately expresses what is observed.

If training in one situation helps someone to cope with many similar problems it is often said that *transfer of training* has taken place (Ch. 8), or just that there is *response generalization*. This can apply even when the 'generative response' notion is unnecessary. If someone has been speaking loudly in a noisy factory and carries on talking loudly at home, we might refer to response generalization, and this would also be so for someone who learned to speak up in assertion training with plenty of carry-over to everyday circumstances. It is often hard to distinguish between response and stimulus generalization, and so it is common for all effects of training or therapy to be lumped together as *generalization*. It is also difficult to distinguish between the generalization of fairly peripheral response effects, like walking faster or speaking louder, and generative responses which involve a greater degree of abstraction. *Semantic generalization* is an in-between stage where, for instance, subjects who have been reinforced for saying 'beach' might occasionally say 'sea-shore' or 'sand' instead; the meaning or other associations of a word may be remembered rather than the word itself. On most real-life occasions all the different kinds of generalization may blend together, so the distinctions between them are not always important.

One very well-known study of positive reinforcement in 'client-centred' therapy made use of response categories which depended on the combined judgement of several clinical psychologists.

61

Therapy sessions consisted of conversations between the therapist and a 'client'. Clinical judgement was the basis for analysing tape-recordings by classifying the therapist's statements according to whether they showed approval or disapproval of the patient. In the same way the patient's statements were sorted into nine different categories such as 'problem orientation', 'anxiety', 'negative feelings' and, most important, 'similarity to therapist' (Truax, 1966). The pattern of when the therapist gave sympathy and agreement correlated with the way the patient gradually changed during the sessions of therapy. The therapist gave most approval when the patient talked about himself sensibly, especially if he used a verbal style like the therapist's own. The end result was that the patient talked sensibly about himself more often, and acquired some of the therapist's ways of speaking. Truax's conclusion was therefore that positive reinforcement by the therapist caused helpful changes in the patient, measured as very broad classes of response. There has been a long debate between the therapist in Truax's investigation, Carl Rogers, and B. F. Skinner about whether it is accurate to describe client-centred therapy in terms of positive reinforcement. Both Rogers and his clients believe that the atmosphere of warmth and

'positive regard' (Rogers, 1955) supplied by the therapist allows the patient to improve without value judgements or specific instructions. Whatever reinforcement there is must therefore be informal and intuitive. But the degree to which the reinforcement concept can be adapted to deal with Rogerian therapy (Meyer and Chesser, 1970) illustrates both the flexibility of the concept and the way in which complex kinds of thought and feeling can be interpreted as response classes.

Conclusion and summary

The cornerstone of operant conditioning is the proposition that behaviour is strengthened by contingent rewards. This is positive reinforcement, which is most visible in the feats performed by laboratory animals rewarded with food. Gradual shaping of new skills or categories of response can be useful in many contexts, especially when combined with additional means of directing behaviour such as prompting or instruction. The motivating powers of reinforcement can supply incentive for many items of conduct and influence a wide range of decisions and choices.

62