PSYCHOLOGY AND LIFE

Richard J. Gerrig
Philip G. Zimbardo
Frode Svartdal
Tim Brennen
Roger Donaldson
Trevor Archer

ALWAYS LEARNING

PEARSON

PSYCHOLOGY AND LIFE

Note the difference in experimental procedure between Pavlov's classical conditioning situation and Thorndike's situation. In Pavlov's situation, two stimuli – the CS and the UCS – are presented together, with no requirement that the learner should do (or not do) anything – the two stimuli are presented independently of behaviour. In contrast, Thorndike's procedure requires a specific response for the desired stimulus (access to food) to occur. Here, the rewarding stimulus is presented dependent on the correct response. Also note Thorndike's interpretation of *how* consequences act in strengthening or weakening behaviour: they act through the feelings of pleasure and satisfaction, and displeasure and dissatisfaction. As will become clear in the next section, such an interpretation is mentalistic and very different from the behaviouristic account offered by Skinner.

Experimental analysis of behaviour

B. F. Skinner embraced Thorndike's view that environmental consequences exert a powerful effect on behaviour. Skinner outlined a programme of research whose purpose was to discover, by systematic variation of consequences of behaviour, the ways that environmental conditions affect the likelihood that a given response will occur:

> A natural datum in a science of behavior is the probability that a given bit of behavior will occur at a given time. An experimental analysis deals with that probability in terms of frequency or rate of responding . . . The task of an experimental analysis is to discover all the variables of which probability of response is a function. (Skinner, 1966, pp. 213–14)

Skinner's analysis was experimental rather than theoretical – theorists are guided by derivations and predictions about behaviour from their theories, but empiricists, such as Skinner, advocate the 'bottom-up' approach. This means they start with the collection and evaluation of data within the context of an experiment and are not theory driven.

To analyse behaviour experimentally, Skinner developed **operant conditioning** procedures, in which he manipulated the *consequences* of an organism's behaviour in order to see what effect they had on subsequent behaviour. An **operant** is any behaviour that is *emitted* by an organism and can be characterized in terms of the observable effects it has on the environment. Literally, *operant* means *affecting the environment*, or operating on it (Skinner, 1938). (Skinner preferred not to use the term 'instrumental' about such behaviour because it implies intention.) Operants are *not elicited* by specific stimuli as classically conditioned behaviours are. Pigeons peck, rats search for food, babies cry and coo, some people gesture while talking, and others stutter. The probability of these behaviours occurring in the future can be increased or decreased by manipulating the consequences they have. If, for example, a baby's coo prompts desirable parental contact, the baby will coo more in the future. Operant conditioning, then, modifies the probability of different types of operant behaviour as a function of the consequences they produce.

Operant conditioning Learning in which the probability of a response is changed by a change in its consequences.

Operant Behaviour emitted by an organism that can be characterized in terms of the observable effects it has on the environment.

What environmental contingencies might cause babies to smile more often? *Source:* Eyewire.

To carry out his new experimental analysis, Skinner invented an apparatus that allowed him to manipulate the consequences of behaviour: the *operant chamber*. Figure 6.10 shows how the operant chamber works. When, after having produced an appropriate behaviour defined by the experimenter, a rat presses a lever, the mechanism delivers a food pellet. This device allows experimenters to study the variables that allow rats to learn – or not to learn – the behaviours they define. For example, if a lever press produces a food pellet only if the rat presses with the right paw, the rat will swiftly learn to press in this way.

In many operant experiments, the measure of interest is how much of a particular behaviour an animal carries out in a

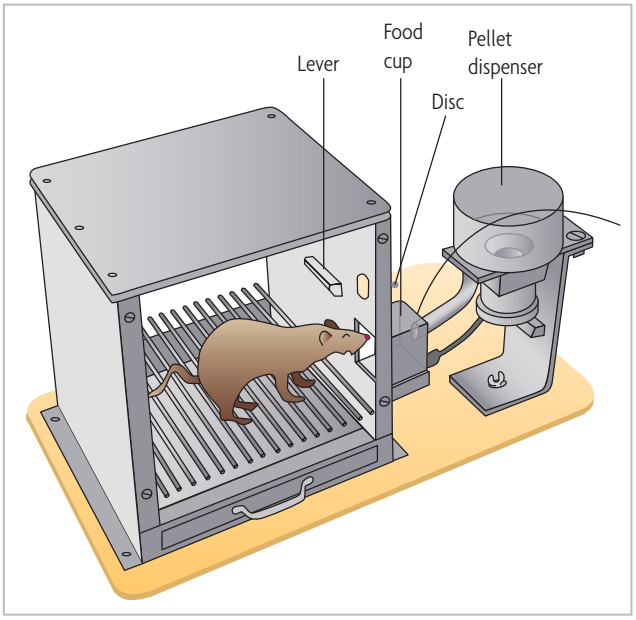

Figure 6.10 Operant chamber. In this specially designed apparatus, typical of those used with rats, a press on the lever may be followed by delivery of a food pellet.

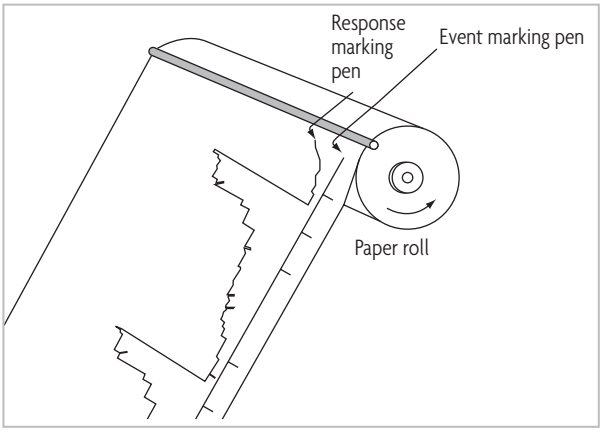

Figure 6.11 The cumulative recorder. The paper moves slowly to the left. Each response is marked by the pen moving up; each reinforcer is marked by a downward dash. When the response marker reaches its top, it is reset to the bottom. Note that a steep curve = many responses per time unit, while a shallow curve = few responses per time unit.

period of time. Researchers therefore record the frequency of the particular behaviour over time in so-called *cumulative* curves (Figure 6.11). These simply display the total number of responses as a function of time. The steeper the curve, the more responses are emitted per time unit. This methodology allowed Skinner to study the effect of reinforcement contingencies on animals' behaviour in a direct and easily visible way.

Reinforcement contingencies

A **reinforcement contingency** is a consistent relation between a response and the changes in the environment that it produces. Imagine, for example, an experiment in which a pigeon's pecking a disc (the response) is generally followed by the presentation of grain (the corresponding change in the environment). This consistent relation, or reinforcement contingency, will usually be accompanied by an increase in the rate of pecking. For delivery of grain to increase *only* the probability of pecking, it must be contingent *only* on the pecking response – the delivery must occur regularly after that response but not after other responses, such as turning or bowing. Based on Skinner's work, modern behaviour analysts seek to understand behaviour in terms of reinforcement contingencies. Let's take a closer look at what has been discovered about these contingencies.

Positive and negative reinforcers

Suppose you wanted your pet rat to press the lever only with the right paw. To increase the probability of this behaviour, you would want to use a **reinforcer**, a stimulus that – when made contingent on a behaviour – increases the probability of that behaviour over time. In this context, 'reinforcement' would refer to the procedure applied by you as an experimenter (delivery of a reinforcer following a required response) *and* the fact that this procedure led to an increased frequency of the response (a learning process).

Reinforcers are defined empirically or functionally, in terms of their effects on changing the probability of a response. If you think about the world, you can probably find three classes of stimuli: those toward which you are neutral, those that you find *appetitive* (you have an 'appetite' for them), and those that you find *aversive* (you seek to avoid them). The compositions of these classes of stimuli clearly are not the same for all individuals: what is appetitive or aversive is defined by the behaviour of the individual organism. Food is an obvious area of individual difference in this regard: for example, although many people find strawberries quite delicious, others find strawberries virtually inedible. If you intend to use strawberries to change behaviour, it is important to consider if strawberries will in fact work as reinforcers.

Reinforcement contingency A consistent relationship between a response and the changes in the environment that it produces.

Reinforcer Any stimulus that, when made contingent upon a response, increases the probability of that response.

Because reinforcers are defined functionally, any behavioural consequence that increases the frequency of that behaviour is termed a *positive reinforcer.* Often, positive reinforcers will be appetitive stimuli, but the important criterion for a positive reinforcer is that it actually increases behaviour. This means that neutral or even aversive events could be considered as positive reinforcers if such events maintain or increase the frequency of behaviours they follow. Note that 'positive' here simply means that such consequences are presented contingent on a response. Also note that you need to distinguish between the *reinforcer* (a stimulus) and *reinforcement* (the fact that behaviour increases in frequency as a function of its consequences). For example, telling a joke may be more likely if the consequence of joke telling is pleasurable laughter. Here, laughter is the reinforcer; actual increased joke telling as a function of laughter is reinforcement.

When a behaviour is reinforced by the removal of an aversive stimulus, the event is called a *negative reinforcer*, and the process *negative reinforcement*. If your headache is relieved by drinking a glass of water, drinking water is the operant behaviour, and escape from the headache is the reinforcing stimulus. There are two general types of learning circumstances in which negative reinforcement applies. In *escape conditioning*, animals learn that a response will allow them to escape from an aversive stimulus, as in the example of the headache. Raising an umbrella during a downpour is another example of escape behaviour. You learn to use an umbrella to escape the aversive stimulus of getting wet. In *avoidance conditioning*, animals learn responses that allow them to avoid aversive stimuli before they even begin. Suppose your car has an irritating buzzer that sounds when you fail to buckle your seat belt. You will soon learn to buckle up to avoid the aversive noise.

To distinguish clearly between **positive** and **negative reinforcement,** remember the following: both positive reinforcement and negative reinforcement *increase* the probability of the response that precedes them. Positive reinforcement increases response probability by the presentation of an appetitive stimulus following a response; negative reinforcement does the same in reverse, through the removal, reduction or prevention of an aversive stimulus following a response.

Recall that for classical conditioning, when the unconditioned stimulus is no longer delivered, the conditioned response suffers extinction. The same rule holds for operant conditioning – if reinforcement is withheld, **operant extinction** occurs. Thus, if a behaviour no longer produces predictable consequences, it returns to the level it was at before operant conditioning – it is extinguished. You can probably catch your own behaviours being reinforced and then *extinguished.* Have you ever had the experience of dropping a few coins into a drinks machine and getting nothing in return? If on one occasion you kicked the machine and your drink came out, the act of kicking would be reinforced. However, if the next few times your kicking produced no results, kicking would quickly be extinguished.

As with classical conditioning, *spontaneous recovery* is also a feature of operant conditioning. Suppose you had reinforced a pigeon's behaviour by providing food pellets when it pecked a key in the presence of a green light. If you discontinued the reinforcement, the pecking behaviour would extinguish. However, the next time you put the pigeon back in the apparatus with the green light on, the pigeon would likely spontaneously peck again. This is an example of spontaneous recovery. In human terms, you might try kicking the drinks machine again with a time lag after your initial extinction experiences.

Positive and negative punishment

You are probably familiar with another technique for decreasing the probability of a response – punishment. A **punisher** is any stimulus that – when it is made contingent on a response – decreases the probability of that response over time. Just as we defined positive and negative reinforcement functionally, positive and negative punishment are also defined functionally. When a behaviour is reduced in frequency as a consequence of the delivery of a stimulus, **positive punishment** occurs (you can remember this meaning of *positive* because something is added to the situation). Touching a hot stove, for example, produces pain that punishes the response, making you less likely to touch the stove next time. When a behaviour is followed by the removal of a stimulus, the process is referred to as **negative punishment** (you can remember this definition of *negative* because something is subtracted from the situation). Thus when a parent withdraws a child's allowance after she hits her baby brother, the child learns not to hit her brother in the future.

Although punishment and reinforcement are closely related operations, they differ in important ways. A good way to differentiate them is to think of each in terms of its effects on behaviour. Punishment, by definition, *reduces* the probability of a response occurring again; reinforcement, by definition, *increases* the probability of a response recurring. For example, some people get severe headaches after drinking caffeinated beverages. The headache is the stimulus that positively punishes and reduces the behaviour of drinking coffee. However, once the headache is present, people often take aspirin or another pain reliever to elimi-

Negative reinforcement A behaviour is followed by the removal of an aversive stimulus, increasing the probability of that behaviour.

Positive reinforcement A behaviour is followed by the presentation of an appetitive stimulus, increasing the probability of that behaviour.

Operant extinction When a behaviour no longer produces predictable consequences, its return to the level of occurrence it had before operant conditioning.

Punisher Any stimulus that, when made contingent upon a response, decreases the probability of that response.

Positive punishment A behaviour is followed by the presentation of an aversive stimulus, decreasing the probability of that behaviour.

Negative punishment A behaviour is followed by the removal of an appetitive stimulus, decreasing the probability of that behaviour.

Table 6.1 The three-term contingency: relations among discriminative stimuli, behaviour, and consequences

	Discriminative stimulus	Emitted response	Stimulus consequence
1. **Positive reinforcement:** A response in the presence of an effective signal produces the desired consequence. This response Increases.	Soft-drink machine	Put coin in slot	Get drink
2. **Negative reinforcement (escape):** An aversive situation is escaped from by an operant response. This escape response increases.	Heat	Fan oneself	Escape from heat
3. **Positive punishment:** A response is followed by an aversive stimulus. The response is eliminated or suppressed.	Attractive matchbox	Play with matches	Get bumed or get caught
4. **Negative punishment:** A response is followed by the removal of an appetitive stimulus. The response is eliminated or suppressed.	Brussels sprouts	Refusal to eat them	No dessert

nate the headache. The aspirin's analgesic effect is the stimulus that negatively reinforces the behaviour of ingesting aspirin.

Discriminative stimuli and generalization

You are unlikely to want to change the probability of a certain behaviour at all times. Rather, you may want to change the probability of the behaviour in a particular context. For example, teachers often want to increase the probability that a child will sit quietly in class, without changing the probability that he or she will be noisy and active during recess. Through their associations with reinforcement or punishment, certain stimuli that precede a particular response – **discriminative stimuli** – come to set the context for that behaviour. Organisms learn that, in the presence of some stimuli but not of others, their behaviour is likely to have a particular effect on the environment. For example, in the presence of a green traffic light, the act of crossing a junction in a motor vehicle is reinforced. When the light is red, however, such behaviour may be punished – it may result in a traffic ticket or an accident. Skinner referred to the sequence of discriminative stimulus–behaviour–consequence as the **three-term contingency** and believed that it could be applied to the analysis of most human behaviour (Skinner, 1953). Table 6.1 describes how the three-term contingency might explain several different kinds of human behaviour.

Under laboratory conditions, manipulating the consequences of behaviour in the presence of discriminative stimuli can exert powerful control over that behaviour. For example, a pigeon might be given grain after pecking a disc in the presence of a green light but not a red light. The green light is a discriminative stimulus that sets the occasion for pecking; the red is a discriminative stimulus that sets the occasion for not pecking. Organisms learn quickly to discriminate between these conditions, responding regularly in the presence of one stimulus and not responding in the presence of the other. By manipulating the components of the three-term contingency, you can constrain a behaviour to a particular context.

Organisms also generalize responses to other stimuli that resemble the discriminative stimulus. Once a response has been reinforced in the presence of one discriminative stimulus, a similar stimulus can become a discriminative stimulus for that same response. For example, pigeons trained to peck a disc in the presence of a green light will also peck the disk in the presence of lights that are lighter or darker shades of green than the original discriminative stimulus.

Discriminative stimuli Stimuli that act as predictors of reinforcement, signalling when particular behaviours will result in positive reinforcement.

Three-term contingency The means by which organisms learn that, in the presence of some stimuli but not others, their behaviour is likely to have a particular effect on the environment.

Using reinforcement contingencies

Are you ready to put your new knowledge of reinforcement contingencies to work? Here are some considerations you might have:

- *How can you define the behaviour that you would like to reinforce or eliminate?* You must always carefully target the specific behaviour whose probability you would like to change. Reinforcement should be contingent on exactly that behaviour. When reinforcers are presented non-contingently, their presence has little effect on behaviour. For example, if a parent praises bad work as well as good efforts, a child will not learn to work harder in school – but because of the positive reinforcement, other behaviours are likely to increase. (What might those be?)
- *How can you define the contexts in which a behaviour is appropriate or inappropriate?* Remember that you rarely want to allow or disallow every instance of a behaviour. We suggested earlier, for example, that a teacher might want to

increase the probability that a child will sit quietly in class without changing the probability that he or she will be noisy and active during recess. You must define the discriminative stimuli and investigate how broadly the desired response will be generalized to similar stimuli. If, for example, the child learned to sit quietly in class, would that behaviour generalize to other 'serious' settings?

- *Have you unknowingly been reinforcing some behaviours?* Suppose you want to reduce the frequency of a behaviour. Before you turn to punishment as a way of reducing its probability, try to determine whether you can identify reinforcers for that behaviour. If so, you can try to extinguish the behaviour by eliminating those reinforcers. Imagine, for example, that a young boy throws a large number of tantrums. You might ask yourself, 'Have I been reinforcing those tantrums by paying the boy extra attention when he screams?' If so, you can try to eliminate the tantrums by eliminating the reinforcement. Even better, you can combine extinction with positive reinforcement of more socially approved behaviours.

It's important to be aware that the reinforcers parents produce can make children's conduct problems, such as tantrums, more likely. In fact, parenting research has identified unknowing reinforcement as one cause of serious behaviour problems in children. For example, Gerald Patterson and his colleagues (G. R. Patterson, 2002; Reid et al., 2002) have outlined a *coercion model* for antisocial behaviour. Family observations suggest that children are put at risk when their parents issue threats in response to small misbehaviours (such as whining, teasing or yelling) without following through on those threats. At some moments, however, these parents would issue harsh or explosive discipline toward the same behaviours. The children appear to learn the lesson that relatively large acts of aggressive and coercive behaviour are appropriate and necessary for achieving goals – leading to a cycle of increasing severity of the children's antisocial behaviour.

Behaviour analysts assume that any behaviour that persists does so because it results in reinforcement. Any behaviour, they argue – even irrational or bizarre behaviour – can be understood by discovering what the reinforcement or payoff is. For example, symptoms of mental or physical disorders are sometimes maintained because the person gets attention and sympathy and is excused from normal responsibilities. These *secondary gains* reinforce irrational and sometimes self-destructive behaviour. Can you see how shy behaviours can be maintained through reinforcement, even though the shy person would prefer not to be shy? Of course, it is not always possible to know what reinforcers are at work in an environment. However, as a behaviour becomes more or less probable, you might try to carry out a bit of behaviour analysis.

One final thought. It's often the case that real-life situations will involve intricate combinations of reinforcement and punishment. Suppose, for example, parents use negative punishment by grounding a teenager for two weeks when he stays out past curfew. To soften up his parents, the teen helps more than usual around the house. Assuming his helping behaviour appeals to the parents, the teen is trying to reinforce his parents' 'reducing the sentence' behaviour. If this strategy succeeds in changing the punishment to only one week, the teen's helping behaviour will have been negatively reinforced – because helping led to the removal of the aversive stimulus of being grounded. Whenever the teen is grounded again (a discriminative stimulus), his helping behaviour should be more likely. Do you see how all the contingencies fit together to change both the teen's and the parents' behaviours?

Let's now take a look at the ways in which various objects and activities may come to function as reinforcers.

Properties of reinforcers

Reinforcers are the power brokers of operant conditioning: they change or maintain behaviour. Reinforcers have a number of interesting and complex properties. They can be learned through experience rather than be biologically determined, and can be activities rather than objects.

Conditioned reinforcers

When you were born, there were a handful of **primary reinforcers,** such as food and water, whose reinforcing properties were biologically determined. Over time, however, otherwise neutral stimuli have become associated with primary reinforcers and now function as **conditioned reinforcers** for operant responses. Conditioned reinforcers can come to serve as ends in themselves. In fact, a great deal of human behaviour is influenced less by biologically significant primary reinforcers than by a wide variety of conditioned reinforcers. Money grades, and various kinds of status symbols are among the many potent conditioned reinforcers that influence much of your behaviour.

Virtually any stimulus can become a conditioned reinforcer by being paired with a primary reinforcer. In one experiment, simple tokens were used with animal learners.

> With raisins as primary reinforcers, chimps were trained to solve problems. Then tokens were delivered along with the raisins. When only the tokens were presented, the chimps continued working for their 'money' because they could later deposit the hard-earned tokens in a 'chimp-o-mat' designed to exchange tokens for the raisins (Cowles, 1937).

Teachers and experimenters often find conditioned reinforcers more effective and easier to use than primary reinforcers because: (1) few primary reinforcers are available in the classroom,

Primary reinforcers Biologically determined reinforcers such as food and water.

Conditioned reinforcers In classical conditioning, formerly neutral stimuli that have become reinforcers.

whereas almost any stimulus event that is under control of a teacher can be used as a conditioned reinforcer; (2) they can be dispensed rapidly; (3) they are portable; and (4) their reinforcing effect may be more immediate because it depends only on the perception of receiving them and not on biological processing, as in the case of primary reinforcers.

In some institutions, such as psychiatric hospitals or drug treatment programmes, *token economies* are set up based on these principles. Desired behaviours (grooming or taking medication, for example) are explicitly defined, and token payoffs are given by the staff when the behaviours are performed. These tokens can later be exchanged by the patients for a wide array of rewards and privileges (Kazdin, 1994; Martin & Pear, 1999). These systems of reinforcement are especially effective in modifying patients' behaviours regarding self-care, upkeep of their environment, and, most important, frequency of their positive social interactions.

Response deprivation and positive reinforcers

Suppose you need to get a child to do something. You don't want to pay her or give her a gold star, so instead you strike this bargain: 'When you finish your homework, you can play with your video game.' Why might this tactic work? According to *response deprivation theory*, behaviours become preferred and, therefore, reinforcing when an animal is prevented from engaging in them (Timberlake & Allison, 1974). For example, water-deprived rats learned to increase their running in an exercise wheel when their running was followed by an opportunity to drink. Conversely, exercise-deprived rats learned to increase their drinking when that response was followed by a chance to run (Premack, 1965). Can you see how the promise of video games after homework follows this same pattern? For a period of time, the child is video game-deprived – the rate at which the child would ordinarily play the video game is restricted below normal. To overcome that deprivation, she will learn to work on her homework.

This analysis suggests two important lessons. First, these examples remind you why you shouldn't assume that the same activity will function as a reinforcer for an animal at all times. You need to know, for example, whether the animal is food-deprived before you attempt to use food as a reinforcer. Second, these examples suggest why virtually any activity can come to serve as a reinforcer. You can experience deprivation along any number of dimensions. In fact, if you didn't allow a child to do homework for a period of time, she would learn other behaviours to overcome homework deprivation.

Critical thinking in your life

Unexpected effects of rewards

Reinforcers are, by definition, events that strengthen the behaviour they follow. If a teacher praises a student for good performance in solving maths problems, praise will be a reinforcer if the student continues his or her good work. In this case, there is a correspondence between the technical definition of reinforcement, and the more colloquial term 'reward'.

However, sometimes rewards do not work as reinforcers. An example is the so-called *overjustification effect*. One of the first demonstrations of this effect (Lepper et al., 1973) compared 3- to 5-year-old children randomly allocated to three experimental conditions. In one condition children were told that they would receive a 'good player' ribbon for drawing with pens. In a second condition, children played without any expectation about a reward; they just got the reward after having played. A third group did not receive the reward and had no expectation to receive one. All of the children played with the pens, which is an enjoyable activity for preschoolers. Later, when observed in a free-play setting, the children who had received a reward that had been promised to them played significantly less with the felt-tipped pens compared to the two other groups. Thus, instead of strengthening the playing behaviour, the expected reward had an opposite effect.

This finding seems to contradict the reinforcement principle. Remember, however, that reinforcement is defined functionally, so technically Lepper et al. just described a situation in which a consequence of behaviour reduces its probability (i.e. punishment). Nevertheless, it is of great theoretical and practical interest that some 'rewards' actually work contrary to their intended effects. Why is this so?

One important factor seems to be the initial probability of the target behaviour. Note that playing with pens is enjoyable for children – the activity is rewarding in itself, it is intrinsically motivated. If an external reward is added to such an activity, it can actually reduce the motivation for that activity. 'Overjustification' in this context implies that something inherently

Intrinsically motivated behaviour Behaviour that is motivated by interest or enjoyment in performing the behaviour itself.