David Hardman

Judgment and Decision Making

Psychological Perspectives

BPS TEXTBOOKS IN PSYCHOLOGY
The British Psychological Society
BPS Blackwell

12 Risk

KEY TERMS

affect heuristic
appraisal-tendency theory
betrayal aversion
catastrophic potential
'do no harm' heuristic
dread risk
expressed preferences
five-factor model of personality
kin selection theory
naturalism heuristic
outrage heuristic
precautionary principle
pseudo-certainty effect
psychometric approach to risk
revealed preferences
risk compensation
risk-as-feelings model
sensation-seeking
sex differences
social amplification of risk
signal value
testosterone
unknown risk
voluntary vs. involuntary risk

CHAPTER OUTLINE

INTRODUCTION

How do you feel about nuclear power? X-rays? Do you smoke, drive, take illegal drugs? Have you been mountain climbing, or would you like to? The study of risk can be quite perplexing. People are often quite willing to accept certain types of risks, but get quite angry about other risks that actually pose less of a threat. It is also a curious paradox of progress, at least in the developed world, that new technologies often make us feel more at risk, while at the same time providing us with lifestyles and lifespans that previous generations could never have imagined.

In this chapter, I shall examine the factors that influence our judgments of risk. As we will see, early cost-benefit analyses of what is regarded as an acceptable level of risk quickly gave way to psychometric analyses. This latter technique has found that people have a range of interrelated intuitions that can be grouped into two dimensions. We shall also look at the influences on risk judgments from emotional factors, social factors, personality, sex, race, and expertise.

THE REVEALED PREFERENCES APPROACH TO RISK

In a 1969 *Science* article Chauncey Starr posed the question, 'What is our society willing to pay for safety?' To answer this question he applied cost-benefit analysis to various technologies and behaviours. As an example of the kind of figures he used (which involved various assumptions), here are the types of data used for assessing motor-vehicle travel:

> The calculation of motor-vehicle fatalities per exposure hour per year is based on the number of registered cars, an assumed 1½ persons per car, and an assumed 400 hours per year of average car use . . . The figure for annual benefit for motor-vehicle travel is based on the sum of costs for gasoline, maintenance, insurance, and car payments and on the value of the time savings per person. It is assumed that use of an automobile allows a person to save 1 hour per working day and that a person's time is worth $5 per hour. (pp.1237–1238)

Starr plotted the level of risks to benefits for various technologies/activities, where risk was defined as the probability of fatalities per person-hour of exposure, and benefit per person was measured on a dollar scale. The ratio of risks to benefits differed for various activities. On the assumption that these figures were a reasonably accurate reflection of people's tolerance for various risks, people appeared willing to accept voluntary risks (for example, smoking, hunting) that were about 1000 times greater than involuntary risks (for example, electric power). Starr also suggested that benefit awareness, as estimated from the level of advertising, increased the levels of risk that the public were willing to accept.

Using his methodology, Starr thought currently accepted levels of risk could be used to predict the risks that would be accepted from other sources (for example, acceptance levels for coal-burning power plants could be used to predict the acceptable level of risks from nuclear power stations).

THE PSYCHOMETRIC APPROACH TO RISK

The two dimensions of risk

Starr's methodology was based on the notion that people's preferences are *revealed* in their behaviour. This idea was questioned in Chapter 8. In particular, it is also questionable whether people's behaviours are based on a rational assessment of costs and benefits. Other authors have examined people's *expressed preferences* in relation to risks and benefits. Fischhoff *et al.* (1978) asked members of the Oregon League of Women's Voters and their spouses (a thoughtful and influential group of private citizens) to rate various technologies/activities for risk, benefit, and maximum level of acceptable risk. They also asked them to rate each item on various risk characteristics: voluntariness of risk, immediacy of consequences, whether the persons exposed have a knowledge of the risks, whether the risks are known to science, the extent of control that a person has over the risk, newness of the risks, whether the risk kills people one at a time or in large numbers, whether the risk is one that arouses dread, and the likelihood that the consequences will be fatal.

There were high intercorrelations between the various risk factors. Using the technique of factor analysis, Fischhoff *et al.* found that two basic dimensions appeared to underlie the perceived relationships between these risk factors. One of these factors can be termed *dread risk* and the other *unknown risk*. Dread risk is associated with the feeling of dread, and with a number of other factors such as catastrophic consequences and fatal consequences. Unknown risk is associated with factors such as being novel, unobservable, and having delayed consequences. A similar study found that the single factor 'number of people exposed' emerged as a third dimension (Slovic *et al.*, 1980). Nonetheless, two underlying dimensions seem to capture most of the variation in responding, a finding that has also been replicated in other countries (e.g. Brun, 1992; Teigen *et al.*, 1988; Vlek & Stallen, 1981). Figure 12.1 shows how these two dimensions map onto a number of risks.

Factors underlying the two dimensions

Voluntary versus involuntary risks Starr's 1969 paper hypothesised that voluntariness mediated perceived risk–benefit relationships. To test this idea, Fischhoff *et al.* (1978) dichotomised their items into the 15 most voluntary and 15 least voluntary risks. Contrary to Starr's hypothesis, the relationship between perceived risks and perceived benefits did not differ for voluntary and involuntary risks. However, in terms of what was regarded as an acceptable level of risk, people were willing to tolerate a higher level for risks that they saw as voluntary. Nonetheless, the extent to which a risk is seen as involuntary also correlates highly with other factors, such as lack of control, global catastrophe, inequity and catastrophe (Slovic *et al.*, 1980). This raises the possibility that it is these other factors that impact on people's judgments, and not involuntariness per se. Indeed, when these factors were removed statistically, voluntariness no longer related to acceptable risk.

Catastrophic potential Slovic *et al.* noted that respondents seemed to show little concern about the risks from diagnostic X-rays, despite the fact that the invisible and irreversible contamination that they produce can lead to cancer and genetic damage. By contrast, people believed that nuclear power posed a greater risk of death than any of the other hazards that they were asked to consider. This perception was linked to the perceived potential for disaster. Presumably, it is – at least in part – the concern about

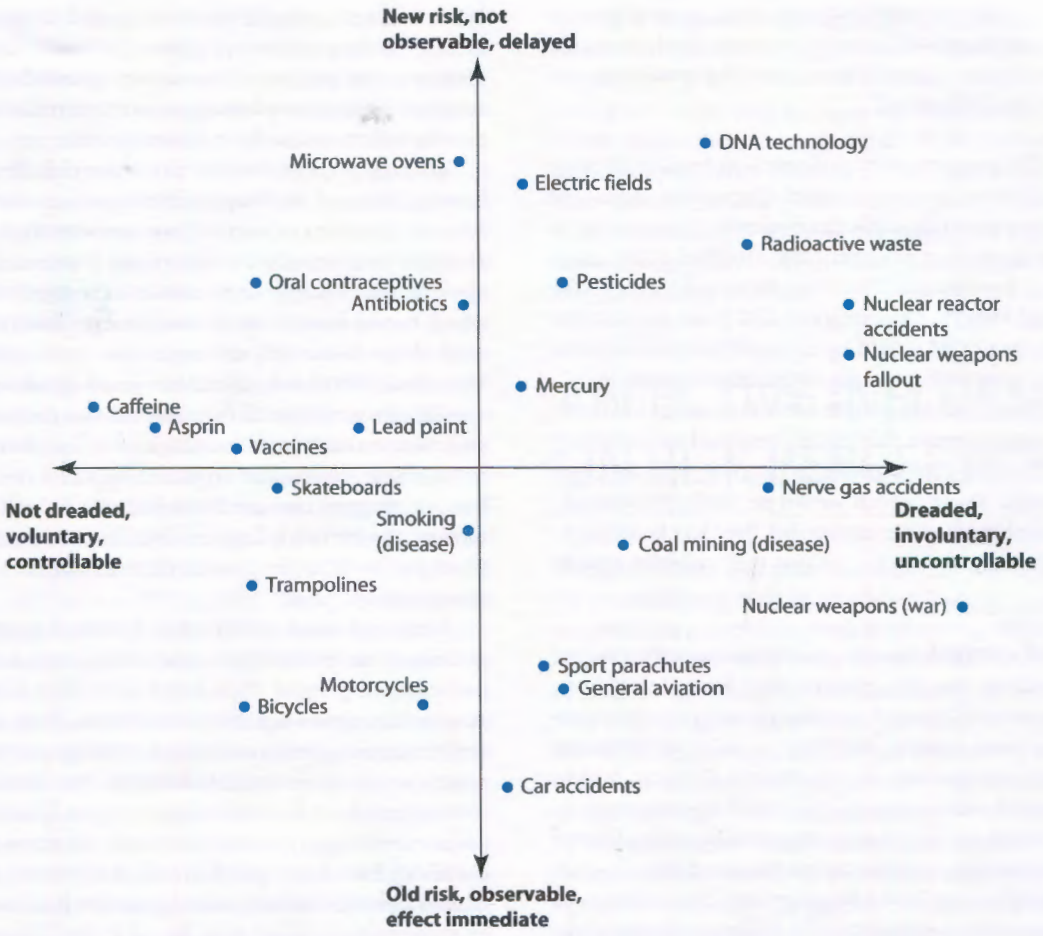


Figure 12.1. *The two-dimensional factor structure*
This is a simplified graphic that does not show all of the risks that have been reported in such studies.

catastrophic potential that lies behind public opposition to nuclear technology.

Slovic *et al.* also pointed out that it is very difficult to demonstrate the improbability of catastrophic accidents because doing so requires vast amounts of data. What little evidence there is may be interpreted within the framework of a person's prior beliefs. Furthermore, certain types of accident, when they occur, may have *signal value* – that is, they may be interpreted in terms of an increased likelihood of similar events in the future and, thus, the need for corrective action. This hypothesis was supported by a questionnaire study that showed certain types of events to be regarded as more informative than others. Consider the following two scenarios. In one, a dam collapse kills 40 people. In the other, a partial core meltdown at a nuclear reactor kills one person inside the plant, but the radiation is contained and does not reach the external environment. The judged levels of grief and suffering were about the same for both incidents (4.9 vs. 4.5 on a seven-point scale). However, the nuclear accident was seen as more informative and more worrying, and people thought that there was a greater need for awareness and a greater effort needed to prevent recurrence.

Known versus unknown risks People tend to be more concerned about unknown risks than known risks. When risks are unknown it is likely that people fear the worst. When people first became aware of HIV/AIDS in the 1980s, almost nothing was known about how the disease was spread, about how many people had the illness, or what the chances were of acquiring it. It is probably fair to say that public concern in Western countries was far greater in those early days than it is now, when we know more about HIV, how to avoid it, and what can be done to assist those who have acquired it.

It is sometimes argued that we should respond to new risks by applying the *precautionary principle*. This is sometimes formulated in different ways, for example:

> In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation. (The United Nations *Rio Declaration on Environment and Development*[1])

> [I]f an action or policy might cause severe or irreversible harm to the public, in the absence of a scientific consensus that harm would not ensue, the burden of proof falls on those who would advocate taking the action. (Wikipedia)[2]

In practice, the precautionary principle – at least in its most stringent formulations – is controversial. Opponents argue that it is a major hindrance to desirable development (or would be, if applied). Thus, supporters of genetically modified (GM) crops point to the huge benefits that this technology could bring to the world's poor and hungry. Opponents of GM crops say that the major beneficiaries of GM would be the multinational corporations that make them, with great risk to the environment.

Finally, one likely factor in people's concerns about unknown risks is the *ignorance aversion* that we encountered in Chapter 7. Recall that not only do people prefer to bet on events that they feel knowledgeable about (which would be perfectly rational), but they prefer to bet on vague events that they feel knowledgeable about rather than chance events that they consider equally probable.

Perceptions of control People are willing to accept a higher level of risk for things that they perceive they have control over. However, as we saw in Chapter 9, people typically perceive themselves as having more control than they actually do (including the perception of control over chance events). This can lead to unrealistic optimism, whereby people believe that good things are more likely to happen to them than to other people and bad things are more likely to happen to other people than to them.

It is likely that there are both advantages and disadvantages to the illusion of control and optimistic bias. Without striving, there can be no achievement, but without the belief that one has some control then striving will not occur. However, when one strives for something a negative outcome is always a possibility. In some domains, negative outcomes may mean physical harm. For instance, a motorist who has an unrealistic perception of his level of control may kill himself (and others) through dangerous driving.

Other intuitions about risk

Naturalism: nature knows best One quite common intuition is that 'natural is good' (e.g. Baron, 1998). Thus, 45 per cent of the general public in Oregon agreed that 'Natural chemicals, as a rule, are not as harmful as man-made chemicals.' Of the same group of respondents, 49 per cent blamed agricultural pesticides for causing malformations in newly born children in a (hypothetical) small community. For a sample of professional toxicologists these percentages were just 13 per cent and 6 per cent, respectively (Kraus *et al.*, 1992).

The naturalism heuristic is perfectly understandable; after all, our very survival is based on the consumption of the water, fruit, vegetables, and meat that our environment provides us. However, the products of the natural world are not always benign. Plants did not evolve because they wanted us to eat them; on the contrary, many plants have their own defence mechanisms against consumption. This is thought to be why people have an aversion to bitter-tasting foods: the aversion is an evolved mechanism to alert us to the possible presence of toxins (Soranzo *et al.*, 2005). However, the same aversion may act against the interests of contemporary humans by making us reluctant to eat foods containing nutrients that combat heart disease and cancer.

We may overestimate the extent to which the natural world is benign. Some of the foodstuffs we consume now are only safe because conscious or unconscious domestication has made them so. Earlier in human history these foods were extremely dangerous. For example, wild almonds contain a chemical called amygdalin, which breaks down to create the lethal poison cyanide. A snack of wild almonds can kill someone who ignores the bitter taste. Diamond (1998) has described how domestication gave rise to the non-lethal sweet almond. It so happens that the occasional almond tree contains a mutation in a single gene that prevents them from synthesising amygdalin. In the wild, these trees die without leaving progeny because birds discover and eat all their seeds. However, when early farmers discovered these trees they would plant the seeds in their orchards, thus beginning the spread of sweet almond trees.

Ames and Gold (1990) have described plants and animals as being in an 'evolutionary war'. Plants have developed natural pesticides to prevent their being eaten, but animals have also developed defences against certain toxins. Thus, people are quite well protected against low doses of carcinogens. This is just as well, since about half the natural chemicals that have been tested on rodents have been found to be carcinogenic. This is about the same proportion as for synthetic chemicals, yet many more synthetic chemicals have been tested overall than natural ones. Of the 27 natural pesticide rodent carcinogens that had been identified by 1990, one or more had been found in the following foodstuffs: anise, apple, banana, basil, broccoli, Brussels sprouts, cabbage, cantaloupe, caraway, carrot, cauliflower, celery, cherry, cinnamon, cloves, cocoa, coffee (brewed), comfrey tea, dill, eggplant (aubergine), endive, fennel, grapefruit juice, grape, honey, honeydew melon, horseradish, kale, lettuce, mace, mango, mushroom, mustard (brown), nutmeg, orange juice, parsley, parsnip, peach, pear, pepper (black), pineapple, plum, potato, radish, raspberry, rosemary, sage, sesame seeds (heated), strawberry, tarragon, thyme, and turnip.

The naturalism heuristic means that people are less willing to pay money to combat harm caused by events perceived as natural, as compared to when the source of harm is human. For instance, 68 per cent of respondents were willing to contribute to save endangered dolphins from pollution, but only 44 per cent were willing to contribute when the danger was described as a new virus. This translated into an average contribution of $18.85 for the pollution danger and $6.35 for the virus (Kahneman & Ritov, 1994; see also Kahneman *et al.*, 1993). Kahneman *et al.* actually proposed a similar, but rather more general, heuristic to explain their results: they suggested that people's willingness to contribute to alleviating harm is determined by an *outrage heuristic*, whereby their feelings are related to the cause of the harm.

Eliminating versus reducing risk Studies of consumers have shown that people are more willing to pay a premium for the elimination of a risk compared to its reduction. This tendency is known

as the *certainty effect* (Samuelson & Zeckhauser, 1988; see also Chapter 7). Such reactions can result in a *pseudo-certainty effect*, as demonstrated by Slovic *et al.* (1982). They found that a vaccine against a disease affecting 20 per cent of the population was considered less attractive when described as effective in half of all cases than when described as providing full protection against one of two exclusive and equally probable (10 per cent in each case) virus strains causing identical symptoms.

The pseudo-certainty effect might be considered rational, in the sense that the removal of one disease means that effort can then be concentrated on ameliorating the threats posed by the alternative disease. By contrast, tackling two distinct diseases might, all other things being equal, pose more of a challenge. However, for certain types of risk it is not clear that the elimination of risk is a desirable aim. In the United States, the 1980 Superfund law concerns the clean-up of hazardous waste that has been left in the ground. The law mandates the complete removal of such waste. However, it has been argued that the greatest expense is incurred during the removal of the last 10 per cent of such waste (Breyer, 1993, cited in Baron, 2000). Thus, the money that is spent on completely removing the final remains of waste from one site could be better spent on removing waste from other areas.

'Do no harm' Ritov and Baron (1992) have suggested that an omission bias occurs with some decisions (see Chapter 8). This is a greater reluctance to risk negative outcomes by one's actions, compared to risking negative outcomes by not taking action. The omission bias can be considered as reflecting a 'do no harm' heuristic (Baron, 1998), a heuristic that may also be behind calls to apply the precautionary principle that we encountered earlier. Baron (1998) has suggested that the reluctance of some people to vaccinate their children may be an instance of omission bias. He gave the example of polio vaccination. The Salk vaccine uses a killed virus whereas the Sabin vaccine uses a live virus (Sabin has actually replaced the Salk vaccine). Neither virus is 100 per cent effective, but whereas Salk occasionally fails to *prevent* polio occurring the Sabin vaccine occasionally *causes* polio to occur. However, Sabin is more effective overall. Nonetheless, Baron noted that 'Nobody ever sued the makers of the Salk vaccine for failing to prevent polio, but many people have sued the makers of the Sabin vaccine for causing polio' (1998, p.112).

The 'do no harm' principle may also lie behind another perplexing finding. Consider an automobile company that uses a cost-benefit analysis, trading off dollars against lives saved, in order to decide whether to implement a safety precaution. On the basis of the analysis, even though a high value was placed on human life, the company decides against the safety precaution. Viscusi (2000) reports that people view such a company more unfavourably than a similar company that imposes the same risk on motorists without conducting a cost-benefit analysis. Sunstein (2005) has suggested that a 'cold heart' heuristic may be operating.

Betrayals Koehler and Gershoff (2003) noted that people are especially averse to risks of death that come from products (such as airbags) designed to promote safety. In one study, people heard that there was a 2 per cent chance that drivers of Car A, with Air Bag A, will be killed in serious accidents. For drivers of Car B, with Air Bag B, there was a 1 per cent chance of death, but an additional chance of one in 10,000 (0.01 per cent) of dying as a result of the air bag deploying. When asked to choose between these two equally priced cars, most people chose Car A – the riskier car. In other words, people seem to be choosing a higher-risk option in order to avoid the possibility of betrayal. Whether or not this example illustrates the use of a moral heuristic gone awry (Sunstein, 2005) as opposed to simply showing that people will incur *some* cost (but not necessarily any cost) to avoid betrayal (Koehler & Gershoff, 2005) is a matter of debate.

AFFECTIVE INFLUENCES ON RISK PERCEPTION

The observation that people's assessments of risk are influenced by an underlying dimension of *dread*, and its associated factors, suggests that perception of risk is influenced by emotion as well as cognition. However, prior to the 1990s there was little discussion of emotion in relation to decision making. It was often assumed that feelings arose as a consequence of cognitive evaluation, but without having a direct input into decisions or behaviour (although the anticipation of emotion was sometimes recognised as a possible input into cognitions).

On the basis of more recent work, there is now reason to believe that the relationship between cognition and emotion is much more interactive. Figure 12.2 shows the *risk-as-feelings* model proposed by Loewenstein *et al.* (2001). This suggests that cognitions and emotions can act upon each other, but also that they can both act directly on behaviour without the mediation of the other (the relationship between cognition, emotion, and decision making will be explored further in Chapter 15).

In Chapter 9 we saw that different emotions have different effects on people's risk judgments and their optimism. Induced anger and happiness produced equal optimism about events that were ambiguous in terms of their certainty and controllability, whereas induced fear led to pessimism. For unambiguous events, only the 'happy' participants were optimistic, whereas the fearful and angry participants were about equally pessimistic (Lerner & Keltner, 2001, Study 3). Lerner and Keltner interpreted their results in terms of *appraisal-tendency theory*. This states that emotions elicit cognitive appraisals that interpret the situation that gave rise to the emotions, and thereby help respond to the situation. However, emotions can persist beyond the situation that elicited them, such that they may affect the interpretation of other situations (Figure 12.3).

Shortly after the attacks on the World Trade Center on 11 September 2001, Jennifer Lerner and colleagues asked people to rate their emotional responses. Then, six to 10 weeks later, the same people were asked to assess the risk to themselves and other Americans from a diverse range of events. People who showed greater anger in the wake of 9/11 tended to give lower ratings of risk. By contrast, people who showed greater fear tended to give higher ratings of risk. These emotional responses also related to

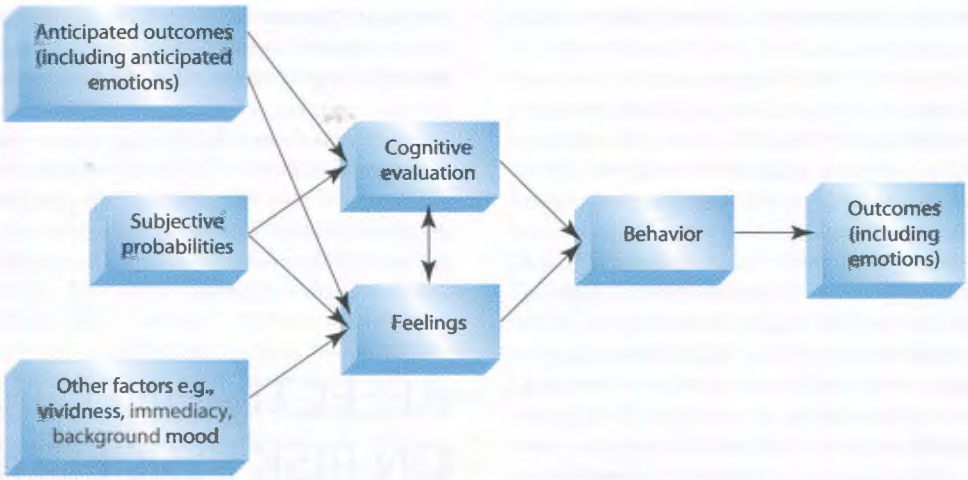
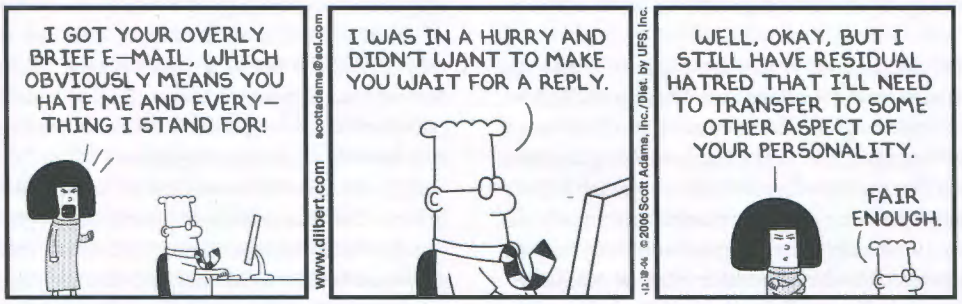


Figure 12.2. *The risk-as-feelings perspective*

Source: Loewenstein *et al.*, 2001.



© Scott Adams, Inc./Dist. by UFS Inc.

Figure 12.3. *Emotions tend to persist beyond the situation that elicited them*

different views about policy: Anger was associated with the wish to deport foreigners who did not possess a valid visa, whereas fear was associated with the desire to build stronger contacts with the Muslim world (Lerner *et al.*, 2003). This study also found that experimentally induced fear led to higher risk estimates than did anger. Women rated the risks as higher than did men, though this did not interact with the experimental manipulations.

Other researchers have proposed that people use an *affect heuristic* to make judgments of risks and benefits. Alhakami and Slovic (1994) found that, if people liked an activity, they tended to rate its risks as low and its benefits as high. On the other hand, if they disliked an activity then they rated its risks as high and its benefits as low. They suggested that people were using their affective feelings to make these judgments. This interpretation of their findings was supported by a further study in which half the participants made risk and benefit judgments under time pressure (Finucane *et al.*, 2000, Study 1). Time pressure reduces the cognitive resources available for analytical deliberation and may increase arousal, making affective processes more salient (Maule & Svenson, 1993). For most items on a list of 23 hazardous facilities or activities, the risk–benefit correlations were more strongly negative for the judgments made under time pressure. In a follow-up study, enhancing the awareness of benefits through the provision of information not only tended to increase the ratings of benefits, but also decreased the ratings of risks. Conversely, enhancing the awareness of risks increased the ratings of risks, but decreased the ratings of benefits (Finucane *et al.*, 2000, Study 2).

THE SOCIAL AMPLIFICATION OF RISK

How are perceptions of risk influenced by social and cultural factors? The psychometric approach to risk has shown that many factors do influence risk perceptions, but only recently have researchers tried to incorporate the wider social and cultural factors into a theory (or, perhaps more accurately, a framework). The

basic idea behind the social amplification of risk framework (SARF) is that:

> risk events interact with psychological, social, and cultural processes in a way that can heighten or attenuate public perception of risk and related risk behavior. Behavioral patterns, in turn, generate secondary social or economic consequences but may act also to increase or decrease the physical risk itself. Secondary effects trigger demands for additional institutional responses and protective actions, or, conversely (in the case of risk attenuation), impede needed protective actions. (Kasperson *et al.*, 1988, p.234)

In short, social amplification involves two stages. The first stage is the transfer of information about the risk or risk event. For risks that are not directly experienced, risk amplification may be influenced by the volume of information, the degree to which information is disputed, the extent of dramatisation, and the symbolic connotations of the information. For example, a risk message about a chemical may be interpreted differently, depending on whether it comes from a pharmaceutical company or a consumer organisation. Both have symbolic value, depending on a person's prior views. A pharmaceutical company message may be regarded as less credible if it arouses the mental image of uncaring big business. The consumer organisation may lead to a more positive image of a body that is on the side of the ordinary person, therefore more attention is paid to the content of the message.

The second stage involves the response mechanisms of society. Response mechanisms may involve the use of heuristics and values, social group relationships, signal value, and stigmatisation. Kasperson *et al.* (1988) refer to risk events having a 'ripple effect'. For example, the Three Mile Island nuclear incident was followed by nuclear plants around the world being shut down and restarted more often for safety checks, despite the fact that these phases of operation are by far the riskiest.

In reviewing 15 years of research into the social amplification of risk, Kasperson *et al.* (2003) noted the importance of *trust* as a response mechanism. Distrust increases the perception of risk, intensifies the public's reaction to risk signals, contributes to the perceived unacceptability of risk, and stimulates political activism (p.32).

PERSONALITY, SEX, RACE, AND EXPERTISE

Personality

Is risk-taking domain-specific or domain-general? One constant finding from several studies is that the intercorrelations between risk taking in different domains tend to be weak. On the face of it, this would appear to indicate that there are not really any risk-seeking or risk-avoiding personalities (e.g. Salminen & Heiskanen, 1997; Weinstein & Martin, 1969). However, a complicating factor here is that people's *perceptions* of risk can vary depending on the characteristics of the situation; in fact, for many people, attitudes to the risks that they *perceive* tend to be quite stable across situations (Weber & Millman, 1997). In other words, when variations in risk perception are controlled for, people who tend to take risks in one domain are more likely to take risks in other domains. In the remainder of this section I shall look at how other aspects of a person's personality may affect their propensity to take risks.

The evolution of personality variations MacDonald (1995) proposed that variations in personality could be viewed as representing alternative strategies for maximising evolutionary fitness. This idea has been explored in more detail by Nettle (2005, 2006), who makes explicit the notion that any given personality trait involves a tradeoff between the benefits and costs incurred by that trait. For example, extraversion is associated with greater mating success, more social support, more physical activity, and more exploration of one's environment. However, extraverts tend to expose themselves to higher levels of risk. People who are hospitalised due to accident or illness tend to be higher in extraversion than those who are not (Nettle, 2005). Furthermore, people who suffer traumatic injury tend to be high in sensation seeking (Field & O'Keefe, 2004), which is one of the underlying facets of extraversion.

In addition, extraverts are more likely to migrate, be involved in criminal or antisocial behaviour, or get arrested. As Nettle (2006, p.625) notes, all of these are sources of risk, 'that in the ancestral environment might have meant ostracism or death'. Because extraverts tend to have more sexual partners, they are more likely to expose offspring to step-parenting, itself a risk factor for child well-being.

The other aspects of personality have similar costs and benefits, though these may not always be as apparent in our contemporary environment as they would have been in our ancestral environment. For instance, neuroticism may have a protective effect by making people vigilant with regard to dangers. However, high levels of neuroticism are associated with negative health issues, such as stress and depression. Nettle (2006, p.626) notes that certain groups of risk takers, such as alpinists and mountaineers, are frequently found to be low in neuroticism (e.g. Egan & Stelmack, 2003; Goma-i-Freixanet, 1991).

There is not a stable optimum level of any personality trait. The costs and benefits of different traits vary across time depending on the nature of the environment, and may also be different for males and females. For example, an important trait in animal species is exploratory behaviour, which has a substantial heritable component. In the great tit species (*Parus major*), bolder females are more likely to survive poor years when food is less abundant, because they engage in greater exploratory behaviour and competition for what resources there are (Dingemanse *et al.*, 2004). However, when food is abundant, bolder females do less well, possibly because they become involved in dangerous and aggressive encounters that have no benefit. Males, who are the dominant sex, show the opposite pattern. Much of their effort is devoted to defending territory. Thus, when food is scarce, competition for territory is relaxed due to the higher levels of mortality, and so the

less bold males do well. However, when food is abundant it is the bolder males that thrive.

The five-factor model of personality and its relation to risk taking A model of personality that has received wide attention is the five-factor model proposed by Costa and McCrae (1985, 1992). It proposes that a considerable amount of the variation in personality can be explained in terms of extraversion, neuroticism, openness, conscientiousness, and agreeableness. Nicholson *et al.* (2005) investigated the relationship between the 'Big 5' personality factors and risk propensity in a variety of domains, using a sample of MBA students and company executives. The risk domains investigated were recreation, health, career, finance, safety (for example fast driving), and social (for example standing for election). Participants were asked to rate themselves in terms of both past and present behaviour.

Higher levels of extraversion were associated with greater risk propensity in all domains, apart from career risk. Similarly, greater openness was associated with greater risk propensity in all domains, apart from safety. By contrast, higher levels of neuroticism, agreeableness, and conscientiousness were associated with lower levels of risk propensity in almost all domains. One exception was a small positive association between neuroticism and health. Nicholson *et al.* also found that sensation-seeking was the primary predictor of risk taking in four domains, and for risk taking overall.

As expected, men showed higher levels of risk propensity than women, but this declined significantly with age. Age-related reductions in risk taking were greatest for the domains of recreation, health, and safety – domains where risky behaviours are most strongly associated with youth.

Sensation seeking As has been noted above, sensation seeking is one of the facets of extraversion in the 'Big 5' model of personality. However, it is worth noting that sensation seeking has been a major topic of investigation in its own right (e.g. Zuckerman, 1979; Zuckerman *et al.*, 1964). The trait of sensation seeking appears to have quite a strong heritable component (Zuckerman, 2005). Risky behaviour in various domains has been linked to sensation seeking, for example: recreational sport (Zuckerman, 1983), driving (Burns & Wilde, 1995), drug use (Franques *et al.*, 2003; Pedersen, 1991), alcohol use and sex (Kalichman *et al.*, 2003). Interestingly, given the contention that personality variations represent cost-benefit tradeoffs, there is research linking low levels of sensation seeking to Parkinson's disease (e.g. Evans *et al.*, 2006).

The most recent version of Zuckerman's Sensation Seeking Scale is Form V (SSSV) (Zuckerman *et al.*, 1978). This contains four subscales:

1. Thrill and Adventure Seeking (TAS). This scale measures the desire to engage in activities that provide unusual sensations of speed or gravity, and which may be a little frightening.
2. Experience Seeking (ES). This measures the desire for stimulation via the mind and senses, through music, art, travel, and psychedelic drugs.
3. Disinhibition (Dis). This measures sensation seeking via social activities such as drinking, parties, varied sexual partners, and so on.
4. Boredom Susceptibility (BS). This represents an intolerance for repetitive experience of any kind.

There have been many studies examining the sensation-seeking profiles of those involved in sports. For instance, two studies have reported that climbers, including alpinists and mountaineers, score higher than controls on ES and TAS (Cronin, 1991; Goma-i-Freixanet, 1991). A recent survey of people involved in sports found that those involved in high-risk sports (e.g. mountaineering) scored higher on all subscales than those involved in low risk sports (e.g. swimming). However, although it has been hypothesised that sensation seeking is often associated with impulsiveness (e.g. Zuckerman, 1994), the high-risk sportspeople were no more impulsive than the low-risk sportspeople. Evidence suggests that sensation-seeking tendencies in sportsmen are often mediated by the skill and knowledge required to cope with the risks involved (Kerr, 1997).

Sex and race[3]

A wide range of studies have shown that men tend to judge risks as smaller than women (Slovic, 2000). Furthermore, these sex differences have been observed among certain groups of experts. Female physical scientists have been found to judge the risks from nuclear technologies as higher than their male counterparts (Barke *et al.*, 1997). Likewise, female members of the British Toxicological Society judged societal risks as higher than did the male members (Slovic *et al.*, 1997).

Although it is tempting to look towards biological explanations Slovic (2000) pointed to other factors that may be involved. Reviewing data reported by Flynn *et al.* (1994), he noted that white males tended to perceive less risk than non-white males and both white and non-white women (the latter three groups all shared fairly similar perceptions). In fact, the white males with the lowest risk-perception scores (about 30 per cent of white males) were better educated than most of the rest of the sample, had higher household incomes, and were politically more conservative. They also tended to hold different attitudes, characterised by 'trust in institutions and authorities and by anti-egalitarian attitudes, including a disinclination toward giving decision-making power to citizens in areas of risk management' (2000, p.401).

In short, Slovic (2000) suggested the possibility that people's risk perceptions may be related to the extent to which they 'create, manage, control, and benefit from' major technologies and activities. Consistent with this notion, experimental research has shown that manipulating a person's feelings of power changes their level of risk taking. Anderson and Galinsky (2006) found that increasing a person's feelings of power (either consciously, or unconsciously via a priming manipulation) led them to perceive risks more optimistically. They also behaved in a more risky fashion in a business scenario, a sexual scenario, and a negotiation scenario.

Nonetheless, it is likely that biological factors do play a part. Indeed, from an evolutionary psychology viewpoint behavioural

differences associated with status are entirely compatible with a biological viewpoint. Consider testosterone, a hormone that contributes to the development and maintenance of masculine features, and is implicated in risk-taking and dominance behaviours (Buss, 1999). In one study, researchers measured the testosterone levels of MBA students, some of whom had been involved in a new business venture start-up prior to their studies (White *et al.*, 2006). Those with entrepreneurial experience showed higher levels of testosterone,[4] as well as a greater propensity for risk taking (as measured by a questionnaire). Another study found that, among male fans of both soccer and basketball, a victory led to an increase in testosterone whereas a defeat led to a decrease (Bernhardt *et al.*, 1998).

In most, but not all, domains men take greater risks than women (see for example the meta-analysis by Byrnes *et al.*, 1999). This can have both anti-social and pro-social outcomes. Consider violence as a case in point. On average, men are more violently aggressive than women; in particular, the majority of murders are committed by men, and other men are usually the victims (e.g. Daly & Wilson, 1990). When examined by age, the number of male murder victims shoots up during the adolescent years, peaks in the mid-twenties, and slowly declines thereafter. Female victims increase slightly in the adolescent years, but are nowhere near as numerous as male victims.

Women's aggression, when it occurs, is more likely to be verbal than physical, so carries less risk. These patterns can be explained by an evolutionary model of intrasexual competition (see Buss, 1999). In brief:

> Males are more often the perpetrators of violence because they are the products of a long history of mild but sustained effective polygyny characterized by risky strategies of intrasexual competition for access to the high-investing sex. The fact that men die on average seven years earlier than women is but one of the many markers of this aggressive intrasexual strategy. (Buss, 1999, p.287)

For women, physical risk taking entails greater potential costs in relation to benefits, except when they are called upon to protect their children (Campbell, 1999). However, their verbal aggression often focuses on the physical attractiveness of the other person, or involves the use of unfavourable names relating to sexual behaviour (for example 'whore'), or involves the spread of rumours designed to damage their reputation.

There are also sex differences in certain pro-social forms of risk taking. For instance, Johnson (1996) analysed data for the 676 recipients of Carnegie heroism awards up to 1995. About one in five of all rescue attempts ended in the rescuer's death. Johnson found that about 92 per cent of the acts were performed by males, and about 60 per cent of the objects of rescue attempts were also male. Of rescue attempts by women, 20 per cent were to assist relatives, as compared to just 6 per cent of attempts by men. Of rescue attempts by men, 68 per cent were to assist strangers, as compared to 47 per cent of attempts by women. These contrasts were even stronger in cases where the rescuer died.

Why do people take risks to help others (and animals too; see Box 12.1)? Hamilton's (1964) *theory of kin selection* proposed that organisms are more likely to make sacrifices for another organism if that other carries copies of their own genes. In particular, sacrifices are more likely to be made for close relatives such as siblings and offspring, as compared to cousins and to strangers. This behaviour is predicted to occur when the costs for the helper are lower than the benefits for the recipient multiplied by the degree of relatedness between the two individuals. Hamilton summed this up in the following formula:

$$rB > C \quad (12.1)$$

where r represents the degree of relatedness, B is the benefit to the recipient, and C is the cost to the helper.

Scenario studies support this model of altruism (e.g. Burnstein *et al.*, 1994; Neyer & Lang, 2003). What is slightly more puzzling is why people should take risks for people who are not related to them. One possibility is based on the fact that, throughout most of human history, people have lived in small, kin-based groups. Perhaps we have not evolved a perfect mechanism for distinguishing among kin and non-kin, because this has not been necessary (van Vugt & van Lange, 2006). In support of this notion is evidence that people are more willing to help unrelated children who share their facial features (DeBruine, 2004), and are more willing to help people who share their name or speak the same dialect (Barrett *et al.*, 2002), or who hold similar attitudes (Park & Schaller, 2005).

Also, taking risks can result in enhanced status and reputation, especially for men. Indeed, Bassett and Moss (2004) found that women prefer risk-taking men as both short- and long-term partners. Farthing (2005) reported that women found heroic risk takers more attractive than did men, but were not attracted by non-heroic risk takers (for example, those involved in risky sports). However, his methodology was criticised by Wilke *et al.* (2006), who reported the results of their own study involving both German and American participants. They found that women gave higher attractiveness ratings to men involved in recreational risk taking and social risk taking. However, risk taking in the domains of health, ethics, and gambling was regarded as unattractive. Wilke *et al.* also found that male participants' ratings of female attractiveness was influenced in the same way, i.e. higher ratings for those involved in social and recreational risk taking. These authors also found that males and females tended to prefer partners who shared their own risk profiles.

Expert versus lay perceptions of risk

Are experts better than laypeople at assessing risk? The answer would appear to be 'Yes'. Slovic *et al.* (1979) report asking people to rank order a list of 30 risks. Three different samples of laypeople ranked nuclear power near to the top of all risks (with two of the groups ranking it as *the* riskiest item), whereas professional risk assessors ranked it twentieth. The risk judgments of all groups correlated with actual frequencies of death, but the relationship was much stronger for the experts. Furthermore, when laypeople gave their own fatality estimates these were no more closely related to their risk judgments than the actual frequencies of death.[5]

It would be a mistake, however, to assume that experts themselves view risk purely in terms of the objective characteristics of

BOX 12.1. RISK TAKING IN ANIMALS

Like humans, animals also take risks. For example, when birds spot a predator they often engage in a 'distraction display' that is designed to attract the predator's attention away from some object, typically the bird's eggs or young. One frequently used display is the feigning of injury, such as performing normal actions but with interruptions. Although some have suggested that such risks are taken for the benefit of the social group (Wilson, 1975), most evolutionary researchers now consider that such actions are taken in the interest of perpetuating one's own genes. This is the basis of kin selection theory (see main text; see also Dawkins, 1976).

Animals have also been known to take risks to assist humans. This story is from de Waal (2005, pp.171–172):

> In 2004, Jet, a black Labrador in Roseville, California, jumped in front of his best friend, a boy, who was about to be bitten by a rattlesnake, and took the serpent's venom. Jet was rightly considered a hero. He wasn't thinking of himself; he was a genuine altruist . . . This shows the risks animals are prepared to take. The boy's grateful family spent four thousand dollars on blood transfusions and veterinary bills to save their pet.

De Waal says that Jet probably considered the boy to be a member of his pack. This example, like human examples of assistance to strangers, most likely originates from mutuality and the assistance of kin (see main text).

The examples above involve animals taking risks in order to help others. There is also evidence that, like people, animals take risks just for the thrills involved. Perhaps (as with people) there are reputational, and hence mating, benefits to be obtained, though I am not aware of any evidence for this. Here are three examples quoted from the animal researcher Jonathan Balcombe (2006, pp.86–88):

1. Orang-utans in Tanjung Putting, Borneo, play a sport that human observers call 'snag riding', which involves hanging onto a falling dead tree, then grabbing a vine or other vegetation to escape before the tree hits the ground.
2. Ravens are inclined to goad dozing wolves, which suggests the exhilaration of danger given the risk of being caught and killed. Wolf authority David Mech and others have seen ravens dive at wolves resting on lake ice, walk up and peck sleepy wolves on their tails, and even alight on their backs. Wolves will lunge at and stalk ravens, who may evade them often at the last minute, as in a game.
3. I have watched bold squirrels repeatedly approach to within an inch of the screen door at the rear of my home, behind which a cat sits riveted and ready to pounce. The squirrels seem fully aware that the cat is there, but are equally confident in their untouchability. The cat actually did pounce on one occasion, causing the squirrel to leap back and vanish in a flash. The saucy rodent returned within two minutes, as if to call on another adrenaline rush.

the entities they are assessing. Carlo *et al.* (1992) conducted a study where a large sample of epidemiologists, toxicologists, physicians, and general scientists received information about three substances. Each subject was read a brief vignette written to reflect the mainstream scientific thinking about one substance. For half of the respondents this substance was referred to only as substance X, Y, or Z. The other respondents were told that the substance was either dioxin, radon, or environmental tobacco smoke (ETS). Revealing the name of dioxin didn't have any significant effect on health risk assessments but the experts were significantly more likely to rate radon and ETS as a serious environmental health hazard when they were told the name of the substance. Carlo *et al.* concluded that experts' evaluations of scientific data may be influenced by values and experiences which might in turn bias estimates of risk.

It would also be misleading to conceive of experts as a homogenous group with a shared consensus about the risks in their domain. The Kraus *et al.* (1992) study, briefly mentioned earlier, showed differences between toxicologists and the lay public, but also substantial disagreement among the experts themselves. The toxicologists were particularly divided about the ability of animal tests to predict a chemical's effect on humans. Toxicologists working in industry saw chemicals as more benign and were somewhat more confident than their counterparts in academia and government in the general validity of animal tests – except when those tests provided evidence for carcinogenicity, when many of the industrial experts changed their minds. Based on these findings, Kraus *et al.* suggested that controversies over chemical risks may be fuelled as much by limitations of risk assessment and disagreement among experts as by public misconceptions.

RISK COMPENSATION

Earlier, we saw that – once variations in risk *perception* are controlled for – risk taking is fairly stable across different domains. Along similar lines, Adams (1995) has described a theory of *risk compensation* in which individuals vary in their propensity to take risks. Risk-taking propensity is influenced by the rewards of risk taking, and 'individual risk-taking decisions represent a balancing

act in which perceptions of risk are weighed against propensity to take risk' (1995, p.15). The more risks that a person takes, the greater the number of rewards and losses he or she incurs (on average). People's perceptions of risks and their propensity to take risks are also influenced by 'cultural filters' in the form of various belief systems. Adams likens the balancing act performed by people to a thermostatically controlled system, in which some people have their thermostat set higher than others.

One of the implications of this model is that when people's perceptions of risks are reduced, they respond by taking greater risks. An example of this risk compensation is the Davy lamp. This device

> which most histories of science and safety credit with saving thousands of lives, is usually described as one of the most significant safety improvements in the history of mining. But it appears to have been a classic example of a potential safety benefit consumed as a performance benefit. Because the lamp operated at a temperature below the ignition point of methane, it permitted the extension of mining into methane-rich atmospheres; the introduction of 'the safety lamp' was followed by an *increase* in explosions and fatalities. (Albury & Schwarz, 1982, cited in Adams, 1995, p.211)

Of course, when people's perceptions of risk are *increased* then the opposite effect is predicted: they should take less risks. In Sweden in September 1987, all motorists were required to switch from driving on the left side of the road to driving on the right. Consequently, pedestrians had to remember to look left rather than right when crossing a road. Considerable publicity was devoted to the dangers involved in this change. In the event, however, Sweden had the safest September on the roads in many years. Eventually, people got used to the new system and 'by November they were back to their normal (acceptable?) rate of killing each other' (Adams, 1995, p.143).

Adams's own test case for the risk compensation model is the impact of motor vehicle seat belt legislation. His review of road death statistics from around the world showed that seat belt legislation has had no appreciable effect on the level of driver fatalities. The only country to show a decrease in driver fatalities following legislation appears to have been the UK, but this reduction seems to have actually been due to a drink-driving campaign rather than to the effect of seat belt use. On the other hand, deaths of rear-seat passengers (who were originally exempt from legislation) increased, as did the deaths of pedestrians and cyclists. In other words, once compelled to wear seat belts drivers felt more safe and drove faster, displacing the risk onto other road users.

A similar lack of positive evidence can be found in relation to the use of safety helmets for cyclists. Indeed, there is evidence that wearing safety equipment actually puts cyclists at greater risk. Walker (2007) fitted a camera to his own bicycle so that he could measure the proximity of overtaking vehicles. He found that vehicles passed closer to his bicycle when he was wearing a safety helmet than when he was not. Walker suggested that motorists expect cyclists without helmets to be less reliable. He also found that vehicles gave him a wider berth when he donned a long-haired wig that made him look like a female in the eyes of approaching drivers.

RISK COMMUNICATION AND MENTAL MODELS

Given the numerous factors described in this chapter that influence risk judgments, it is little surprise that the communication of risk continues to pose a major challenge for governments, companies, and public bodies. Morgan (1993, p.29) described a traditional method for communicating risk, in this case the risk to US householders from radon:

> The EPA prepared this brochure according to traditional methods: ask scientific experts what they think people should be told and then package the result in an attractive form. In fact, people are rarely completely ignorant about a risk, and so they filter any message through their existing knowledge. A message that does not take this filtering process into account can be ignored or misinterpreted.

Aside from the background knowledge that people may bring to bear on a risk communication, there are other issues too, some of which we have encountered in previous chapters. For instance, given the difficulty that people sometimes have in thinking about probabilities, would it be better to use natural frequencies instead – for example, '10 people out of every 1000' instead of '1 per cent' (see Chapter 3)? What about low-probability risks? Zeckhauser and Viscusi (1990) pointed out that with an estimated 10^{-7} annual risk it would take many years of widespread observation even to learn whether the risk is of an order of magnitude that is higher or lower. As carcinogenic risks are often coupled with long time lags and a complex (multiple factor) causality, precise inferences may not be possible.

A widely held view in current risk research is that public risk perceptions should not be regarded as an annoyance that needs correcting but as informative regarding people's concerns, and something that needs to be taken into account when devising risk communications (e.g. Pidgeon *et al.*, 1992). A fruitful recent approach has been to investigate people's mental models (see Chapter 6):

> We have concluded that the only way to communicate risks reliably is to start by learning what people already know and what they need to know, then develop messages, test them and refine them until surveys demonstrate that the messages have conveyed the intended information. (Morgan, 1993, p.29)

Morgan and colleagues have described how they developed a brochure to communicate information about radon gas in private homes (Bostrom *et al.*, 1992; Morgan, 1993). This involved the use of open-ended interviews beginning with the request, 'Tell me about radon.' As the interviews proceeded the questions got more specific. After about a couple of dozen people had been interviewed no further new ideas emerged, so the researchers used the responses to devise a closed-form questionnaire that was then administered to a much larger sample.

14 Cooperation and Coordination

KEY TERMS

behavioural game theory
chicken
common knowledge of rationality
communication
convention
culture
dictator game
diffusion of responsibility
empathy
ethics
fear
free-rider
game theory
generosity
greed
Nash equilibrium
prisoner's dilemma
public goods game
punishment
social value orientation
tit-for-tat
trust
ultimatum game

CHAPTER OUTLINE

INTRODUCTION

Human interactions can be characterised as involving cooperation, coordination, and competition, sometimes in combination. This chapter looks at the first two of these, although we will also see some instances of competitive behaviour. I shall begin by briefly considering game theory, which is a mathematical discipline devised to analyse how rational people would behave in interactive situations (behavioural game theory is concerned with how people actually behave). We will return to some of the game-theoretic ideas throughout the chapter, but without getting into any deep mathematical waters.

In the second main part of the chapter, I shall look at cooperation: how it has been studied, why it exists at all, and the factors that influence whether or not people will behave cooperatively in a given situation. In the third part of the chapter, I shall look at situations where people need to coordinate their behaviour. This is distinguished from mere cooperation in the sense that the situations concerned allow for more than one 'best'[1] combination of behaviours by the individuals concerned, hence they must coordinate on a single combination.

GAME THEORY AND BEHAVIOURAL GAME THEORY

In a task known as the *ultimatum game* two participants bargain over an amount of money, such as $10. A Proposer has to offer some proportion of the $10 to a Responder. If the Responder accepts the offer, then she gets to keep that sum and the Proposer keeps whatever he has left.[2] If the Responder rejects the offer then neither party gets to keep any money. In the experimental situation there is no communication between Proposer and Responder, so there is no discussion or negotiation about the offer made. In fact, there are normally very tight experimental controls such that Proposers and Responders are anonymous to each other. Take a moment to think about the sum that you would offer if you were in the Proposer's position, and what you would accept if you were the Responder.

From the point of view of analytical *game theory*, described below, both participants should act in a way that maximises their self-interest. The Responder should accept any amount that is offered, because however small this is she will still go away with more than she had before. The Proposer should anticipate this and offer the smallest amount possible ($0.01, unless the study stipulates some other 'smallest unit of exchange').

You may not be surprised to learn that people actually behave in a rather different way than that suggested by this analysis. Camerer (2003, pp.50–55) has reported the results of many 'one-shot' studies from around the world, involving monetary stakes of different sizes. The results are highly consistent: the modal and median offers are usually 40–50 per cent and the means are 30–40 per cent. Offers of 40–50 per cent are normally accepted, whereas offers below 20 per cent are rejected about half the time.

Game theory is the analysis of interactions between rational agents. The term 'game' does not literally mean a game as in football or hockey; rather, it simply refers to a situation involving two or more agents, each of whom has two or more *strategies* (courses of action) available to them, where strategies are associated with payoffs, and the payoffs also depend on the action that the other person(s) takes. In reality, people may not demonstrate perfect rationality, but game theorists expect that people will converge towards rational solutions as they gain experience with a game.

Game theory assumes that players in a game share common knowledge of the rules. Where this is not the case then game theory cannot be applied. Game theory also assumes *common knowledge of rationality*, meaning that players assume the rationality of other players.

As with the ultimatum game, studies of other games have often found that people do not behave in accordance with the predictions of game theory, though as mentioned above this may change with experience. *Behavioral game theory*, sometimes also called *psychological game theory*, is the study of how people actually behave in interactive situations.

COOPERATION

Extending our cooperation to other people always involves some sacrifice on our part. We give up time that we could be devoting to other purposes, we invest cognitive or physical effort, and we may also contribute financially. Cooperation in everyday life takes many forms, from helping a friend to move house, to paying taxes, belonging to a union (and possibly engaging in industrial action), to recycling one's household waste, and so on. Of course, not everyone cooperates in such behaviours. Understanding why anyone should cooperate at all, and how we can encourage greater cooperation, has been widely investigated.

The prisoner's dilemma and public goods games

One of the tasks that has been used to investigate cooperation is the *prisoner's dilemma*. This is a very influential problem that mirrors many real-life situations, but was first studied by Merrill Flood of the RAND corporation (Flood, 1952, cited in Poundstone, 1992). Here is an example of a prisoner's dilemma.

Suppose Dan and Joe have been arrested by the police for their part in a serious crime. They have been put in separate cells and cannot communicate. The police don't have enough evidence to convict them of the most serious charge unless at least one of the men confesses. If neither confesses, then they can each expect a year in prison on a lesser charge. Each man gets a visit from a police officer who tells them that they can escape prison by implicating their partner, who will be sent to prison for 10 years. The catch is that if both men implicate each other then they will each get five years in prison. These outcomes are shown in Table 14.1.

The strategies open to Dan and Joe are usually referred to as cooperation or defection, where cooperation means not telling the police anything, hence not implicating your partner. Defection means implicating your partner. Suppose, for argument's sake, that you were in Dan's position. What would you do?

Game theory makes a very specific prediction about what Dan and Joe both do if they are rational people. Suppose Dan believes that Joe will stay quiet. If Dan also stays quiet then he will get a year in prison. However, he can improve on this outcome by implicating Joe, thereby not going to prison at all. But what if Dan thinks that Joe will defect? If Dan stays quiet, he will get 10 years

Table 14.1. *The prisoners' dilemma. Each cell shows the prison sentence (years) for Dan and Joe, respectively, if that combination of strategies is chosen*

		Joe	
		Implicate Dan	Stay quiet
Dan	Implicate Joe	5, 5	0, 10
	Stay quiet	10, 0	1, 1

in prison. But he can improve on this outcome by implicating Joe. In other words, implicating Joe is the best strategy for Dan. As you can see from Table 14.1, by the same process of reasoning on Joe's part his best strategy is also to implicate his partner.

Many games, though not all, have a *Nash equilibrium* (or more than one in some games). This is a combination of strategies that cannot be improved upon, as long as the other player sticks with their chosen strategy. In the prisoner's dilemma, joint defection is the Nash equilibrium, and because there is only one equilibrium point mutual defection is regarded as the dominant strategy. Sure, the outcome is less than ideal for both Dan and Joe, but it is the best they can get *if they are both rational actors pursuing their self-interest*. In reality, people cooperate more than game theory predicts – about a third of the time, and more so on repeated plays of the game. Nonetheless, they do not do as well as they might do because there is always the temptation to defect at some point, which leads to retaliation by the other player.

A similar task to the prisoner's dilemma is the *public goods game* or resource allocation task. In a typical task, individuals are provided with an endowment and then given the option to contribute some, all, or none of that endowment to a pool of resources. After a round during which contributions may have been made, whatever is in the pool of resources is increased by some proportion. Then the resources are divided up among all participants. As with the prisoner's dilemma, the collectively rational action is to contribute to the pool whereas individually rational action is to withhold contribution. This has parallels with many real-life situations, such as (in the UK) the temptation not to buy a television licence while taking advantage of the public broadcasting service – the BBC – that has been funded by the people who did pay their licence fee.

Alternatively, a task may give people the opportunity to withdraw resources from a limited pool, thus risking the depletion of a public good. This also has parallels with many contemporary problems, such as concerns over energy consumption and overfishing by competing trawlers (see Hardin, 1968). In these situations, the collectively rational solutions are to use less energy and to catch less fish, but in each case the individually rational solution is the opposite.

The evolution of cooperation

Why should cooperative behaviour exist at all? To understand how cooperation might evolve in a population, Axelrod (1984) pitted various strategies against each other in a series of repeated-play prisoner's dilemma computer tournaments. Most of these strategies had been submitted by game theorists, psychologists, economists, and other academics. In his discussion of the strategies, Axelrod distinguished between 'nice' rules, which always begin by cooperating, and 'nasty' rules, which always begin by defecting. In two of the tournaments, the strategies were played for points. In a third round, the reward for successful strategies was 'offspring', such that some strategies thrived and others became 'extinct'.[3]

Axelrod found that the sophistication of the strategies was not an indicator of their success. For example, one of the most sophisticated strategies was called the *Downing* strategy. This began by assuming that the other program would be unresponsive to its own cooperation or defection, but then amending this assumption on the basis of how the other program actually did respond to cooperation and defection. If the other program tended to punish defection but reward cooperation, then Downing would settle for cooperation. Otherwise, it tended to defect. Unfortunately, the opening assumption of unresponsiveness led Downing to get punished for its own early defection. By contrast, the most successful strategy was also the simplest: *tit for tat* (TFT), submitted by Anon Rapoport. TFT begins by cooperating and every subsequent action copies the action of the other player. In other words, TFT punishes defection but rewards cooperation.

In the 'evolutionary' competition, overly nice strategies were preyed upon by nasty strategies, to the point of extinction. However, when the nasty strategies had no prey left they too died out, paving the way for TFT. Thus, Axelrod's computer tournaments showed how the capacity for cooperation can evolve within a population.

The astute reader will by now have noticed that I have shifted from talking about one-shot games to talking about repeated games. Many real-life interactions may be more like repeated games, because we often engage repeatedly in similar interactions, often with the same person or organisation. In fact, because we often do not know when we will cease to interact with someone our interactions may be considered indefinitely repeated games. Importantly, the notion of finite equilibria that applies in one-shot games does not apply in indefinitely repeated games. In fact, for players who are patient any individually rational payoffs can be supported by an equilibrium; this result – not explored further here – is known as the Folk theorem.

In the next part of this chapter I shall examine the factors that influence the extent to which people cooperate.

Consideration of others

About a third of participants cooperate on a single-shot prisoner's dilemma. This is more than would be expected if people were rational actors behaving according to self-interest (as in the game-theoretic analysis). Can this cooperation be attributed to a kind of moral imperative on the part of the participants? Shafir and Tversky (1992) thought not. They asked participants to play a series of one-shot games, each against a different person. Participants were told that they had been allocated to a bonus group that would occasionally be told their opponent's strategy prior to making their own choice. If the third of participants who normally cooperate in a one-shot game are behaving according to a moral imperative, then we might expect the highest rate of cooperative responses to occur when it is *known* that the other person has cooperated.

Shafir and Tversky found that 37 per cent of games resulted in cooperation when the other person's strategy was *not* known, which – as we have seen – is about the usual rate of cooperation. However, contrary to the expectation based on moral imperative, only 16 per cent of responses were cooperative when it was known that the opponent had cooperated (3 per cent of responses were

cooperative when it was known that the opponent had defected). Shafir and Tversky argued that people fail to think through the standard prisoner's dilemma. Specifically, they do not consider all the hypothetical outcomes, so do not know how they would behave if they knew their opponent's strategy. The cooperative responses that do occur may involve a degree of wishful thinking.

Is cooperation facilitated by taking the perspective of the other person? Epley *et al.* (2006) reported a series of studies involving simulated and actual resource-allocation negotiations (for example, one study involved role playing the representative of a fishing organisation in negotiations with other such groups regarding the reduction of harvesting levels). All participants began by stating what they thought was a fair allocation. Half of the participants were then asked to consider what resource allocation other groups might consider fair for themselves and then asked again what they thought a fair allocation would be for all groups. Taking the other groups' perspective into account in this way led people to reduce what they thought a fair allocation was. However, when asked to state how much they would actually take for themselves (or when given the opportunity to actually take some resources), these participants took a larger allocation for their group than did those who had not considered the other groups' perspectives. It appears that taking others' perspectives led people to generate egoistic theories of their likely behaviour, thus leading them to behave more egoistically themselves.

Similar results have been reported by Caruso *et al.* (2006), including one study using the prisoner's dilemma task (Experiment 3). This task included a control group, as well as participants who were asked to take the cognitive perspective of the other person, and participants who were asked to empathise with the other person. The rate of defection was the same for empathisers and control participants (40 per cent vs. 41 per cent, respectively), whereas there was a 68 per cent defection rate among those who had taken the cognitive perspective of the other person.

Another study found that reducing people's capacity for thinking made them behave more fairly in a resource problem (Roch *et al.*, 2000). In this study, participants were given the opportunity to withdraw resources from a common pool, and asked to verbalise their thoughts as they did so. However, people in a high cognitive load condition were asked to hold in mind an eight-digit number while they did the task. These participants were less likely to verbalise task-relevant thoughts, and were more likely to draw an equal share from the pool, whereas those without the extra cognitive load referred to the task more often and tended to take more than their 'fair share'.

In short, these studies lead to the slightly disturbing conclusion that greater cooperation is not more likely to occur when people think harder about a task, at least not in laboratory studies. The following section reviews evidence that different levels of cooperation are actually related to the motives and different values that people possess.

Fear, greed, and punishment

Several studies have examined the effect that *fear* and *greed* might have on rates of cooperation in social dilemmas. Fear, of course, could mean two things in this context. It could mean fear of being taken for a sucker, for example if one behaves in an environmentally friendly fashion when no one else does. Alternatively, it could mean fear of being punished for acting in a self-interested manner, for example being fined for behaving in an environmentally unfriendly manner.

Fear appears to partly explain the generosity of most Proposers in the anonymous ultimatum game. If this generosity is related to the fear that Responders will reject small offers, then we can gauge the extent of this fear by seeing what happens if we take away the Responder's power of rejection. This is the basis of the *dictator game*. It turns out that Proposers do offer less when there is no possibility of rejection, but they still offer more than they would if they were behaving out of pure self-interest (for example, about 20 per cent in studies by Forsythe *et al.*, 1994). However, there is also evidence that dictators are even less generous when great pains are taken to convince them that their responses are anonymous and that their identity is hidden from both the recipient and the experimenter (e.g. Hoffman *et al.*, 1994). Thus, greed may be the flip side of fear: people often behave in a greedy fashion once the fear is removed.

Along similar lines, Ahn *et al.* (2001; see also Dawes *et al.*, 1986; Yamagishi & Sato, 1986) found effects of fear and greed in a repeated-play prisoner's dilemma, but only when participants were randomly paired in each game, and not when the same people played each other in each game. These authors linked greed and fear to particular payoff relationships in the task, as shown in Table 14.2. The four possible outcomes are temptation (T), reward (R), punishment (P), and sucker (S), where $T > R > P > S$. Greed is based on the payoff that can be obtained if you defect when the other player cooperates ($T - R$). Fear is based on the cost to you of cooperating when the other player defects ($P - S$). Ahn *et al.* asked participants to play a series of prisoner's dilemma games in which the size of the $T - R$ and $P - S$ relationships was manipulated. Initially, fear and greed appeared to have little effect. However, when these relationships were normalised, by dividing each by $T - S$, then both fear and greed were related to the degree of cooperation. In fact, greed had a stronger effect on behaviour than did fear.

Fear is also an important factor in a situation where there is a possibility of punishment. Fehr and Gächter (2002) found that above-average contributors to an anonymous public goods game were willing to punish below-average contributors, even though this entailed a cost to themselves. The effect of this was to substantially increase the level of contributions above what was observed in non-punishment periods of the game. Fehr and Gächter considered

Table 14.2. *A generalised view of the payoffs in prisoner's dilemma tasks (P: punishment; R: reward; S: sucker; T: temptation)*

		Player 2	
		Defect	**Cooperate**
Player 1	Defect	P, P	T, S
	Cooperate	S, T	R, R

the punishment to be altruistic because it benefited the group at a cost to the punisher. However, they also found that punishment was motivated by anger, which might suggest that the benefit to the group was merely an incidental effect of revenge.

Evidence of a revenge motive has been seen in a neuroimaging study by Singer *et al.* (2006). They found different types of brain activation when people observed pain being experienced by a partner who had previously behaved fairly and one who had behaved unfairly in an economic game. Seeing a fair partner experience pain led to activation in pain-related brain areas (the fronto-insular and anterior cingulate cortices). When seeing an unfair partner experience pain this elevated activity was slightly reduced for women, whereas men no longer showed significant activity in these areas. This decreased activity for men was accompanied by increased activity in the nucleus accumbens, an area associated with reward processing. Furthermore, the degree of activity in this area correlated with men's rated desire for revenge.

Differences in values

Other research indicates that people's values may affect behaviour on social dilemma problems. Along similar lines to the concepts of fear and greed discussed earlier, Van Lange (1999) proposed that people can be categorised according to their *social value orientation*. *Prosocial* people wish to maximise joint gain and equality in outcomes, *individualists* are interested only in maximising their own gains, and *competitors* wish to maximise relative gain (meaning the difference between one's own and the other's outcome). In one investigation, participants were interrupted halfway through a repeated prisoner's dilemma game and asked to think about the choices they had made and outcomes they had received, and to list some alternatives that were better or worse than what had actually happened. Higher numbers of upward counterfactuals ('It could have been better') were associated with an increased level of cooperation in the second half of the game. By contrast, downward counterfactual thinking ('It could have been worse') was associated with a decrease in cooperation in the second half of the task (Parks *et al.*, 2003).

Furthermore, Parks *et al.* found that prosocials generated the most upward counterfactuals and fewest downward counterfactuals, competitors showed the reverse pattern, and individualists generated the same number of each type. Social value orientation was thus linked to the post-thoughtlisting rate of cooperation, but this relationship was entirely due to the nature of the thoughts that the different types of people had generated. In short, inducing competitive individuals to think about how things might have worked out differently is likely to make them less, not more, cooperative.

Where do our social value orientations come from? Studies reported by Van Lange *et al.* (1997) found that prosocial individuals exhibited a more secure attachment style; that is, they found it easier to get closer to others and to allow others to get closer to them, and did not worry about being abandoned. Furthermore, they found evidence that prosociality may develop as a result of one's family background, because prosocial people reported having more siblings, especially sisters. It is likely that people in large families are exposed to more situations in which resources need to be shared, although perhaps prosociality has a heritable component and prosocial people have more children. In any case, Van Lange *et al.* reported a study of different age groups that found higher proportions of prosocial individuals as age increased. They suggested that younger adults more often find themselves in competitive situations, whereas older adults more frequently experience situations requiring them to give help to others or to receive help from others. This certainly indicates that experience can lead people to become more prosocial.

However, despite the evidence for greater cooperation by prosocials in laboratory tasks, it is not yet clear how far this extends to real life. A survey carried out at US gas stations and connecting points for buses found that social value orientation did not predict people's perception of the environmental impact of cars, nor their preference for public or private transportation. By contrast, people who engaged in more thought about future consequences tended to perceive a greater public impact of cars and to prefer public transport (Joireman *et al.*, 2004). This result would appear to contradict the conclusion of the section before last, but perhaps what matters is not how much one thinks, but what one thinks about (in this instance the future, as opposed to considering what someone else might be thinking or what we might have done differently in the past).

On the other hand, there is some slightly startling evidence that economics majors behave more in accordance with self-interest than do non-economics majors both in real life and in the laboratory. They are more likely to free-ride in public goods tasks (Marwell & Ames, 1981), to offer less and be willing to accept less in the ultimatum game (Carter & Irons, 1991), to give less to private charities (Frank *et al.*, 1993), and to defect more often on the prisoner's dilemma (Frank *et al.*, 1993). In the last study, although both economics majors and non-economics majors defected less as they proceeded from their freshman to their senior year, the trend towards cooperation was less among the economics majors.

Other evidence suggested that this might be at least partly due to the training that economics students receive, as opposed to more selfish students opting to study economics. Frank *et al.* gave the same questionnaire to three different groups of students at the start of the semester and again at the end of the semester. The questionnaire assessed whether students would behave honestly or dishonestly in certain scenarios (for example by reporting or not reporting a billing error) and whether they expected that other people would behave honestly in the same scenario. One group of students was studying introductory astronomy and the other two were taking different introductory microeconomics classes. However, one of the microeconomics professors emphasised game theory in his course, whereas the other didn't.

The astronomy class showed a tendency towards greater honesty by the end of the semester. However, the microeconomics students who were taught game-theoretic concepts showed less honesty and more cynicism about others at the end of semester. The other microeconomics class showed no clear change, with slightly less honesty on one item, greater honesty on two items, and no change on another.

In short, as with the research by Van Lange and colleagues, it looks as though certain values can be learned.

Culture, cooperation, and economic success

Some of the clearest differences in ultimatum game behaviour have been identified in relation to culture. Henrich *et al.* (2005) conducted games in 15 small-scale societies around the world, including peoples living in mountainous tropical forests, high-latitude deserts, and savanna woodlands. Each society was based around one of the following: family groups, family groups plus extended ties, villages, bands, clans, clan chiefdoms, or multiclan chiefdoms. In the ultimatum game, the mean offer between these small-scale societies showed more variation than it did between different industrialised societies. The smallest mean offer was 26 per cent (the Machiguenga of Peru), whereas the highest mean offer was 57 per cent (the Lamelara of Indonesia). The variation in rejection rates was much less, and rejection rates were strikingly low. Four societies had no rejections at all, including the Quichua of Ecuador, where the mean offer was just 25 per cent.

Henrich *et al.* identified two factors that were associated with much of the between-group variation. The first factor was *payoff to cooperation*: cooperation was stronger in societies where there was more cooperation with non-immediate kin (for example in hunting for whales). The second factor was *market integration*: there was more cooperation in societies where there was greater trading via markets.

It is interesting that societies with higher payoffs to cooperation and greater market integration should show higher levels of cooperation in the ultimatum game, levels that are more like those observed in industrialised societies. Although we often think of business as a competitive, and even ruthless, human activity, it actually could not survive unless there were cooperation and trust between the people involved (this theme is explored in more detail by Surowiecki, 2004). When one person pays another for some good or service, he trusts that person to provide what he has paid for. Should that good or service not be provided, then the person defaulting on the arrangement risks damage to his reputation and the loss of future customers. The link between cooperation and economic success is explored further in Box 14.1.

One unusual aspect of the data reported by Henrich *et al.* concerned the Au and Gnau of New Guinea. In these societies, Proposers occasionally made offers above 50 per cent, yet these offers were often rejected. The authors drew a parallel between this behaviour and the culture of gift-giving that is prevalent in these villages and throughout Melanesia. In these societies, the acceptance of gifts, even unsolicited ones, creates an obligation on the recipient to return the favour. When the favour is asked for it may be of a different kind, such as support in a political alliance, and it may be at a time inconvenient to the recipient. For this reason, large gifts may be refused.

Trust, generosity, and communication

Trust Earlier we saw that fear of rejection influenced offers in the ultimatum game, fear of partner's defection motivated some defections in the prisoner's dilemma, and fear of other free-riders led some people to also free-ride in the public goods task. We also saw that fear was not a factor in the public goods task when people were interacting with their friends, rather than strangers, and the level of contribution increased so long as payoffs were based on the lowest contributing member or the average level of contribution. This suggests that cooperation might be increased where trust in one's partner or other members is increased.

Trust has been defined as *the willingness to accept vulnerability based upon positive expectations about another's behaviour* (Dunn & Schweitzer, 2005). Berg *et al.* (1995) devised a game specifically to measure trust. In this game, an Investor was given $10 and told she could either keep this or invest as much as she liked. If she wished to invest some money, then the investment would be placed in the hands of an anonymous Trustee in a separate room. The invested money would earn interest. However, it was entirely up to the Trustee as to how much of the invested money would be passed back to the Investor. Berg *et al.* found that Investors invested about 50 per cent of their endowment on average. When the investment tripled in value, the average repayment by Trustees was about 95 per cent of the initial investment (about a third of the resulting amount after interest). However, there was considerable variation in the amount returned, with about half the Trustees returning nothing or just $1. The amount returned can be considered a measure of trustworthiness.

Trust and trustworthiness appear to vary considerably by country. Camerer (2003, citing Ensminger, 2000) has noted that the Orma herders of Kenya invested about 40 per cent on average but returned only 55 per cent:

> Kenya is considered one of the more corrupt countries in the world, measured by indices of 'transparency', which guess the extent of bribery, bureaucratic corruption, and black market trade, so it is encouragingly consistent that this simple game shows low levels of trust also. (2003, p.87)

People are more likely to trust people who have developed a good reputation. In the prisoner's dilemma study reported by Ahn *et al.* (2001), discussed earlier, people who played against the same person on each game cooperated more often than those who played a different person each time (42 per cent vs. 32 per cent, respectively). Cooperating on all four games was twice as frequent among those who were paired on all four games compared to those who were randomly matched (31 per cent vs. 15 per cent).

In another study (Delgado *et al.*, 2005), participants in a trust game were provided with positive, negative, or neutral information hinting at their (fictional) partner's moral character. These three biographies, respectively, referred to the partner as having rescued a friend from a fire, having tried to sell tiles from the ill-fated space shuttle Columbia, and having missed a doomed airplane flight. Although participants were told that their partner's responses (share or keep) might or might not be consistent with their description, those who were partnered with the 'good' partner rated them as more trustworthy prior to the game and shared more money with them during both early and late stages of the game, although after the game the good partners were no longer rated as any more trustworthy than the bad or neutral partners.

BOX 14.1. COOPERATION AND ECONOMIC SUCCESS

A study by Paciotti *et al.* (2005) has linked more closely the extent of cooperation within a society and its economic success. They compared the Sukumu ethnic group in Tanzania with the Pimbwe. The Pimbwe are the indigenous ethnic group and live in Rukwa, in the south-west of Tanzania, near Lake Tanganyika and the border with Zambia. The Sukuma originate from the north of Tanzania, but in the 1960s and 1970s they migrated to all areas of the country, including Rukwa, where they live outside the villages of the Pimbwe. The Sukuma are farmer-herders, whereas the Pimbwe are small-scale farmers who also fish and eat wild game.

The Sukuma have a successful justice system, the Sungusungu. This originated in Sukuma villages in the north of Tanzania during the 1980s as a group devoted to fighting back against cattle raiders from Uganda. The Sungusungu system has since been replicated in Sukuma villages throughout Tanzania, and has expanded to deal with property crime, debt disputes, adultery, and witchcraft. The Sungusungu are quite sophisticated, having various procedural rules and keeping extensive documentation (although by Western standards their processes are not always fair and can be quite brutal).

Why are the Sukuma so successful, whereas the Pimbwe have many social and economic problems and find it harder to meet their daily needs? Paciotti *et al.* suggest that the Sukuma's success is due to their high propensity for cooperation:

> Sukuma families have a reputation across Tanzania for exceptional hospitality and generosity. Visitors from their own and other ethnic groups are welcomed with lavish spreads of food. In contrast to the Sukuma, Pimbwe families are quite suspicious of others, living and eating in tightly bounded family groups and rarely opening their homes to people outside their family or clan. (2005, p.59)

Although the Pimbwe do have social institutions, these only promote social interaction and cooperation within clans and villages. As a result, the Sukuma are competing successfully with other ethnic groups and possibly Sukuma culture is beginning to replace Pimbwe culture, especially with intermarriage between Pimbwe and Sukuma and the learning of the latter's cultural traditions.

Paciotti *et al.* presented the ultimatum game to individuals from the Sukuma and the Pimbwe. Half of the participants were paired with a person from their own village and half were paired with someone from a different village. The mean offers are shown in Figure 14.1. The within-village offers by the Sukuma are the highest yet recorded (61 per cent on average) and even offers to someone of a different village are above 50 per cent. By contrast, the Pimbwe offered considerably less, especially to people from a different village.

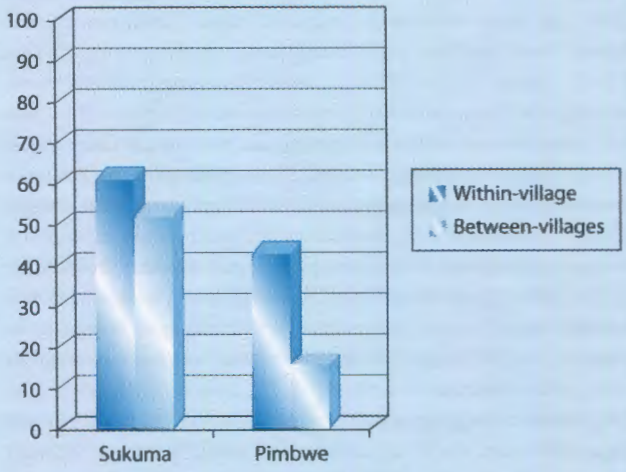


Figure 14.1. *Mean (%) amount offered by the Sukuma and Pimbwe in the ultimatum game*

This study also used functional magnetic resonance imaging (fMRI) to record brain activity during the decision and outcome phases on each trial. A brain region associated with learning (the caudate[4]) showed activity in response to partners' choices but not while the participants themselves were deciding. For neutral partners, caudate activity differed depending on whether their decision was to share or keep. For the non-neutral partners, especially the good partner, there was little difference in activity as a result of outcome decisions. Thus, it appears that prior moral perceptions reduced the extent to which learning occurred as a result of feedback.

When the participants themselves decided not to pass money to a good partner, or to share with a bad partner, there was increased activity in the cingulate cortex, an area associated with conflict monitoring. Decision making also activated the ventral striatum – associated with making predictions and anticipating outcomes – but there was only a weak indication that activity differed according to the type of partner and type of decision.

Generosity We saw earlier that the TFT strategy is highly successful for maximising outcomes in the prisoner's dilemma. However, one drawback to TFT is *noise*. Noise refers to an erroneous response. For example, in real life we are sometimes let down by other people, or we ourselves let other people down, through circumstances beyond our control rather than intentionally (for example the car breaks down on the way to a meeting).

The danger is that noise causes two interacting TFT players to enter into a cycle of never-ending retaliation.[5] Computer simulations have shown that cycles of recrimination can be avoided by adding an element of generosity or forgiveness to TFT (Bendor *et al.*, 1991; Kollock, 1993). TFT + 1 is a strategy that behaves slightly more cooperatively than the interaction partner did on the previous trial. Using real participants, Van Lange *et al.* (2002) found that TFT + 1 led to even higher levels of cooperation than TFT. Furthermore, the intentions of a partner using TFT were judged as less benign under conditions of noise, whereas partners using TFT + 1 were judged equally benign under conditions of noise and no noise.

However, in some situations people may not be able to act in a generous way (for example if generosity requires resources that one cannot afford to provide). In such circumstances communication can overcome the problem of noise, particularly for people who show low levels of dispositional trust who otherwise tend to respond more negatively to noise.

Communication Tazelaar *et al.* (2004) examined how communication would affect people's responses to a social dilemma where noise was introduced. Participants played a game that involved passing some coins (out of an endowment of 10) to a partner, to whom the coins were worth more in value. Unknown to participants, their 'partner' was a computer program playing TFT or TFT + 1. In this game TFT + 1 involved passing back one more coin than the participant had previously passed across. On a few trials, noise was introduced by giving the participant fewer coins than the computer had 'intended'. However, some participants also received an occasional communication from their bogus partner. For example: 'I wanted to give you six coins, but the computer changed my decision. I think you only received three coins.' Tazelaar *et al.* (2004, Experiment 1) found that communication eliminated the otherwise detrimental effects of noise on cooperation, and led people to view their partner as just as benign as in a no-noise condition.

A second experiment examined whether communication had a bigger effect on people who scored low on dispositional trust, as compared to those scoring high. The results are shown in Figure 14.2. The introduction of noise reduced cooperation among those low in trust when there were no accompanying messages, but people with high levels of trust were unaffected. However, when the noise was accompanied by communications from the 'partner', then there was just as much cooperation from low trusters as from high trusters. Accordingly, participants low in trust rated their partner's intentions as less benign when there was noise and no communication. However, noise accompanied by communication led these participants to rate their partner as just as benign as under the no-noise condition, and as just as benign as the high-trust participants rated their partners.

Dunn and Schweitzer (2005) have reported several studies that explore the influence emotions can have on trust. This is important because decisions about trust are often made in affect-rich contexts. Their results showed that emotions can influence trust in an individual, even though that individual was not the cause of the emotion being experienced. In particular, emotions that can be strongly aroused by other people (for example anger and

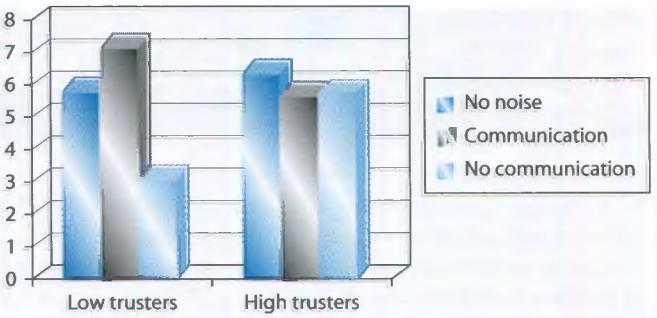


Figure 14.2a. *Mean level of cooperation (coins given), according to dispositional trust and communication condition*

Source: Tazelaar *et al.*, 2004, Experiment 2.

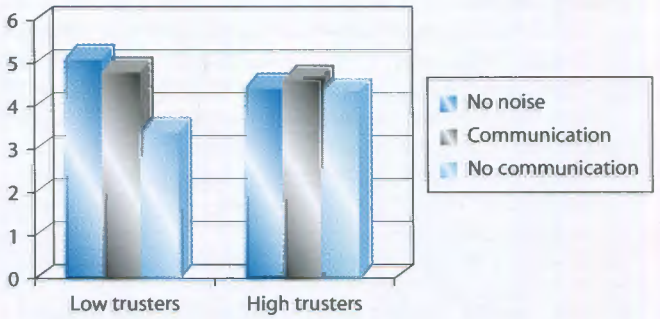


Figure 14.2b. *Mean ratings of benign intent, according to dispositional trust and communication condition*

Source: Tazelaar *et al.*, 2004, Experiment 2.

gratitude) have a bigger effect on trust than emotions that are characterised by personal control (pride and guilt) or situational control (sadness).

Some of Dunn and Schweitzer's studies elicited emotions in participants by asking them to write in detail about (a) things that might make them feel a particular emotion, and (b) a previous situation that had aroused that feeling. In another study, film clips were used to arouse emotions. After the emotion-induction procedure, participants made trust judgments about an unfamiliar co-worker or acquaintance. Across the studies, anger led to the lowest levels of trust whereas happiness and gratitude caused the highest levels of trust. However, pride and guilt (personal control emotions) appeared to have little effect on trust judgments. In these studies, the effect of emotions on trust appeared to be specific to individuals that the participants were not familiar with. When participants rated their trust in a person familiar to them they were unaffected by emotion (see Figure 14.3).

We have seen that communication can help overcome the problems caused by noise in social dilemmas. Other studies have found that cooperation is increased when group members are allowed to communicate (for a review see Sally, 1995). Various reasons have been proposed as to why communication increases cooperation, though Kerr and Kaufman-Gilliland (1994) concluded that only two explanations are supported. One explanation says

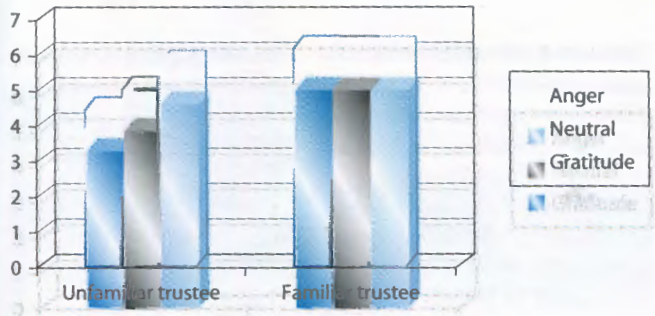


Figure 14.3. *Mean trust ratings for an unfamiliar or familiar student*

Source: Dunn & Schweitzer, 2005.

that communication promotes group solidarity, and the second explanation is that communication allows group members to express their commitment to mutually cooperate.

However, Wilson and Sell (1997) have pointed out that misleading communications could potentially decrease cooperation, if people do not live up to their promises. They conducted a public goods study which varied the ability of participants to announce how much they were going to contribute and the ability to see the past contributions of group members, thus giving rise to four conditions: (a) information about past investments only, (b) information about announcements only, (c) information about past investments and announcements, and (d) information about neither past investments nor announcements.

Wilson and Sell found that average group contributions were highest when there was no information at all and lowest when there was only one type of information available. Information about announcements and past behaviour did increase contributions, but not to the level of the no-information condition. At the start of the experiment, contributions were very high when both types of information were available, but this condition showed the strongest decline in contributions over the course of the experiment (declines also occurred in all other conditions).

More than half of the announcements (53.4 per cent) made by participants were 'lies'; that is, they did not contribute as much as they said they would. Participants contributed exactly what they had said they would in just 27.8 per cent of instances. However, in these latter cases, just over half the time (51.1 per cent) participants were signalling that they would contribute nothing. As Wilson and Sell pointed out, this is hardly the basis on which participants could build a set of cooperative strategies.[6]

Many tasks that have allowed communication between group members have nonetheless tried to keep them physically separate, with each person seated in front of a computer monitor and not easily able to see other people. We might expect that greater cooperation would occur if people were allowed to talk face to face, because nonverbal cues may be used to communicate cooperative intent. Indeed, this is what research indicates (Kurzban, 2001; Roth, 1995). One study has even reported that university staff contributed more money to an honesty box in their coffee room when a pair of eyes, rather than a control image, was displayed on a cupboard door above the box (Bateson *et al.*, 2006).

Scharlemann *et al.* (2001) found that greater cooperation was produced in a one-shot trust game involving strangers after participants saw a photograph of their partner smiling, as opposed to not smiling. Furthermore, whilst male participants were more cooperative towards female images, females were the least cooperative towards other females. Even more strongly predictive of cooperation than smiling was a rating of cooperativeness that a separate group of people had assigned to the images. In other words, people are more inclined to cooperate with people who others agree *look* cooperative.

COORDINATION

People are continually coordinating their behaviour. They manage to avoid bumping into each other (usually) on busy pavements, they behave in a fairly orderly way on underground trains, they arrange to meet friends, and so on. Airplanes take people from one place to another with very few collisions, companies provide the goods that meet customer demand, and so on. How do we manage to do these things so effectively?

'Chicken'

Game theorists have represented these kinds of problems in various games. A very well known game is called *chicken*. The game of chicken was made famous in the 1955 movie *Rebel Without a Cause*. In this film two teenage boys drive their cars towards the edge of a cliff, leaving it until the last moment before jumping out. The first person to jump out is considered 'chicken'. Subsequent movies involving tearaway teenagers showed the game of chicken as it is more often thought about; two teenagers drive cars towards each other (Poundstone, 1992). The first one to swerve away from the other is 'chicken'. Of course, you do not even need to get in a car to play chicken. Merely by turning down an invitation to play the game, you have *de facto* played the game and lost.

The game of chicken is shown in Table 14.3, where the payoffs are represented as points. The worst outcome for both is if both players drive straight, shown in the table as (−2, −2). This is considered to be mutual defection. There are two Nash equilibria in chicken, and these are shown in the top right-hand cell and bottom left-hand cell. Both involve one player driving straight and the other swerving. For example, if Player 1 drives straight then Player 2's payoff from swerving is −1. Player 2 can't improve on this by driving straight because that has a payoff of −2. Likewise, if Player 2 swerves, then Player 1 cannot improve on driving straight. Thus, this combination is an equilibrium point. The players' problem arises because there are two equilibria: how do they coordinate which of the two to settle for?

One way to play chicken is to convince your opponent that you are not going to swerve. You could do this through macho boasting, physical acts of machismo, or plain irrationality. Kahn (1965) suggested throwing bottles of whisky out of the car, wearing dark glasses to make it plain that you cannot see much, or

Table 14.3. *The game of chicken*

		Player 2	
		Swerve	**Drive straight**
Player 1	Swerve	0, 0	−1, 1
	Drive straight	1, −1	−2, −2

Table 14.4. *The free-rider problem as a chicken game*

		Player 2	
		Work hard	**Slack**
Player 1	Work hard	2, 2	1, 3
	Slack	3, 1	0, 0

throwing the steering wheel out of the car as you reach a high speed. Of course, if your opponent doesn't notice this, then you have a problem! Not for nothing has chicken been referred to as the 'prototypic dangerous game' (Colman & Wilson, 1997). Indeed, Colman and Wilson used chicken to show how evolutionary processes could give rise to antisocial personality disorder (APD). Briefly, his model showed that the payoff for a 'dangerous' strategy exceeded that for a 'cautious' strategy when the population contained a middling to high proportion of cautious players. However, whereas evolution appears to have equipped people to be both cooperative and non-cooperative in various degrees, depending on personality and circumstances, only about 2 per cent of the population exhibit APD.

By now, you may have realised that chicken is not just about driving cars. There are many situations where different parties confront each other and need to decide whether either of them is going to back down. One widely studied real-life situation is the Cuban missile crisis, briefly mentioned in the previous chapter. To recap: when it was discovered that the Soviet Union was in the process of installing nuclear missiles in Cuba, President Kennedy's military advisers wanted the USA to launch a preemptive military attack on Cuba. Kennedy's administration decided that this was too risky; however, in opting to seek a peaceful solution to the crisis they themselves risked looking weak. Therefore, the enforcement of a naval blockade around Cuba made the Soviet Union realise the need to agree a negotiated settlement, because suddenly the Kennedy administration looked tough, thereby raising the possibility that it *might* be willing to engage in military action.

This kind of brinkmanship occurs often in world affairs. Some leaders even make a public show of irrational behaviour. Of course, the danger is that neither side gives way, and escalation leads to the outcome that is worst for everybody. Few people realise just how close the Cuban missile crisis came to nuclear war. In 2002 it was revealed that a Soviet submarine had considered firing a nuclear-tipped torpedo, but one of the three officers involved in the decision prevented them from reaching the unanimity that was required to fire the missile (Dixit & Skeath, 2004, p.484).

Helping behaviour

Suppose the farmers on two neighbouring farms are putting up a fence between their properties, in order to prevent their livestock from wandering onto their neighbour's land. It is hard work, and each farmer is tempted to slack off a bit and let the other person do the bulk of the work. This situation is shown in Table 14.4. It is actually a chicken game (I have simply added two points to each payoff in Table 14.3). The temptation for one person to slack off in a task like this is known as the *free-rider problem* or *social loafing*. Actually, when there are only two people involved one person is quite likely to notice if the other is slacking off. However, the chicken game, like other kinds of game, can involve multiple players.[7] When that is the case, it is harder to spot that any one individual is slacking off.

In some situations it is probably better to use the term *diffusion of responsibility* rather than 'free-riding', because the latter implies a slightly more malign intent, whereas that is not always the case with people who fail to act. Diffusion of responsibility occurs when there is ambiguity about who should provide help to someone in need of assistance. Experimental studies have found that help is less likely to be given as the number of potential helpers increases. In one study, people were less likely to assist someone who appeared to be having an epileptic seizure when there were other people present (Darley & Latané, 1968).

It should be noted that people often do go to the help of others and normally do give assistance to someone who is screaming. There are many factors that influence helping. These include noticing that something is wrong (which can depend on the nature of the event and how much of it was seen), the social norms that apply within the context, as well as the potential costs and rewards associated with helping (see Schroeder *et al.*, 1995). However, the fragility of helping behaviour in the face of situational factors is well documented, such as in the 'good Samaritan' study reported by Darley and Batson (1973). In this study, theological students who were in a hurry frequently failed to provide assistance to someone who appeared to be in distress. In some cases, they literally stepped over the man, who was slumped in the entrance to a building. For some people, a kind of 'cognitive narrowing' appeared to have occurred: they simply hadn't noticed the man. Other participants did appear to be concerned about the man but were also worried about letting down the person they were on their way to see.

An emotion that has often been linked to helping behaviour is *empathy*, an emotional reaction involving compassion, concern, and tenderness. It has been suggested that empathy is actually a selfish emotion, being based on a merging of images of the self and other person (Cialdini *et al.*, 1997). Nonetheless, strong feelings of empathy are more likely to lead to helping behaviour (e.g. Batson, 1998; Dovidio *et al.*, 1990). However, other research indicates that empathy only motivates helping in-group members, whereas

helping out-group members is motivated by attraction. For instance, Krebs (1975) found that people experienced stronger physiological reactions and were more altruistic towards someone highly similar, as opposed to less similar, to themselves when that person showed signs of pain or pleasure during the course of a game.

There is evidence for in-group/out-group differences in the motivations for helping other people. Stürmer *et al.* (2005, Study 1) found differences in the motivations of homosexual and heterosexual HIV/AIDS volunteers who were assigned as a 'buddy/home helper' to a male client. Homosexual volunteers were more likely to view their clients as in-group members, but did not differ from heterosexuals in their perceptions of their clients' needs. For homosexual volunteers, measures of helping and time spent with the client were both predicted by empathy, but not by measures of attraction, such as liking and respect for their client, perceived friendliness of the client, and so on. For heterosexual volunteers the reverse was true: helping and time spent with the client were predicted by attraction but not by empathy. In fact, heterosexuals reported giving a higher level of practical help, on average. Heterosexuals also reported a longer duration of service, and this was predicted by attraction, not empathy. Empathy was a marginally significant predictor of service duration among homosexuals, but attraction had no predictive value.

Similar results were found in an experimental follow-up study. Participants who believed they were taking part in an internet communication study were more likely to indicate willingness to assist a partner who seemed distressed when that partner was of the same sex (in-group) as opposed to the opposite sex (out-group). For the in-group participants, helping intentions were predicted by empathy, but not by measures of interpersonal attraction, oneness, sadness, or distress. For the out-group participants, interpersonal attraction predicted stronger helping intentions. Also, higher levels of participant distress were associated with weaker helping intentions. Sadness and oneness had no predictive value.

Solving coordination problems through the use of conventions

For many repeated forms of interaction people have developed conventions that most people understand operate for the benefit of all. By internalising such conventions people are able to coordinate their activities in an effortless fashion. Consider public transport, such as buses and underground trains. Typically, you do not need to purchase a seat-specific ticket, nor do you need to enter into negotiations with other passengers as to where you can sit. The convention that everyone observes is *first come, first served* (FCFS).

In a series of studies, Milgram and his colleagues inadvertently found how deeply ingrained the FCFS convention is. For example, one study (reported in Milgram, 1977) found that New York subway commuters were often surprisingly willing to give up their seats in response to a request (68 per cent complied in one study). However, what was much harder than getting people to agree was actually mustering the courage to ask them in the first place. The students involved felt anxiety, tension, and embarrassment, and often were unable to go through with asking the question.

Likewise, a study of people's responses to queue-jumping (Milgram *et al.*, 1986) produced significant effects on the people asked to do the jumping, many of whom

> procrastinated at length, often pacing nervously near the target area, spending as much as a half hour working up the 'nerve' to intrude. For some, the anticipation of intruding was so unpleasant that physical symptoms, such as pallor and nausea, accompanied intrusions. (1986, p.686)

What these findings show is not merely that conventions exist, but that we internalise them. Indeed, the queue itself has been described as a 'social system' (Mann, 1969; Schmitt *et al.*, 1992). Evidence suggests that people's objections to queue-jumping are less to do with the costs to themselves, which might be a negligible loss of seconds or a couple of minutes, and more to do with moral outrage at the violation of norms and values on which the queue is based. As listed by Schmitt *et al.* (1992, p.806), these include egalitarianism, orderliness, principles of fairness, and duties of justice. If people's objections to intrusions into queues *were* based on the costs to themselves, then it should not matter whether a five-minute delay is caused by someone who pushes in front of you or by a service provider, such as a museum security guard, who delays you right by the entrance because he only allows in a certain number of people every five minutes. In a series of studies, Schmitt *et al.* showed that people are much more bothered by the former than the latter.[8]

SUMMARY

Human interaction is sometimes analysed using game theory. This predicts how rational people acting in their own interests would behave, assuming that the people they are interacting with are also rational and share common knowledge of the situation. Behavioural game theory is concerned with explaining how people actually behave, which is sometimes different from the predictions of analytical game theory.

Examples of 'games' analysed by researchers are the ultimatum game, the prisoner's dilemma, and the public goods game. In the latter two, there is a conflict between individual and collective rationality. Many real-life situations involve such tensions, as with the collective need to reduce fishing catches and the interest of the individual fishing boats in maximising their catch.

There is evidence that people do not always fully think through the possible consequences of their actions on social dilemmas. However, inducing them to think more (for example about other people's strategic options) does not seem to help. Indeed, one task found that people behaved more fairly when they thought less.

A considerable body of research shows that variation in responding is related to motivational factors, such as fear or greed,

and social value orientation (prosociality, individualism, or competitiveness). Some of the largest variations in responding on the ultimatum game have been observed between cultures. There is evidence that the degree of cooperativeness within a culture is related to economic success.

Trust, generosity, and communication have been shown to be important factors influencing cooperation in laboratory dilemmas. In particular, these can prevent defections occurring when responses are disrupted by noise (i.e. erroneous responses that were not intended by the actor).

Other interactions (also represented in games) involve coordination of behaviour between people. A 'prototypic dangerous game' is *chicken*, which involves a stand-off between two or more people. Each person wishes to follow a different strategy from the other person. A person may attempt to influence another by way of various signals, such as acting tough. It has been suggested that the recurrence of real-life chicken situations throughout history has caused the evolution of antisocial personality disorder.

Chicken can also be used to model situations of mutual helping behaviour, in which one or more parties are tempted to free-ride if their behaviour cannot easily be observed. In other situations, where a person in distress may need assistance, diffusion of responsibility may be a better term to use – where other potential helpers are involved, individuals are more likely to hold back from giving assistance where there is ambiguity as to who is responsible for doing so. Social psychology research has also identified various other factors that influence the incidence of helping behaviour.

In general, the existence of conventions helps people to solve a range of coordinations. This chapter has looked at the conventions that people hold in relation to seating and queuing, and has shown the depth of feeling that people have about threats to those conventions.

QUESTIONS

1. Think of a real-life situation – other than those mentioned in this chapter – that could be represented as a prisoner's dilemma or a resource dilemma. Show this in the form of a diagram.
2. Discuss the idea that people might be more cooperative in social dilemmas if they were to try and imagine things from the other party's point of view.
3. What is the weakness of Tit-For-Tat as a driver of cooperative behaviour and how can this be overcome?
4. Does it pay to always be nice?
5. Why does our memory for other people matter in relation to social dilemmas? What would it be like if we were unable to create new memories of our social interactions?
6. Design an experiment to examine the impact of in-group/out-group identification on helping behaviour.
7. What is the chicken game?

NOTES

1. 'Best' is not quite the right word, but I will address this concept more accurately when we get to it.
2. Some studies pay all participants, whereas other studies pay two or more randomly chosen Proposer–Responder pairs.
3. For more on the evolutionary implications see also Dawkins, 1976[2006], Chapter 12.
4. This will be discussed further in Chapter 15.
5. Of course, in real life we can usually just withdraw from contact with a person who 'defects'.
6. An interesting variation on the public goods game, involving competition both within and between groups, has been reported by Goren and Bornstein (2000).
7. To keep things simple I am not showing the game representations of multi-person situations, but I refer the interested reader to Dixit and Skeath (2004).
8. For a fascinating study of behaviour in very long queues, see Brady (2002).

RECOMMENDED READING

Poundstone, W. (1992). *Prisoner's dilemma*. Oxford: Oxford University Press. William Poundstone describes the development of game theory, and its application to the Cold War and to public policy.

Ridley, M. (1996). *The origins of virtue*. New York: Viking Press/Penguin Books. Matt Ridley explains how our positive moral instincts arose from evolution.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few*. London: Little, Brown. This book, recommended in Chapter 13, also contains material relevant to this chapter.