Chapter 1

The Hi-Lo Paradox

1. The Game of Hi-Lo

You and another person have to choose whether to click on A or B. If you both click on A you will both receive £100, if you both click on B you will both receive £1, and if you click on different letters you will receive nothing. What should you do?

It is obvious that the only rational choice is to click on A. Yet oddly, game theory has no explanation of *what makes* A-choices rational.

This is an example of the game of Hi-Lo. In the general case of Hi-Lo, each of *n* players chooses one item from the same finite set of alternatives without consultation. With each alternative goes a prize, and one alternative's prize is greater than all the others. If all choose the same alternative all get the prize that goes with it, and if not everyone chooses the same alternative none gets anything. Figure 1.1 shows the payoff matrix of a Hi-Lo in which there are two players, 1 and 2, and two alternatives, A and B, with associated prizes of 5 and 1.

You are to play Hi-Lo, and it is common knowledge that you and your coplayer are intelligent people. It seems quite obvious that you should choose A. However, the question why it seems obvious, and the related question of why people almost always do choose A, have turned out to be anything but easy to answer. This chapter explores these questions. The answer I shall offer has far-reaching implications for game theory and, I shall argue, for our conception of ourselves as social beings.

		Player 2	
		A	*B*
Player 1	*A*	5, 5	0, 0
	B	0, 0	1, 1

Figure 1.1. Hi-Lo

It may be thought that such situations are artificial and rare, that 'we should be so lucky!' But Hi-Los are very prevalent in our lives. It is just that the Hi-Lo structure of the payoffs is often not transparent. I first give examples of Hi-Los, of varying transparency, in varied human activities. Next I describe what is known about the choices people make. As we might expect, the evidence is that in cases in which the structure is transparent, people are overwhelmingly successful in coordinating on (A,A), and that we have great facility in Hi-Lo. It appears, too, that there is a strong intuition that choosing A is the only rational thing to do.

Example 1. Running a single. In the game of cricket two batsmen—the striker and nonstriker—stand at two 'wickets', one at each end of a twenty-two-yard strip. The striker tries to hit a ball projected towards his end of the pitch far enough that each batsman can get to the wicket at the other end before the other side retrieves the ball and strikes one of the wickets with it, in which case one of the batsmen is said to be 'run out', a serious setback for his side. If the batsmen succeed their side's score goes up by one 'run'. If they don't run, both sides' scores are unaltered. If one runs and the other doesn't, it's likely that the runner will find himself run out. If the striker hits the ball only a short distance, then even if both run, one of them is likely to be run out, but if he makes a good hit and they both run, they are likely to add one run to their side's score. Suppose the hit is good enough. Each batsman may either run or stay; so the payoff matrix has the form of figure 1.2.

Example 2. Focal coordination. As we have seen, Schelling games in option form are typically Hi-Lo. [MB is referring to material that would have been in the unwritten chapter III. The material is discussed by the editors in section 4 of the introduction.]

Example 3. Who fetches which? Lizzie and I fail to talk in the morning about who is going to fetch Julian from his school and who will fetch Emily from hers; and it's now too late to get in touch. It is common knowledge between us that I have a meeting near Emily's school, that she will be near Julian's, and that we share the objectives of fetching

		Nonstriker	
		run	*stay*
Striker	*run*	1, 1	–20, –20
	stay	–20, –20	0, 0

Figure 1.2. Running a single

		Me: fetch Emily	Me: fetch Julian
Lizzie	fetch Julian	2, 2	–2, –2
	fetch Emily	–2, –2	1, 1

Figure 1.3. Who fetches which?

		Player 2: wide	Player 2: stay
Player 1	pass	5, 5	0, 0
	shoot	0, 0	1, 1

Figure 1.4. Vision

both children on time and minimizing total time spent fetching. So we face the problem in figure 1.3.

Example 4. Vision. At a certain moment in a football match, Player 1 has the ball, and the obvious move is to shoot, but defenders are blocking his path. The chances of finding the net are therefore only slight, but if Player 2 stays where he is he may well keep possession if the ball rebounds. There is also another option. If Player 2 runs wide, a pass from Player 1 to Player 2 would allow Player 1 to find space in front of goal, Player 2 to make a return pass and Player 1 to have a good sight of goal. Player 1 and Player 2 face the problem of figure 1.4.

In Example 1 the options are obvious—run or stay. But here they are not. It takes what footballers call 'vision' for Player 1 to be aware of the pass-wide option for his side, and either 'vision' or 'telepathy' for Player 2 to see his part in the move, if there is no chance for Player 1 to attract his attention. Decision problems are often, as here, not given exogenously in their entirety but are in part made—by the creative perception of possible options by the decision-maker.

Example 5. Telling the truth. An act utilitarian is someone who in each choice situation in which she finds herself chooses with the aim of maximizing *U*, the sum of the utilities of all individuals in some reference group. If the outcome of her choice depends on the choices of certain others, the decision situation is a game, and if she and these others are all act utilitarians, and there is complete information, it is a pure

Player 1 \ Player 2	B	D
T	2, 2	–2, –2
F	–2, –2	2, 2

Figure 1.5a. Truth telling: With strong disbelief

Player 1 \ Player 2	B	D
T	2, 2	0, 0
F	–2, –2	0, 0

Figure 1.5b. Truth telling: With weak disbelief

coordination game. Hodgson (1967) argues that in deciding whether to tell the truth, act utilitarians confront a coordination game of the form of a simple Schelling game. The speaker Player 1 can tell the truth (T) or a falsehood (F); the hearer Player 2 can believe (B) or disbelieve (D). It is assumed that U is high (2, say) if the truth is believed and low (-2, say) if a falsehood is believed. In some cases believing the truth and disbelieving the falsehood have exactly the same effects (they do if the truth is a proposition P, the falsehood is $\neg P$, and disbelieving means believing the negation), and so do believing the falsehood and disbelieving the truth, so the payoffs are those of figure 1.5a. Hodgson argues that act utilitarians are left not knowing whether or not to tell the truth, and rejects this form of utilitarianism because of this indeterminacy. Gauthier (1975) notes, however, that if a disbeliever merely fails to form a belief at all, rather than believing the opposite of what is asserted, then the outcome of (F,D) is worse than that of (T,B). In this case the matrix is that of figure 1.5b, a (weak) Hi-Lo. On Gauthier's interpretation of disbelief, if A is the unique rational solution of Hi-Lo then act utilitarianism gives clear advice about truth telling. As to whether A is the unique rational solution, that is the topic of this chapter.

Example 6. Calling a catch. A second Hi-Lo drawn from cricket illustrates a fundamental form of organization. One basic function of organization is to amplify the information on which individual actions are based. This allows a profile of actions that is better geared to the state of the world than any of those attainable by the players without organization. If the procedural cost of the organization is not too great, the enlarged game in which players have the option of conforming to this profile is a Hi-Lo in which conforming to it is A.

When a batsman hits the ball in the air, if a fielder catches it before it touches the ground the batsman is 'caught out', which gives a significant boost to the fielding side (a payoff of 20). If the ball is 'skied', catching it means tracking its flight while running towards it. This is a difficult task which absorbs all the fielder's attention. If two fielders (F1 and F2) are near and both go for the catch there may be a collision, and the likely upshot of this is that the batsmen score two runs or so, and a fielder may be injured (a payoff of −5). If neither goes for the ball the likely upshot is two runs or so (a payoff of −2).

Typically each of the fielders will have some idea of whether he or the other is better placed to make the catch, but can be wrong. We model this by supposing that there are two equiprobable states of the world, ω_1 and ω_2. In ω_1, F1 is better placed and will make the catch with high probability (1.0) if he is unhindered, and F2 would make it with a lower probability (0.75). In ω_2 the players' positions are reversed. Whatever the state, each player gets right which state it is with probability 2/3 and gets it wrong with probability 1/3.

To begin with, I suppose that each fielder has two options: catch if you think yourself better-placed (B), and catch if you think yourself worse-placed (W). This is a simple Hi-Lo, in which each of these is in Nash equilibrium with itself (figure 1.6a). (The mathematical derivation of the payoffs is given in the appendix to this chapter.)

Some cricketing teams use the following calling routine. The captain calls the name of the fielder he thinks has the better chance, and the captain is always right about this. A fielder therefore has a third option: obey the call (C). But C carries a procedural cost: attending to the call takes precious seconds and reduces the chance of success if you do end up going for the catch—if unhindered the better-placed has probability of 0.5 and the worse-placed 0.375. The upshot is

		F2	
		B	*W*
F1	*B*	9, 9	5.8, 5.8
	W	5.8, 5.8	7.3, 7.3

Figure 1.6a. Calling a catch: Without organization

		Player 2 C	B	W
	C	10, 10	8.8, 8.8	2.7, 2.7
Player 1	B	8.8, 8.8	9, 9	5.8, 5.8
	W	2.7, 2.7	5.8, 5.8	7.3, 7.3

Figure 1.6b. Calling a catch: With organization

shown in figure 1.6b. (Again, the mathematical derivation of the payoffs is given in the appendix to this chapter.)

The situation is a Hi-Lo in which (B,B) and (W,W) are still equilibria, but better even than (B,B) is (C,C): the extra information more than compensates for the procedural cost of (C,C).

Example 7. Deciding together. Suppose that two people have exactly the same aims, and have the opportunity to decide together what to do. Suppose these common aims are all the aims either has that could be affected by their decision. Suppose the problem they jointly face is a matching game with options A, B and C. For example, a colleague and I might have the opportunity to decide together whether to take a visiting speaker to lunch at restaurant A, B or C, and are concerned only to eat well, within budget and in time for the seminar. Deciding together is exchanging information, weighing considerations, agreeing on the judgement that a certain alternative is best and deciding to adopt that alternative. If the two people decide together, then at the stage of this process at which they reach agreement on a judgement they face a Hi-Lo.

Example 8. Calling a run. Example 1 was about the situation in cricket in which it is clear to both batsmen that a single run is 'on'. But in other cases this may require judgement, and the two batsmen may not be equally well placed to make it. Some cricketers use the following practice for deciding whether to run. If the striker hits the ball 'forward of square', so that it subtends an angle less than 90 degrees with the line of the pitch, he calls Yes or No; if he hits the ball backward of square, the nonstriker calls. Both run if and only if the call is Yes. Call this the Variable Caller practice. It puts the decision in the hands of the batsman better placed to make a good judgement quickly. To conform with the Variable Caller practice, a batsman implements the following strategy (V): if you hit the ball forward of square, or your partner hits it backward of square, call Yes if you think a run is on, otherwise No; if you hit it backward, or your partner hits it forward, do not call; obey your partner's call; if he fails to call, stay. Suppose, simplifying, that to their common knowledge the two options each batsman considers are Run if you think there's time, and Variable Caller (V). Then the situation is a Hi-Lo in which (V,V) is the efficient pair. This example resembles Example 6 except that the judgement controlling the basic action each player should take—taking a catch or staying in Example 6, here running or staying—was there made by a third party and here made by one of the players themselves. It is an example of a self-organizing team, and an archetype of a very large class of organizations.

Example 9. Unlocking. Figure 1.7 shows a situation of a type that was common in the town in which I live during the 'manhole years' that followed the privatization of public utilities and the laying of cable. I myself was in car 3. In the figure the diagonal arrows represent the direction signals cars 1, 2 and 4 were making; the horizontal arrows mean that cars 3 and 5 were making no direction signals. The five cars all came to a halt, everyone in view of everyone. A few seconds later we started to move: first, car 2 turned right into Abingdon Road, car 1 remaining stationary; next, car 4 turned right into Whitehouse Road

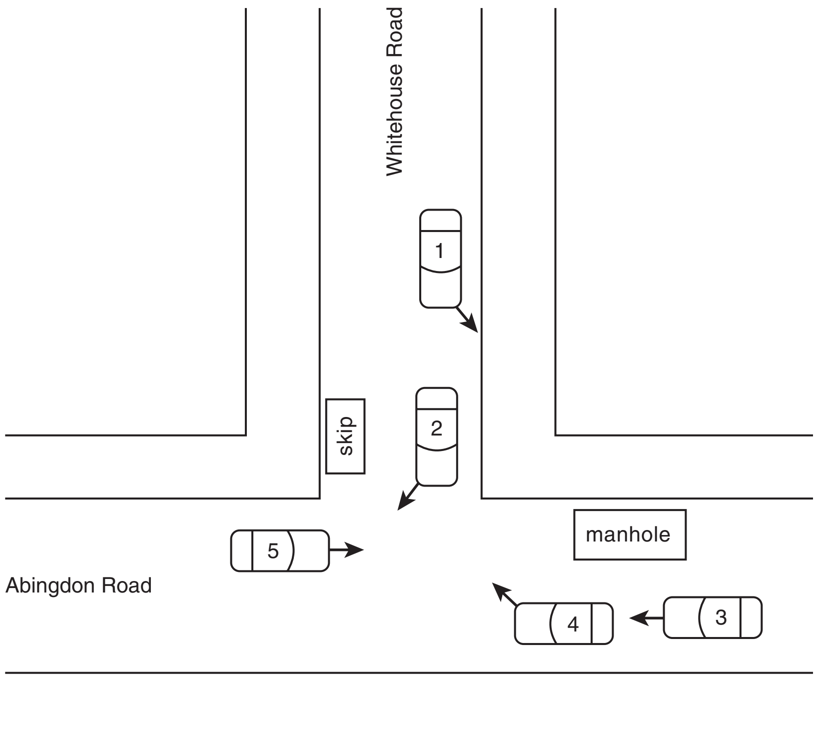

Figure 1.7. Unlocking

and threaded between the skip (dumpster) and car 1; next, car 3, my own, went straight on down Abingdon Road; next, car 5 went straight on up it; finally, car 1 passed the skip and turned left. Assuming that everyone wanted to be on his or her way in as little time as possible, this traffic unlocking problem is a Hi-Lo. The unlocking process that took place took perhaps 45 seconds. There were other possible combinations of movements, one involving car 4 forcing cars 1 and 2 to reverse. All were likely to take longer for each of the drivers. The five of us solved the Hi-Lo in a matter of seconds.

These examples begin to show the importance of situations of Hi-Lo form in our lives. They arise whenever there is a common purpose and a best method of furthering it, and mixing methods is worst of all. Game theory has made us subtly sensitive to conflicts of private motive, not least when people subscribe publicly to a common aim; but this should not blind us to the huge domain of decisions we make in which aims are exactly or approximately the same. This does not require that they be stably so, only that there are occurrent frames in which they are. And this is normally the case when we carry out the normal activities of members of teams and other organizations to which we belong. Within the big game in which we may contemplate our department or firm or family as an entity outside ourselves, with whose goals our own purposes may conflict, we daily play subgames in which we frame ourselves as functioning members of some such system. When we do we are usually playing Hi-Lo.

2. The Data

There are these two broad empirical facts about Hi-Lo games: people almost always choose A, and people with common knowledge of each other's rationality think it obviously rational to play A. Call these the *behavioural fact* and the *judgemental fact* about Hi-Lo.

In cricket, when it is clear that there is a safe run to be taken, and the payoffs are therefore as in figure 1.2,[1] it is unheard of not to take the run. In Schelling games, the only formal explanation we have of the coordination success we observe—variable frame theory—depends on the assumption that people play A in the induced Hi-Lo. In these examples, people appear to be disposed to make A-choices even though the Hi-Lo element in their problem may not be transparent. Their problem is not displayed to them as a Hi-Lo, but we can still explain their behaviour as resulting from A-choices in an implicit Hi-Lo; moreover, if they do implicitly confront Hi-Los, their behaviour entails that they make A-choices in them. In the Hi-Los which emerge when we decide together which alternative is best, as in Example 7, to agree, genuinely, that restaurant A is best and yet not to decide to go there seems cockeyed. It seems as absurd as concluding by yourself that a certain alternative would be best for all the aims that weigh with you, and then not choosing that alternative.

When it comes to laboratory tasks or text-book examples in which the form is transparent, the evidence that people are inclined to choose A is extremely commanding. Consider a Hi-Lo game 1, presented to two players for money prizes of £10 and £1. It seems so obvious that everyone will choose A that there were apparently no experimental tests before 2001 of the hypothesis that people choose A in overt Hi-Lo games. Then, in the course of an experimental programme to which I shall return, Gerardo Guerra and I asked each subject in a pair to choose one of the cards in a display (figure 1.8), in which the prize for matching on each of the items is shown below it. Although the differences among prizes were here quite small, fifty-eight subjects out of sixty-four, or 91 per cent, chose the £6 card.[2]

Most of us appear to have an overwhelming intuition that it is rational to play A, and moreover that this is obvious. The conviction seems to be independent of how rational we think our coplayer: in particular, it

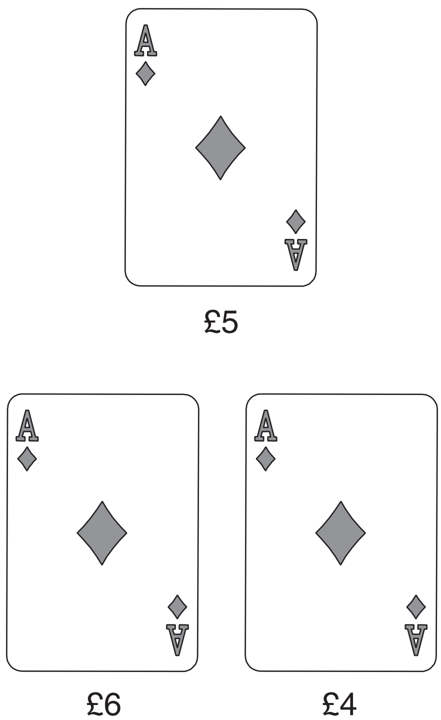

Figure 1.8. Choose a card

applies just as much when we think our coplayer highly rational as when we think he is simple-minded. Moreover, it applies just as much when there is common knowledge of the rationality of us both. This is what I have dubbed the judgemental fact about Hi-Lo. The judgemental fact is illustrated by the widespread endorsement of the intuition by theorists considering play between symmetrically rational players. Several assert or imply that A is obviously rational (Lewis 1969; Gauthier 1975; Farrell 1988; Harsanyi and Selten 1988; Anderlini 1999; Crawford and Haller 1990; Bacharach 1991; Sugden 1995; Janssen 2001a).[3]

One might expect the explanation of these two facts about Hi-Lo, the behavioural and the judgemental fact, to be straightforward. But these facts prove to be mysterious, and their explanation anything but that.

3. The Hi-Lo Paradox

3.1 What Game Theory Predicts

The basic solution concept of standard game theory, Nash equilibrium, delivers two solutions to Hi-Lo: (A,A) and (B,B). (B,B) is just as good a solution as (A,A), for it satisfies just as well the logic of Nash equilibrium, namely that beliefs plus instrumental rationality yield facts which match the beliefs. *If* I think you will choose B then the act it's best for me to choose is B, and *if* you think I will choose B then it's best for you to choose B, so both our hypothesized expectations are correct. (B,B) expectations get confirmed by best-replying agents.

Because (A,A) and (B,B) are both Nash equilibria, the Nash equilibrium solution concept predicts little or nothing. Whether little or nothing depends on how we interpret the claim that these equilibria are solutions. The strong form of the claim is that there is some equilibrium that all rational players (who are commonly known to be rational) will play. The weak form is that every such player will play her component in some equilibrium. The strong form predicts little, and the weak form nothing. In either form, game theory fails to explain why everyone chooses A.

3.2 The Paradox

The predictive failure is bad news for game theory in two different ways. First, because game theory aspires to explain actual behaviour, and here fails to explain a particularly strong regularity of behaviour. But second, because game theory is a theory of rational choice and so should deliver solutions which accord with, not actual behaviour, but another sort of empirical data, namely our intuitions about what is rational. To be sure, our intuitions are both partial and unreliable evidence about the rationality of an action. We should not expect people to be able to identify the reasoning principles that govern their conclusions, even when these principles are sound: for example, most people easily and reliably reason in accordance with *modus ponens*, but almost no-one can tell you that it is *modus ponens* that sanctions their conclusions. We know something makes sense (playing A, par excellence); but we do not know what sense it makes. And our intuitions may not be reliable: we may be deceived in finding some theoretical or practical proposition persuasive; one sort of evidence for this is that on reflection it loses its initial aura of obvious rightness. But the more stable under reflection our intuitions are, the stronger they are, and the more they are shared by serious thinkers about the subject matter, the greater is the challenge to theory to encompass them.

A paradox is a mismatch between such high-quality intuitions and the deliverances of an accepted theory. The clash between the obvious rationality of the choice A and the inability of game theory to single out A is a paradox. It is a weak paradox, because game theory does not predict that A will not be chosen. It is a paradox nonetheless, a failure of theory to agree with intuition.

3.3 Refining Equilibrium

Such paradoxes are not uncommon in game theory. Some have been dealt with in the past by 'refinement' of the Nash equilibrium solution concept. The broad object of the equilibrium refinement programme of the 1970s and 1980s was to eliminate equilibria which were intuitively suspect, by adding to the Nash solution principle further rationality conditions on players' choices of a decision-theoretic character. To eliminate in this way a suspect Nash equilibrium (A,X), we seek a decision-theoretic argument that A is not, on closer examination, or on a more careful specification of the game, a rational choice against the choice X.

A simple example is the elimination of Nash equilibria in which one act is weakly dominated. Suppose the payoffs are as in figure 1.9.

		Player 2	
		X	*Y*
Player 1	*A*	1, 6	1, 3
	B	1, 4	2, 5

Figure 1.9. A game with a weakly dominated Nash equilibrium

(A,X) is a Nash equilibrium (A is *a* best reply to X), but A is weakly dominated by B since B is at least as good as A against all feasible acts of Player 2, and better against one, Y. Now it is argued (Selten 1975) that even rational players, having made a choice, may with nonzero probability do something else by accident. In the present case, consider a state in which Player 1 is certain that Player 2 will choose X and Player 2 is certain that Player 1 will choose A. Then Player 2 will indeed choose X, but there is some probability ε that her 'hand will tremble' and she will *play* Y. This means that the expected utility of A for Player 1 is lower than that of B, so it is not rational for Player 1 to choose A. The weak-dominance pattern in A's payoffs means that this is true even for infinitesimal ε. Selten concludes that to be a rational solution, a Nash equilibrium must not include a weakly-dominated act.

Neither this requirement nor any other refinement which has commanded general acceptance has the effect of eliminating (B,B) in Hi-Lo. Two refinements proposed by Harsanyi and Selten (1988), called *risk dominance* and *payoff dominance*, do eliminate (B,B). As we shall see in a minute, however, risk dominance is suited to describe the choices of players whose rationality is bounded in a certain way, but not to players with common knowledge of full rationality. Payoff dominance is another matter. Its logic is closely related to the main thesis of this book. I shall argue that in some circumstances the outcomes it selects will tend to be generated by the deliberations of fully rational agents. But the rational considerations that generate it are quite different from the decision-theoretic ones that the equilibrium refinement school was searching for. The conclusion: the Nash solution principle implemented using the most stringent version of standard rationality fails to deliver A. (B,B) is not only an equilibrium; it is a perfectly good equilibrium. But A is obviously the unique rational choice: this is the paradox.

It may be protested that there is no real paradox here because the game-theoretic argument for B is flawed. It first supposes that each player expects B, then proceeds to show that these expectations are part of a consistent constellation of beliefs and acts. But it does nothing to show where the expectation might come from in the first place. Indeed, such an expectation is intrinsically implausible, so it postulates the intrinsically implausible. How could anyone expect a rational player to play B?

Indeed it seems that no-one could. But the problem is to show why—why rationality excludes B.

It must be understood that game theory does not predict B; it merely fails to exclude B. There is nothing fallacious about that, only something incomplete. The programme game theory sets itself is to plot all outcomes *consistent with* all its postulates about the players' reasoning and knowledge. Ultimately, the reason why (B,B) is a solution is that it is consistent with all the facts about rationality that game theory can muster. (B,B) is a solution because game theory has mustered no fact about rationality that excludes B. Given this absence, the rigorous project of game theory obliges it to allow belief in B, and then B as a response to this belief. Given the absence, so far from criticizing game theory for allowing B and the equilibrium (B,B), we should praise it for its clear-headed acknowledgement of the consequences of its own limits. We should only regret that our current formal theory of interactive behaviour has a serious gap, no response to the question 'What's wrong with B?' or, equivalently, 'What's right with A?'

4. The Response

Since game theory fails to explain why people would choose A in Hi-Lo, any explanation of A-choices must say they are not the result of game-theoretic reasoning in Hi-Lo, or at least not of this alone. Game theorists, philosophers, psychologists and others have sought to explain A-choices. The theories advanced have fallen into the same two broad categories as theories that have tried to explain C-choices in the Prisoner's Dilemma: respecification theories and bounded rationality theories. [Throughout the text MB uses the standard notation of 'C' and 'D' for the strategies "Cooperate" and "Defect" in the Prisoner's Dilemma.] Respecification theories explain behaviour usually thought of as an A-choice in a Hi-Lo by saying that the chooser is not in fact playing Hi-Lo but some related game G in which game theory does predict A. Bounded rationality theories explain A-choices in terms of limits on or lapses in rationality. In the Prisoner's Dilemma literature we find respecification theories in which G is, for example, an indefinite repetition of the Prisoner's Dilemma, or a Prisoner's Dilemma with transformed payoffs, and bounded rationality theories in which players have limited depth or use magical reasoning.

There is a third sort of theory, which may be called revisionary; revisionary theories aspire to explain the target behaviour (here, A) as rational, but rewrite accepted doctrines of what is rational. Revisionary theories may add new principles of rational choice, like equilibrium selection theories, or extend the framework by adding new primitives, like variable frame theory, or challenge previously accepted canons, as some evidentialists deny the instrumentalist canons. The boundary between revisionary theories and bounded theories may be contentious even in the first two cases, for the extension championed by the revisionist may be seen by others as involving invalid reasoning. The theory advanced in this book is a revisionary theory which is not immune from the possibility of such a reaction.

5. Respecification Theories

Farrell (1988) and Anderlini (1999) remodel the game in which A is chosen as a game in which players can communicate with each other before choosing between A and B. Anderlini suggests that our intuition that it's rational to play A is due to an implicit assumption that players will manage to communicate their intentions. Farrell introduces a new rationality postulate (a refinement, called 'neologism-proofness') which applies in such games. The key idea is that some claims are 'self-signalling'—the speaker would like them to be believed if and only if they're true; Farrell postulates that it's rational to believe a self-signalling claim. Clearly both 'I'll play A' and 'I'll play B' are self-signalling. Intuitively, in Hi-Lo this makes it rational for the sender to send the message 'A'.

Aumann and Sorin's (1989) respecification theory is actually not for Hi-Lo but for a broader class of games, common-interest games, which includes Hi-Lo games as a subclass. A common-interest game is any in which there is a profile that is Pareto-superior to all others. A common-interest game need not be a coordination game. An example is Rousseau's Stag Hunt, shown in figure 1.10. [The story behind the game is that the two players are individuals living in a presocial 'state of nature'. Each chooses independently whether to hunt for rabbits (R) or deer (S, for 'stag'). Deer hunting requires concerted action by both individuals, while either can hunt rabbits on his own. Both individuals do better by hunting deer together than by hunting rabbits separately, but hunting deer alone is the least productive activity of all.]

		Player 2: *S*	Player 2: *R*
Player 1	*S*	2, 2	–1, 1
	R	1, –1	1, 1

Figure 1.10. The Stag Hunt

If a common-interest game is a coordination game and symmetric, it is a Hi-Lo, and (A,A) is the unique Pareto-optimal profile of the definition. If it can be shown that in common-interest games in general the Pareto-optimal profile is played, it is shown a fortiori that (A,A) is played in Hi-Lo. Aumann and Sorin consider an arbitrary common-interest game and suppose that it is repeated indefinitely, and that with some chance each player has 'bounded recall'—she remembers past play only *n* rounds back. They show conditions under which, in this setting, the probability that a player plays A tends to 1 over time.

Both these models explain A-choices in a Hi-Lo game by interpreting this game as the kernel of a larger game. The phenomena they study are of intrinsic interest, but they only help explain the A-choices we are interested in (and the feelings people have about them) if the situations in which these A-choices are made really do feature communication or repetition. Looking again at the examples of section 1 it is clear that most of them do not actually have these features. In particular, one-shot laboratory experiments with subjects behind screens certainly do not. Is it not possible, though, that even if they do not, nonetheless the players treat them as if they did in the sense that they use heuristics appropriate to cases in which they do, perhaps because they are used to such cases? But this 'assimilation' hypothesis is of no help in the repetition case, because (A,A) is predicted only in the nonexistent long run; the model does not predict that the prospect of this long run encourages the use of A now.

In the communication case assimilation cannot be excluded out of hand. It is not inconceivable that, in deliberating in some Hi-Los, a player asks herself what message her coplayer would send if she were able, and best-replies to this counterfactual, imagined message. This hypothesis, however, is not advanced by Farrell and Anderlini themselves. It is also empirically implausible in some paradigm cases, and fails to satisfy the generality aim. In some Hi-Los, such as those that arise in fast-moving games, quick decisions are needed, and the roundaboutness of the reasoning on this imagined-message hypothesis makes it implausible. It doesn't ring true that a footballer wonders, even implicitly, what his teammate would signal in order to further his interest in winning the game if he had the chance to signal. And in cases in which one of the actions in the efficient profile is itself a message, as in Calling a Run, and in the vast array of organizational Hi-Los of which this is the archetype, the hypothesis means that players have thoughts about quite bizarre matters: the receiver deliberates that the sender would, if she had the chance, send a message saying that she was going to send a message, an amusing but improbable possibility.

6. Bounded Rationality Theories

In the literature that seeks to explain the behavioural fact—A-choices—we find two kinds of cognitive shortcomings: magical thinking and depth limits. I first describe what these are, then how they might explain A.

6.1 Magical Thinking

An agent's reasoning is called *magical* (Elster 1989, pp. 195–202) if it fallaciously attributes causal powers to her own decisions. Decision theory and common sense both hold that one sort of reason that can rationalize an action x is a *causal* reason, namely that doing x would *bring about* a better end-state than its alternatives would.[4] Someone who thinks magically claims to have reasons of this kind, and this much of her thinking is sound; her mistake is in her view of what causes what.

The specific mistake she makes in so-called evidential or diagnostic reasoning comes from the fact that *what she decides* is good evidence for something else. The decision-maker now slides from the idea that her choice is a good sign of the other's choice to the idea that it is a cause of it. Very often, perhaps typically, in our decision-making lives, what makes a choice x be evidence for an outcome o is precisely that x tends to cause o, and it seems likely that it is the normality of this case that lures us into the slide. But although it's normal it's far from guaranteed, so sliding is unsafe. Consider the case of smoking and cancer. Suppose you hear on the news that the conditional probability of contracting cancer is higher given that you smoke than it is given that you give up smoking: $P(\text{cancer} \mid \text{smoker}) > P(\text{cancer} \mid \neg\text{smoker})$. It is natural to conclude that giving up is a good idea, and what lies behind this is the presumption that the statistical relationship comes from a causal link from smoking to getting cancer. But this presumption might be false: it might be that the relationship comes from a common cause, for example, that the tendency to smoke and susceptibility to cancer are both expressions of a certain gene.

In the game case a player's choice of x can be evidence that another player will choose y. But games are by definition situations in which different players' acts are causally independent. This part of the definition of a game is due to the fact that game theorists want to model decisions made without physical communication, and assume that there is no such thing as telepathy. So if Player 1 believes, rightly or wrongly, that her choosing x is evidence that Player 2 will choose y, believing that this evidence is due to a causal link must be a mistake.

There is more than one scenario in which Player 1's choice can be evidence of Player 2's. The simplest is the case of similar players. If

Player 1 knows that she and Player 2 are 'like-reasoning', she can argue: 'Whatever choice x I make, Player 2 will make the same one; so I should choose x so that (x,x) is as good as possible.' This argument 'masks off' nonsymmetric pairs of choices and leaves Player 1 just scanning the main diagonal. In the Prisoner's Dilemma it leads a player to choose C, and in Hi-Lo to choose A. It is fallacious, because choosing C will make no difference to whether Player 2 chooses C or D (Lewis 1979). What is true is that if at the end of deliberation Player 1 has chosen C, it is probable that Player 2 will have too. But this fact may not be legitimately used by Player 1 during a deliberation not yet completed.

Bacharach and Colman (1997) propose a way in which Player 1 can see her choice as evidence of Player 2's in which Player 1's model of Player 2 is slightly richer, containing a best-reply step. Player 1 argues: 'Whatever I choose, Player 2 will "read my mind", and then do the best thing for himself given my choice.' This is a masking move which lets Player 1 confine her attention not to diagonal points (x,x) but to the points $(x, B(x))$, where $B(x)$ is Player 2's best reply to x, which is assumed unique. Colman and Bacharach suggest that a player goes on with the piece of magical thinking: 'So I should choose the x that makes the pair $(x, B(x))$ as good as possible'. The choice which it delivers is called the Stackelberg act. But this is not the end of the story. The Stackelberg act, say x^*, need not be a best reply against the expected act of the opponent $B(x^*)$—that is, $(x^*, B(x^*))$ need not be an equilibrium. It seems likely that in these cases the player may reject the Stackelberg act on reflection.[5] Colman and Bacharach conjecture that a Stackelberg act will tend to be chosen when but only when it has the equilibrium property. Colman and Stirk (1998) found evidence of this in a set of twelve 2×2 games, nine having and three not having the Stackelberg equilibrium property.

In Hi-Lo, A is the Stackelberg act and has the equilibrium property, so both these magical reasoning theories, the similar-player theory and the Stackelberg theory, explain the behavioural fact.[6] Both are bounded rationality theories, because magical reasoning involves a mistake.

Another explanation of A-choice, by Jacobsen (2001b), subtle and imaginative, claims to explain the choice as rational, but also seems to involve a hidden evidentialist element. In the Jacobsen theory each player selects a 'plan'. Player 1 selects the pair (x,y); x is her choice in her actual role (as Player 1), y is her choice for the act she would perform if she were to be Player 2. This seems to be another evidentialist theory. Think of me, choosing between plan (x,y) and (x',y'). According to Jacobsen, by the axiom Janssen (2001a) calls Internal Consistency, if I choose (x',y') rather than (x,y) this implies that my expectation of your act is y' rather than y. The reason given is that my expectation of what you will do depends on the corresponding part of my plan, that is, on what I decide I would do if in the other role. It is indeed reasonable for me to take this as evidence about what you will do. But since I cannot causally affect what you will do, I cannot use this epistemic consequence of my deciding on y' rather than y as a reason to change from x to x'. My changing from y to y', though evidence that you will do y' and evidence against the event that you will do y, does nothing to bring it about that you will do y' and not do y.

6.2 Depth Limits

An old idea for explaining A-choices is 'equiprobability'. The thought is that if Player 1 has no idea what Player 2 will do, then it's better for Player 1 to do A. For if she has no idea, then her personal probabilities for Player 2's doing A and B must be equal, and then A gives her a higher expected payoff (2.5 instead of 0.5 in the figure 1.1 example). This model is *asymmetric*: Player 1 treats Player 2 as unlike herself. For if Player 2 were like her in terms of his initial information and reasoning powers, he would presumably go through the same reasoning and would also choose A. So Player 1 cannot halfway rationally both think Player 2 is just like her and have no idea what Player 2 will do.

Harsanyi and Selten's principle of risk dominance turns out to boil down to the equiprobability principle—and so in my view to be inappropriate as an unbounded rationality solution concept. So I shall classify it as a bounded rationality explanation of A, even though that may not be how Harsanyi and Selten intended it. Harsanyi and Selten's general theory of rational solutions of games adopts the postulate that solutions must be Nash equilibria, then seeks 'refinements' of Nash equilibrium to deal with cases when there is more than one unrefined Nash equilibrium. Whereas the equiprobability argument is essentially decision-theoretic and selects an act directly, the risk-dominance theory first selects an equilibrium and then predicts that each player will play her part in it. For the class of symmetric 2×2 games, which includes Hi-Lo, Stag Hunt and the Prisoner's Dilemma, the risk-dominance principle is that an equilibrium is a solution only if it is not 'risk-dominated'; where, if E and E' are any two equilibria, *E risk-dominates E'* just if adherents of E do better than adherents of E' against coplayers equally likely to be adherents of either. If there are two equilibria with no acts in common, like (A,A) and (B,B) in Hi-Lo, or (S,S) and (R,R) in Stag Hunt, the risk-dominance principle gives the same act-choices as the equiprobability argument.[7] In Hi-Lo, risk dominance yields A.

Harsanyi and Selten themselves say little about why we should expect people to adhere to E only if doing so gives higher payoffs against an equal-probability mix of E and E' adherents. In evolutionary game theory, risk-dominance ideas arise naturally in studying the emergence of one equilibrium rather than another; the processes studied involve a hard-wired E population invaded by random waves of hard-wired E'-adherents, and what counts is the relative objective success of E-adherence and E'-adherence against the resulting probability mix of E and E'.

But here we are interested not in hard-wired but in deliberating agents. The probabilities in the mix are probabilities in the head of Player 1, specifically elements of Player 1's model of Player 2.

We can see better what is involved in the equiprobability theory by casting it in the framework of Stahl and Wilson's 'level n' theory of games (1995). This is a bounded rationality theory of how real players play games quite generally. In it, players reason strategically at different 'levels'. A level 0 player has no model of her coplayer; level 1 players believe they are playing level 0 players, and maximize expected payoff on this belief; level 2 players think a coplayer is level 0 or level 1 with probabilities adding to 1, and maximize expected payoff; and so on. Further, level 0 players are assumed to pick an option at random. Thus level 1 players of Hi-Lo are led to play A by what is precisely the logic of the equiprobability model. The equiprobability theory interpreted thus as an application of level 0 theory is a bounded rationality theory par excellence. A player models her coplayer as bounded—ultra-bounded, because entirely devoid of strategic reasoning; and herself thinks that half the population plays B, a grotesque belief considering that virtually 100 per cent play A. Other empirical evidence also goes strongly against this theory. Stahl and Wilson (1994) found experimentally that over a range of ten games the fraction of subjects who were level 0 was insignificant and the fraction who were level 1 was 24 per cent.[8]

7. Salience Theories

7.1 Applying Variable Frame Theory to Hi-Lo

One idea for explaining why people play A—the behavioural fact—is that it is due to the salience of A.[9] Schelling thought that, however it may be that a particular coordination equilibrium's salience makes it get played, in Hi-Lo this mechanism gets (A,A) played because giving the highest prize makes (A,A) salient. Probably the commonest response of game theorists today to the question why people choose A is an equilibrium-selection version of this idea: the equilibrium (A,A) gets selected because it is the salient equilibrium. This was Harsanyi and Selten's opinion. They point out that each player prefers (A,A) to (B,B),[10] and conclude that we should adopt (A,A) as the solution because the players 'should not have any trouble in coordinating their expectations' at (A,A). (Harsanyi and Selten 1988, p. 81). The suggestion of these theorists is, I think, that salience can explain A-choices by itself, that is, without our having to call in other theories of A-choice, such as bounded rationality, respecification, or payoff dominance. In this section I investigate this suggestion.

The only nonhandwaving theory we have of how rational players solve coordination problems is variable frame theory [described in section 4 of the introduction]. Applied to Schelling games, it works by using salience characteristics to turn these into Hi-Lo games. Could variable frame theory perhaps be reapplied to Hi-Lo games to show that salience characteristics of A and B explain A-choices in them? This might give a nonhandwaving theory of A-choice as an effect of framing.

[It turns out that, provided we allow there to be a family of predicates which includes *giving the highest prize* but not *giving the lowest prize*, we can model the asymmetric salience of A and B in variable frame theory, and that for interesting bands of parameter values we get unique variable frame equilibria. If *giving the highest prize* is highly salient, there is a unique equilibrium in which people almost always play A. To be specific, consider a Hi-Lo game with k basic acts, A, B, ... Let the payoffs be a in (A,A), b in (B,B), . . . , with $a > b > \ldots$. Let the universal frame be $F = \{F_0, F_1\}$ where F_0 is the generic family {*thing*} and F_1 = {*highest prize*}; for each player, the extension of *highest prize* in the set of act-descriptions is A. Assume that the availabilities of these families are such that $v(F_0) = 1$ and $v(F_1) > 0$. Then there is a unique variable frame equilibrium in which players for whom F_1 comes to mind opt for *choose the highest prize* while other players opt for *pick a thing*. This implies that an arbitrary player plays A with probability $v(F_1) + (1 - v[F_1])/k$.]

In this theory, the mechanism through which being highest-prized has an effect on choice is that it makes A salient—A is chosen in virtue of this salience. (Coordination on it has to be reasonably well paying, but only above average, not necessarily highest.) I shall call this mechanism for generating A-choices the *pure salience mechanism*. The theory has a perhaps surprising implication: if an act giving a lower prize were for some reason or other highly salient, the low-prize act would be chosen rather than A.

[This can be shown in a simple extension of the previous model. Suppose that each player makes his choice between the options A, B, . . . by pressing one of k buttons. The A button is salient through its association with the highest prize, but another button, R, is salient because it is the only red one. Let z be the average of the prizes a, b, . . . and assume that r, the prize associated with R, is greater than z (which is possible only if $k > 2$). The reason for this assumption will become clear shortly. Let the universal frame be $F = \{F_0,F_1,F_2\}$, where $F_0 = \{thing\}$, $F_1 = \{highest\ prize\}$ and $F_2 = \{red\}$. Assume $v(F_0) = 1$, $v(F_1) > 0$ and $v(F_2) > 0$.

First, consider a player whose frame is $\{F_0\}$. For her, the only act-description is *pick a thing*; thus she picks among the k buttons at random. Next, consider a player whose frame is $\{F_0,F_1\}$, and thus whose decision problem is {*pick a thing*, *choose the highest prize*}. In the game as he perceives it, *choose the highest prize* is the unique best reply to both of these act-descriptions. Now, consider a player whose frame is $\{F_0,F_2\}$, and thus whose decision problem is {*pick a thing*, *choose the red*}. In the game as this player perceives it, *choose the red* is the unique best reply to both of the act-descriptions he can recognise. (To reach this conclusion, we need the assumption that $r > z$.)

Finally, consider a player whose frame is $\{F_0,F_1,F_2\}$, and thus whose decision problem is {*pick a thing*, *choose the highest prize*, *choose the red*}. As in the case of the frame $\{F_0,F_1\}$, *pick a thing* is strictly dominated by *choose the highest prize*. (Similarly, as in the case of the frame $\{F_0,F_2\}$, *pick a thing* is strictly dominated by *choose the red*.) So the only act-descriptions that can be optimal for this player are *choose the highest prize* and *choose the red*. Recall that players whose frame is$\{F_0,F_2\}$ opt for *choose the red*. Since the unique best reply to *choose the red* is *choose the red*, if the probability that an opponent has the frame $\{F_0,F_2\}$ is sufficiently high, *choose the red* is optimal for the player with the frame $\{F_0,F_1,F_2\}$. More specifically, let v_T ('T' for 'thing') be the probability that a player has the frame $\{F_0\}$; let v_H (for 'highest') be the probability that a player has the frame $\{F_0,F_1\}$; let v_R (for 'red') be the probability that a player has the frame $\{F_0,F_2\}$; and let v_U (for 'universal') be the probability that a player has the frame $\{F_0,F_1,F_2\}$. That is, $v_T = (1 - v[F_1])(1 - v[F_2])$, $v_H = v(F_1)(1 - v[F_2])$, $v_R = v(F_2)(1 - v[F_1])$, and $v_U = v(F_1)v(F_2)$. Then it is straightforward to calculate that *choose the red* is unambiguously optimal for a player with the universal frame if[11]

$$(v_T/k + v_R)/(v_T/k + v_H + v_U) > a/r. \quad (1)$$

That is, if (1) holds, a player with the universal frame maximises her expected payoff by opting for *choose the red*, irrespective of her beliefs about the behaviour of other players with the same frame as herself. If this is the case, the probability that an arbitrary player presses the red button is $v_T/k + v_R + v_U$, while the corresponding probability for the A button is $v_T/k + v_H + v_U$.

This result is only illustrative: it shows sufficient conditions for there to be a unique variable frame equilibrium with a clear prediction of R, assuming a very stripped-down set of possible frames.[12] In order for (1) to hold, the availability of the 'red' family F_2 must be sufficiently greater than that of the 'best payoff' family F_1 to offset the relatively smaller payoff to be had from coordinating on redness rather than on bestness. For example, suppose that $v(F_1) = 1/3$, $v(F_2) = 2/3$, and $k = 3$. Then (1) tells us that, in variable frame equilibrium, if r is greater than $11a/14$, players who recognise both redness and bestness will press the red button. The probability that an arbitrary player presses the red button is then 20/27.]

There is empirical support for a pure salience mechanism: Guerra and I gave subjects Hi-Los with $k = 4$, prizes of £6, £5, £4 and £3, and the £5 choice made salient [by the presence of the spade at the top right of the display]. Figure 1.11 shows such a task. [Of thirty-two subjects,

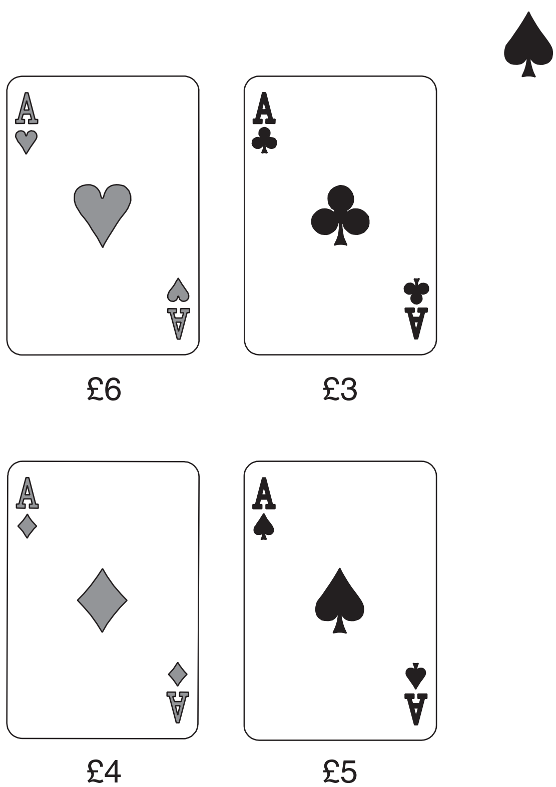

Figure 1.11. Hi-Lo with an inefficient saliency

twenty-three chose the £6 card and eight chose the £5 one.] Making a lower-paying choice more salient significantly dented the hegemony of A, as it should do if all there is going for bestness is its salience.[13]

If the pure salience mechanism were the whole story, then in a context such as the figure 1.11 task, in which another button happens to be sufficiently more salient than the best-paying one, we should feel no qualms about going for it. But this reaction does not match up to our intuitive apperceptions—these seem to be that our attention should be directed towards the best-paying object, even if it's far from the most salient, *because it is the best-paying one*.

Any mechanism in which the best-paying properties have a positive effect on A-choices independently of the salience of A I shall call a *pure bestness mechanism*.

The pure salience mechanism and the pure bestness mechanism are quite different. They are different even if, as usually happens, the best-paying choice and the most salient choice are one and the same. They differ in that they specify different reasons for pressing A. The distinction exemplifies a general feature of the intensionality of reasons. An action may have properties π and π', but if I choose it because it is the π act, then even if I know that the π act is the π' act, it may be false that I choose it because it is the π' act. Indeed I may choose it in spite of its being the π' act.

The fact that the acts supported by pure bestness reasons and salience reasons are highly correlated makes it harder to tell which is operative in naturally-occurring empirical contexts. But because this correlation is contingent and not necessary, it affects not one whit the conceptual distinction.

My own account contains both a pure salience mechanism and a pure bestness mechanism—it says that bestness promotes A-choices both through making A salient and because bestness properties sometimes influence people to choose A independently of salience. Which mechanisms are at work on a given person on a given occasion depends, I shall say, on certain basic features of how she frames the problem. There are two frames. In one salience as such is operative; the judgemental fact is explained by reasons which operate in the other frame.

7.2 Explaining the Judgemental Fact

The bounded rationality theories I have described seem to fail to explain the judgemental fact. But the salience theory, though it does not provide an explanation of this fact, points us in the right direction. For it prompts the question: what property of A *makes* it salient? Suppose we found that the salient-making property was a reason-giving property. Two things might follow: (i) we might have an explanation of the judgemental fact, as being due to the salience of a reason-giving property; (ii) salience itself might drop out as an explanation of the behavioural fact, as an inessential by-product of the reason-giving property.

8. A Germinal Theory

8.1 Demands on a Theory of A

There are three requirements for a good theory of why people play A in Hi-Lo: (i) that it imply observed behaviour, that is, the almost universal choice of A in normal circumstances; (ii) that it do so intelligibly to us, which (to the extent that A intuitively and stably seems to us the only rational thing to do) involves displaying A as uniquely rational—that is, giving principles of rationality which are themselves persuasive, and showing they dictate doing A; and (iii) that it be part of a unified theory of a wide range of problems, not just Hi-Lo—for example, all problems of cooperation.

Of the theories of play in Hi-Lo I have been describing, some address only 'What draws people to A?' and so try to satisfy (i) but don't try to satisfy (ii); others concern the question 'What seems right about A?' but in the form 'What seems right about A to players who may be less than fully rational or think their coplayers are?' The judgemental fact to be explained under (ii), however, is not this but that A stably seems right when there is common knowledge of rationality.

Since we have seen that the principles of standard decision theory and game theory fail to predict A, we are driven to conclude either that the apparent rationality of A is an illusion, or else that the rationality of A derives from principles of rationality that are not included in decision theory. This conclusion means that, to meet (ii), a revisionary theory is required. The new principles in such a theory might conflict with established ones, but they need not—they might constitute a revision only in the sense of an extension. Should we aim at this, or should we judge rather that we are under an illusion? Consider for a moment the illusion theory; it proposes that thinking that one ought to play A is a mistake: that all one is entitled to think is that A is rationally permitted—not forbidden, but not positively indicated either. I stick my neck out and posit that A-choosing is rational. Since I reject the illusion theory, this means I must look for a revisionary one. Playing A is rational, I claim, in virtue of considerations and principles not yet identified (the *Revisionary Conjecture*).

8.2 A Conjecture

My own proposed solution of the Hi-Lo paradox is an application of a revisionary theory of how people play games (the 'variable agency' theory) which I will describe in chapter 4. At this point I will describe some broad features of my explanatory strategy, and some reasons why I think we are pushed towards them by various pieces of evidence, including the failures of the previous assaults on the Hi-Lo problem I've been recounting. I will then point out some challenges a theory of the kind I propose will have to meet.

I suggest that (*Conjecture 1*) a typical person facing any game may have an inclination to reason about what to do in it in a certain way, which for the moment I will label reasoning in *mode-P*. I suggest also that (*Conjecture 2*) in the game of Hi-Lo this inclination is very strong. Mode-P reasoning has two features:

> F1. The player ranks all act-profiles, using a Paretian criterion.
> F2. She takes herself to have a reason to enact her component in the highest-ranked act-profile.

Feature 1 makes mode-P reasoning have something in common with a familiar sort of reasoning about games, the reasoning we engage in when we discuss the 'collectively rational' outcome. A theorist doing this also ranks all act-profiles by assessing the payoff profiles they generate, and on most accounts she does this in accordance with some form of the Pareto principle. We'll see in due course the reason for this similarity.[14] I shall label F1 *collective profile ranking*.

Feature 2 says that if Player 1, say, ranks (*x*,*y*) top, she thinks she should do *x*. I label this feature of mode-P reasoning *projection*, because in mathematical parlance the player's choice is the projection of the whole *n*-tuple onto the dimension of her act.

In ordinary, best-reply reasoning, a player ranks outcomes according to her own component of the payoff-profile of each outcome. In F1, saying that a player uses a Paretian criterion means that if profile *p* is Pareto-superior to profile *p*′ then she ranks *p* higher than *p*′. Because Hi-Lo is a coordination game, in Hi-Lo this results in exactly the same ranking of profiles as ranking them by her own component. But this is not true in some other games, and in these F1 implies a change in the player's ranking; that is, in these cases mode-P reasoning implies a payoff transformation. For example, the utilitarian function, [which sums the players' payoffs], is one way of meeting the Paretian criterion, and in the Prisoner's Dilemma it gives the ranking (C,C) $\succ$ (C,D) ~ (D,C) $\succ$ (D,D), while the own-component ranking is (D,C) $\succ$ (C,C) $\succ$ (D,D) $\succ$ (C,D).

In the case of Hi-Lo, mode-P deliberation yields A as its unique conclusion. The effect of the Paretian ranking requirement in F1 is that (A,A) is the top element in the player's profile ranking, and F2 then implies that she chooses A. So the first piece of evidence that is explained by this theory is the behavioural fact about Hi-Lo, the fact that there is a very strong tendency to play A. In games in general, the requirement means that an agent in mode-P always selects the unique Pareto-optimal profile when there is one, as there is in all common-interest games. In other cooperation games, such as Prisoner's Dilemma, it restricts but does not fully determine the profile choice; more needs to be said about the player's profile ranking before we know which profile out of (C,C), (C,D) and (D,C) is chosen.

In mode-P a player, Player 1, undergoes not only, in general, a payoff transformation but also a 'reasoning transformation'. F2 implies that she abandons the usual way of reasoning to a best act, best-reply reasoning. Instead, she first selects an *act-profile*, then selects as her act her component in it. This procedure has a quite distinct logical form. Suppose for simplicity that Player 1 has a definite belief about Player 2's act (as distinct from a probabilistic one)—say, that Player 2 will do y'. Let U_1 be a payoff function representing Player 1's ranking in mode-P. Best-reply reasoning has the form: choose the act x that maximizes $u_1(x,y')$. [MB uses U_i to denote player i's ranking of act-profiles and u_i to denote player i's individual payoff.] Mode-P reasoning has the form: choose the act x such that (x,y) maximizes $U_1(x,y)$. In Hi-Lo, best-reply reasoning produces A if y′ = A and B if y′ = B; mode-P reasoning produces A.

8.3 Motivation: Game-Theoretic Indeterminacy

The basic problem we confront is how to explain the apparent determinacy of rational choice in Hi-Lo. A theory should reflect intuitions of rationality by revealing a rational basis for it. If these intuitions are determinate then so should the theory be. Note that this meta-theoretical principle is not the same as the one proposed by Harsanyi and Selten, namely, that game theory should be determinate, full stop: 'Clearly a theory telling us no more than that the outcome can be any one of these equilibrium points will not give us much useful information. We need a theory selecting one equilibrium point as the solution of the game' (Harsanyi and Selten 1988, p. 13).

Contemporary game theory apparently fails because there are two equilibria rather than one. And in versions of it that don't contain the Nash equilibrium principle the failure is worse—all choices are, for example, rationalizable.[15] This is a simply a case of a general property of standard game theory: it is generally underdetermined. That may be fine—we should not prejudge the determinacy of reason. For example, if there are Buridanic cases, [where the agent is unable to identify a strictly best option], then in them our basic rational theory *should* be indeterminate.[16] But in Hi-Lo it's not fine, but monstrous, since the choice of A is so obviously right.

The ultimate source of indeterminacy is, it seems clear, the rational agent's attempt to achieve a best act-profile by choosing only one component of the profile. The agent is modeled as deliberating over—seeking the optimum value of—what she herself controls—her own act. So she must take what she does not control, but is controlled by the other, as parametric, and make a hypothesis about the value of this parameter. But there is nothing that ties down her hypothesis except what she can learn from thinking about the other doing just the same sort of thing. This leads to a notorious circle.

Standard game theory gets rid of much of this indeterminacy by adding the assumption of Nash equilibrium. This is essentially the assumption that each player guesses right. I am not in favour of this assumption for the purposes of rational game theory, but we needn't argue about this here; Nash equilibrium fails to confer determinacy anyway, so there is no need to dispute the principle in order to show that standard theory fails for Hi-Lo.

There are two possible ways out. One, pursued above, is to link the player's own act, in the player's mind, to the whole profile in such a way as to give her reason to think she has indirect control over the whole profile. The classic expression of this is evidentialism, which, however, involves fallacy and so provides at best a boundedly rational explanation of A-choices.

The other, which I pursue, is to consider the deliberator as directly choosing a whole profile. This does not mean that she can implement her choice! That would be at least as magical as evidentialism. But on the other hand she is free to ask herself what would be the best profile—just as the social choice or collective rationality theorist does. And what is also true is that the agents can, *between them*, implement their profile choices—they can if all make the same one and if all implement their components. This process may remind the reader of processes that groups of individuals sometimes pursue in in real time. I shall return to this empirical observation.

This is the broad motivation for mode-P reasoning—for the choice of an act by profile evaluation, selection and projection. I note that for the theory to yield profile choices that can all be implemented does not require that players use Paretian rankings, but only that all players' rankings should have a common top element. But the Paretian feature is one way of yielding such agreement in a very broad class of games (including all common-interest games), and there are other motivations for it to which I shall come in a moment.

We can see in another way why F2—profile selection and projection—is needed, that is, why F1—Paretian profile evaluation—would not be enough by itself. In the case of Hi-Lo, F1 makes no difference to the agent's ranking of profiles; so if a player reasons in the usual, best-reply way, we have got nowhere. Putting it otherwise: to explain the facts about Hi-Lo, a traditional payoff-transformation theory which turns (A,A) into the best profile for an individual player is useless, because (A,A) is the best profile for her anyway.

By the same token, in such cases as Hi-Lo one could for certain purposes do without F1—Paretian profile evaluation—and stay with the agents' original evaluations of profiles. One could do without it inasmuch as F2 would still give the same answer, A. However, we are looking for a plausible theory, not an arbitrary device for predicting A. And it turns out that the theory which makes sense of F2 also makes sense of evaluating profiles in a way that reflects their social virtues, such as Pareto optimality, that is, also makes sense of F1. An additional reason for F1 is that, as we have seen above, in other games F2 by itself would *not* give the same answer. (For example, in the Prisoner's Dilemma each would choose D.)

8.4 The Paretian Ranking Requirement

Why should it be assumed that, even if agents do rank profiles, their ranking is Paretian? Suppose we ask, 'What is it about A that seems to make it the obviously rational choice?' When we consult our intuitions, or examine the responses of subjects to this question, the answers seem to revolve round the property of A that the outcome (A,A) is *best for both* players. Another way of saying this is that (A,A) is the unique Pareto-optimal profile.[17]

The idea that the unique Pareto-optimality, or best-for-bothness, of (A,A) enters into the player's choice of A is also found among the *obiter dicta* of game theorists. Two of the most explicit (as on much else) are Harsanyi and Selten. After they have filtered out nonequilibria, they are still left with two profiles, and ask what reason we could have for predicting (A,A) rather than (B,B). Their answer is that (A,A) is salient in virtue of its Pareto-superiority or 'payoff dominance'. They do not address the question of why payoff dominance should be a salient property. One possible answer is that it is a salient reason-giving property of profiles. F2 is an expression of this interpretation.

To the extent that payoff dominance is reason-giving, the fact of its salience drops out of the theory. It need not do so completely, however, for reasons may be more or less powerful. If the reason payoff dominance gives for choosing A is weak, it may be that the tendency to choose A is boosted by a salience mechanism in which payoff dominance operates as a cause of A-choices by a quite different route. This is just what happens in the theory of this book.

In this theory, certain factors in a situation of interaction make it more likely that an agent will reason in a profile-based way *and* that, if she does, she will order profiles in a Pareto-respecting way. These features include payoff features. Among these payoff features are a common interest, but also harmony of preferences. Hence when preferences are completely shared, as in Hi-Lo, the payoff factors have maximal strength. This part of the theory also has some initial support both from behavioural evidence and from intuitions about rationality. Observed rates of A play are higher than rates of S play in typical Stag Hunts, which in turn are higher than rates of C play in typical Prisoner's Dilemmas.[18] And it seems to be part of our intuition about Hi-Lo that A is all the more compelling because not only is (A,A) best for both of us, but we share exactly the same ordering of all outcomes. Our pro-A intuitions are stronger than our pro-S intuition in Stag Hunt or our pro-C intuitions in Prisoner's Dilemma. Unlike in games like Stag Hunt, where there is conflict of interest over paired comparisons other than those involving the best point, in Hi-Lo there is unopposed potential for mutual gain. The only thing that stands in the way of our realizing it by playing in (A,A) is lack of means of overt communication of the contents of our deliberations; there are no 'real forces'.

All the past theories that attempt to rationalize A-choices treat them, in the mind of the agent, as a way of bringing about (A,A) rather than some other profile. But the effect of the Paretian requirement in F1 is just this—it makes the target profile be (A,A) rather than some other. So, it may be asked, is there any point in bringing in the Paretian requirement? The answer is yes, because of the nonextensionality of reasons and the need to model agents' operative reasons. In these theories, the reason why the pair (A,A) is an aim for Player 1 is the property of (A,A) that it maximizes; it isn't its Paretian property (that it is uniquely Pareto-optimal). It might be that a player chooses A for the reason that (A,A) has one of these properties without choosing it for the reason that it has the other. This is true even when, as here, she knows that whatever has one has the other.[19]

What sort of a reason for choosing the profile (A,A) does its Pareto-optimality give? Why should an agent who ranks profiles rank them in a Pareto-respecting way? The very name suggests that such an agent is swayed by considerations of collective or systemic rationality with respect to the collectivity consisting of both (in other cases, all of the) players. It will be part of my thesis that she is.

8.5 Challenges the Theory Faces

A theory which contains 'mode-P reasoning' faces many serious challenges. First, closure: as it stands it appeals exogenously to an inclination to enter mode-P; we need a proper theory of the factors that determine the inclination to reason in this way. Second, collective profile ranking: it remains to be explained why, if and when people start to reason in mode-P, they should be swayed by considerations typical of collective rationality. Third, projection: even granted that an agent should decide that a certain profile p should be realized, why does this give her a reason to do her part in p; given that she favours (A,A), why does this mean she should do A? The question immediately raises a doubt about the whole idea of mode-P reasoning. It's no good, usually, playing your part unless others will too; it will not by itself achieve the desired profile, and it may well be counterproductive in terms of your ranking of profiles. For example, if P2 will in fact choose B, choosing A produces a worse profile than choosing B does if your ranking is the shared ranking given by both u_1 and u_2. Deliberating in mode-P can be futile or worse than standard-mode deliberation if other players are not deliberating likewise. Thus any inclination to do so needs to be tempered by suitable caution. How to include such tempering with caution is a fourth challenge for the theory.

*APPENDIX
Payoffs in Calling a Catch

In the Calling a Catch game, a *state of the world* can be defined by three components: whether F1 really is better-placed (denoted by 1) or F2 is better-placed (2), whether F1 thinks he is better-placed (*b*) or worse-placed (*w*), and whether F2 thinks he is better-placed (*b*) or worse-placed (*w*). This gives eight states, for each of which the prior probability can be calculated straightforwardly. For example, (1,b,b) is the state of the world in which F1 really is better-placed, but each player thinks that he himself is better-placed; the probability of this state is $(1/2) \times (2/3) \times (1/3) = 2/18$. Each player chooses one of the strategies B (go for the catch if and only if you think you are better-placed), W (go for the catch if and only if you think you are worse-placed), or—an option that is allowed only in the second

TABLE 1.1
Payoffs for calling a catch

		Payoff if strategy profile is					
State of the world	Probability	BB	BW	WW	BC	WC	CC
(1,*b*,*b*)	2/18	−5	20	−2	20	−2	10
(1,*b*,*w*)	4/18	20	−5	15	20	−2	10
(1,*w*,*b*)	1/18	15	−2	20	−2	20	10
(1,*w*,*w*)	2/18	−2	15	−5	−2	20	10
(2,*b*,*b*)	2/18	−5	15	−2	−5	10	10
(2,*b*,*w*)	1/18	15	−5	20	−5	10	10
(2,*w*,*b*)	4/18	20	−2	15	10	−5	10
(2,*w*,*w*)	2/18	−2	20	−5	10	−5	10
Expected payoff		9	5.8	7.3	8.8	2.7	10

version of the game—*C* (obey the captain's call). Since the positions of the players are symmetrical, it is sufficient to consider only six strategy profiles: BB (both players choose B), BW (F1 chooses B, F2 chooses W), WW, BC, WC and CC. Given the state of the world and the strategy profile, each player's behaviour is determined. If one and only one player goes for the catch, the payoff for the team is 20 if this player really is better-placed and is acting on B or W; it is 15 if this player really is worse-placed and is acting on B or W; and it is 10 if this player really is better-placed and is acting on C. (A player who really is worse-placed is never called by the captain.) If neither player goes for the catch, the payoff is −2. If both go for the catch, the payoff is −5. Table 1.1 shows, for each state of the world, its probability and the payoff, given each strategy profile. The expected payoff for each profile is shown in the final row.

Notes

1. Sometimes a safe single is not taken because this makes the other batsman the striker, and he may be weaker than the current striker. But this means the payoffs are not as shown.

2. We will come back to possible explanations of why it was 91 per cent rather than 100 per cent. Earlier tasks were matching games in which all prizes

were equal, but one was made salient. In the figure 1.8 task, the £5 card is salient in the spatial configuration. As we shall see in chapter 4, Bacharach and Guerra found that salience can override prize size.

3*. MB was planning to add here some direct evidence that people think A obviously rational.

4. The main tradition in game theory also holds that this is the only kind of reason that can rationalize an action, but agreeing that it is one sort that can does not commit one to this view.

5. Figure 1.a illustrates this case: T is Player 1's Stackelberg act, L is Player 2's, but (B, L) and (T,R) are the only equilibria.

		Player 2	
		L	*R*
Player 1	*T*	1, 1	3, 2
	B	2, 3	1, 1

Figure 1.a. A Stackelberg-nonsoluble game

6. In fact in every common interest game the Stackelberg acts are in equilibrium, and this is the common-interest equilibrium, so the conjecture predicts S in Stag Hunt, however risky S may be. The conjecture is therefore too strong as it stands; however, it remains possible that Stackelberg reasoning occurs, but can be overridden by other considerations.

7. For example, in the Stag Hunt of figure 1.10, (R,R) is risk dominant, because adhering to (S,S) against a coplayer equally likely to be an adherent of (S,S) or (R,R) implies playing S against equally probable S and R, and similarly for adhering to (R,R).

8. It might be possible to exploit the equiprobability idea in a way that does not run into these difficulties by thinking of equiprobability as the initial belief state of a player in an iterative process. Harsanyi (1975) suggests that players generally come to their final beliefs about what each other will do through such a process, called the tracing process. Each begins with a prior distribution over the other's acts, then computes her best reply, then her coplayer's best reply y to this; she then bends her prior somewhat towards y, and repeats the whole process. At each stage the reasoning of a player is level 2, which is a modest but decent level, modal in the distribution found by Stahl and Wilson (1994); but the players also have some inconsistent beliefs: when Player 1's probabilities for Player 2's act are π_t, she also thinks Player 2 has beliefs on which π_t is not optimal, and that Player 2 is a rational optimizer (Bjerring 1977). In other respects the process is too demanding.

9*. In this subsection, the material enclosed by square brackets revises MB's exposition so as to maintain consistency with the treatment of variable frame theory in the introduction. These changes are matters of notation and theoretical detail, not of substance.

10. Mutatis mutandi—their example is not Hi-Lo but a Stag Hunt.

11*. Consider a player who has the universal frame. With probability v_T, her opponent has the frame $\{F_0\}$ and opts for *pick a thing*. With probability v_H, her opponent has the frame $\{F_0,F_1\}$ and opts for *choose the highest*. With probability v_R, her opponent has the frame $\{F_0,F_2\}$ and opts for *choose the red*. To make it as difficult as possible to show that *choose the red* is optimal, assume that if the opponent has the universal frame he opts for *choose the highest*. Then the expected payoff from *choose the red* is $v_T\, r/k + v_R\, r$, while that from *choose the highest* is $v_T\, a/k + (v_H + v_U)a$. The former is higher if (1) holds.

12*. Because this result is concerned only with sufficient conditions for the optimality of *choose the red*, it does not require the principle of payoff dominance to be used. By presenting his analysis in this form, Bacharach avoids having to resolve the issue of how payoff dominance should be understood when the availability of a family of predicates is less than 1. This issue is discussed in section 4 of the introduction, in relation to the game of Large and Small Cubes.

13*. This experiment is described more fully in section 8 of chapter 4.

14. We also find a hint of feature 1 in the equilibrium selection literature, in which it sometimes seems that it is the players, rather than the theorist, who are supposed to select among equilibria—in which case they are also selecting among profiles. But this is nowhere spelled out, and is not the standard interpretation of 'equilibrium selection'. Harsanyi and Selten themselves occasionally veer towards the players-as-selectors interpretation. In discussing payoff dominance they write: 'Clearly, among the three equilibrium points of the game, U_1U_2 is the most attractive one for both players. This suggests that they should not have any trouble coordinating their expectations at the commonly preferred equilibrium point U_1U_2' (Harsanyi and Selten 1988, pp. 80–81). Here Harsanyi and Selten seem to come near to proposing a profile-selection theory, that is, a theory in which players vet and evaluate alternative profiles, and this evaluation governs their decisions.

15*. The concept of rationalizability was introduced to game theory by Bernheim (1984) and Pearce (1984). In a game for two players P1 and P2, a strategy r is rationalizable for P1 if all of the following conditions are satisfied: (i) r maximises P1's expected payoff, given *some* probability distribution over P2's strategies; (ii) every strategy of P2's that is assigned a strictly positive probability in (i) maximises P2's expected payoff, given *some* probability distribution over P1's strategies; (iii) every strategy of P1's that is assigned a strictly positive probability in (ii) maximises P2's expected payoff, given *some* probability distribution over P1's strategies; and so on. Rationalizability differs from Nash equilibrium in not requiring the two players' beliefs to be consistent with each other. In Hi-Lo, for example, it is rationalizable for P1 to choose A (in the belief that P2 will very probably choose A) while P2 chooses B (in the belief that P1 will very probably choose B).

16. True, even for Buridan cases we need, as agents, higher-order reasons, which tell us what to do when lower-order reasons fail to pinpoint. But this does not mean that theory should at all costs introduce further, tie-breaking reasons into the basic theory. If we should, then they are not Buridan cases, *contra hypotesi*—and there are never any Buridan cases.

17. The profile (x^*,y^*) is uniquely Pareto-optimal if and only if for all (x,y) $\neq$ (x^*,y^*), $x^* > x$ and $y^* > y$.

18*. MB planned to insert a footnote documenting this claim.

19. Indeed, this is what classical decision theory obliges us to say. Imagine a situation in which I know that by choosing x I bring it about that my coplayer does (for instance, I know he copies me out of sycophancy), and suppose that (A,A) maximizes both my and my mimic's payoff, both u_1 and u_2. According to decision theory the only reason I can have for action is to bring about consequences I prefer, and u_1 fully represents my preferences about consequences. If I happen to be a Player 2-sympathizer and prefer that she gets what she wants (or a Player 2-antipathizer and prefer that she doesn't), these preferences are already 'in' u_1.