

THE IN A HURRY SERIES

Game Theory

in a Hurry

Strategy when others are deciding too

Max Brocklesby

BOOKSINAHURRY.COM

Contents

The Whole Thing in One Page 4

Why You Should Care 5

The Core Ideas 7

How It Actually Works 20

What People Get Wrong 31

Use It 36

Terms 41

Go Deeper 45

Notes and Sources 47

Bibliography 54

The Whole Thing in One Page

The public image of game theory is a pair of prisoners in separate cells, each wondering whether to betray the other. That puzzle is famous because it is compact. It is also a poor picture of the subject. Game theory is not a collection of traps in which selfish people behave badly. It is the mathematics of conditional choice: what should I do when the result of my action depends on what you do, and you know the same is true of you?

That change sounds small. It removes the floor beneath ordinary decision-making. If the weather will be wet whatever you wear, choose for the weather. If a competitor may cut its price because it expects you to cut yours, your best price depends on its expectation of your expectation. The choice is no longer sitting in front of either player. It is suspended between them.

A game-theoretic model therefore begins by naming the players, the actions open to them, the order in which they move, what each knows, and how each values the possible outcomes. A strategy is then a complete plan for every situation the player might face. Once those pieces are visible, you can look for best responses, discard actions that are worse whatever anyone else does, and ask whether the plans fit together.

The best-known fit is Nash equilibrium. In equilibrium, no player can improve by changing alone while everyone else stays put. That is a test of mutual consistency, not a certificate of wisdom. An equilibrium can waste money, reward aggression, exclude outsiders or leave everyone worse off than another available outcome. There may be several equilibria, giving the theory too many answers, or none in pure strategies, so equilibrium analysis must admit randomisation before the plans can fit.

The calculation also depends on whose interests entered the model. A stable arrangement for firms, voters or states can push costs onto workers, minorities or future generations who were never represented as players.

Timing then changes the game. A threat made today matters only if carrying

it out tomorrow will still make sense. Commitment works by removing tomorrow's freedom. Hidden information adds beliefs, signals and screening. Repetition lets reputation and retaliation pull future consequences into the present. Cooperation can emerge, but only under conditions the model must state rather than admire.

Then the field turns itself around. If people respond strategically to rules, do not ask only how to play the game. Ask who wrote it. Auction formats, school admissions, organ exchange, voting procedures, contracts and regulations can all change what participants find worth doing. Mechanism design begins with the desired result and searches for rules under which pursuing private aims helps produce it.

The cost is severe compression. Real people misread one another, care about fairness, follow norms, make mistakes, hold conflicting aims and change the game while playing it. A model can clarify a strategic structure only after somebody has chosen what counts as a player, a choice and a payoff. Game theory disciplines anticipation. It does not abolish judgement.

That is the book.

Why You Should Care

A patient needs a kidney. A relative is willing to donate, but the blood type or immune match is wrong. Somewhere else, another patient and donor face the opposite problem. In systems where payment for organs is prohibited, an ordinary price market cannot clear the shortage. Yet a carefully designed exchange can link incompatible pairs so that each patient receives a compatible organ from someone else's donor. Add more pairs and the possible chains multiply faster than human intuition can comfortably follow. The problem is medical, logistical and moral. Its operating structure is also a game: people hold private information, have outside options, move at different times and may leave if the rules make participation unsafe.

A mathematically attractive chain can still fail if a donor withdraws after another operation has occurred, so timing, trust and contingency planning

belong inside the design rather than in an appendix.

That is why game theory matters. It is mathematics built for situations in which the objects being analysed can watch the analysis, form expectations and respond. A bridge does not change its load because the engineer predicts it will stand. A bidder may change a bid because the auctioneer publishes a rule, because rivals have read the same rule, and because each knows the others have read it. Strategy begins where prediction enters the thing predicted.

You meet that structure whenever another person's decision changes the value of yours. Two firms choose whether to start a price war. Drivers select routes and collectively create congestion. Employers and applicants send signals about quality. Governments promise rewards, threaten sanctions and wonder whether the promise or threat will be believed. Couples divide unpaid work. Neighbours decide whether to contribute to a shared repair. Platforms write rules for sellers who immediately search for profitable edges. None of these is solved by asking what would be best in isolation, because isolation is the one condition missing.

The subject also corrects a common form of bad advice. People are often told to communicate better, cooperate, be courageous or stop behaving irrationally when the surrounding incentives punish exactly that behaviour. If a worker who reports a fault bears the cost while everyone shares the benefit, silence may be a best response to a bad system. If a peace agreement asks one side to disarm first while the other retains the ability to attack, goodwill does not remove the strategic problem. Moral language can identify what people ought to do. It cannot by itself make doing it safe.

Game theory gives you a disciplined way to ask harder questions. What can each person choose? What do they know? What do they believe others know? Which threats would be carried out? What happens after a deviation? Is the apparent solution stable when one participant reconsiders? Could the same rules produce several self-consistent outcomes, leaving history, convention or accident to select among them? Which people bear consequences without having a move?

It also teaches restraint. A payoff number is not a human motive. Rationality in a model usually means consistency with specified preferences, not intelligence, goodness or emotional coldness. An equilibrium describes a pattern that resists unilateral change. It does not tell you to approve of it. Experiments repeatedly find that fairness, punishment, culture, identity and mistakes affect play, while field institutions contain legal and political details a clean matrix omits.

The reward is not a trick for outsmarting everyone. It is the ability to see when a problem lives between decisions rather than inside one person, and when the strongest move is to alter the information, timing, commitment or rules before anyone moves at all.

The Core Ideas

Your Move Has No Value by Itself

Suppose a small company is considering a new shop. In an ordinary investment calculation it would estimate rent, wages, demand and margin. Now add one fact: a larger rival may open nearby if the newcomer enters, but may stay away if entry looks unprofitable. The new shop's value depends on the rival's response. The rival's response depends on what it expects the newcomer to do and whether fighting entry would cost more than sharing the market. A business plan has become a game.

That is the threshold. A decision problem asks which action performs best against an environment treated as given. A strategic problem asks which action performs best when other decision-makers can respond. Their choices are part of your environment, and yours is part of theirs. The dependence can involve two people or millions, open conflict or quiet coordination, money or any other consequence the participants care about.

A game-theoretic model strips the situation into parts. The players are the decision-makers. Their actions are the moves available at a particular moment. Their strategies are complete plans for how to act. The information structure states what is known, hidden or inferred. Timing says who moves

when and what can be observed. Payoffs represent how the model says each player values the resulting outcomes. Change any element and a recommendation may reverse while the people and physical setting remain unchanged.

The payoff numbers require care. For a choice among certain outcomes, it may be enough that higher numbers preserve the player's ordering. Once mixed strategies or other lotteries are compared, ranks alone are insufficient: the numerical gaps must support expected-utility comparisons. The model is not measuring happiness with a ruler. It is representing preferences strongly enough for the calculation being attempted.

Payoffs need not be cash or moral approval. Safety, status, fairness and revenge can enter if they affect choice and are specified before outcomes are explained. Otherwise a modeller can rescue any failed prediction by rewriting preferences after the event.

Each other piece carries a trap. A firm, ministry, political party or household may be treated as one player only when its internal disagreements do not drive the result. An action is not an outcome: choosing a price does not determine profit until customers and competitors respond. A strategy is not the move that happened. It also specifies conduct after branches never reached.

Rationality is narrower than everyday speech suggests. In the standard model, a rational player chooses consistently in pursuit of the represented preferences and uses beliefs as the model specifies. That does not make the player wise, informed or kind. A perfectly consistent person can pursue a foolish aim. A model can be internally flawless while its players, information or payoffs misdescribe everyone involved.

The boundary matters as much as the calculation. A pollution bargain between two firms may look efficient if nearby residents are absent. A wage negotiation can miss workers who cannot afford to wait. Someone chooses who counts as a player and which consequences count. The mathematics begins after that judgement. It cannot repair an omission it has been

instructed not to see.

Game theory therefore starts with humility disguised as formalism. Before asking for the best move, establish whose choices alter whose results. The first strategic error is often not choosing badly. It is pretending the other chooser is weather.

A Strategy Plans for Other People's Moves

In ordinary language, strategy means a broad intention: grow slowly, defend the centre, avoid escalation. In game theory it has a harsher meaning. A strategy is a complete contingent plan, specifying what a player would do at every information set that might arise, including situations never reached in play.

That distinction looks pedantic until a threat appears. A retailer may announce that it will match any rival's price cut. The announcement is an action. The strategy includes what the retailer will do after a cut of 1 per cent, 20 per cent, a temporary promotion, a regional discount or a product change. The rival does not respond to the slogan alone. It asks which promised reactions are likely to occur under the conditions that would trigger them.

Once strategies are specified, the cleanest question is the best response. Hold the others' strategies fixed. Which strategy gives this player the highest payoff? The answer can change when the others change. If a tennis receiver always leans towards the backhand side, serving wide to the forehand may be best. If the receiver adjusts, the server must adjust too. Strategic reasoning is a loop of conditional answers, not a ranking of moves detached from opponents. A strategy profile collects one strategy for every player. The central tests ask whether any member, taking the rest of that profile as given, has a reason to depart.

For a two-player simultaneous game, the normal form lays the strategy choices in a table. Rows belong to one player, columns to the other, and each cell lists two payoffs, usually with the row player's first. Compare down each column to find the row player's best responses, then across each row

to find the column player's. A cell carrying both marks is a pure-strategy equilibrium. If no cell does, mixed strategies may still make the plans fit.

Sometimes one strategy is best whatever anyone else does. It is then dominant. More often the useful result is negative: one strategy is worse than another against every possible choice by the others. It is strictly dominated and can be removed. Iterated elimination repeats the process, because deleting an inferior strategy may expose another. This can simplify a crowded game before any equilibrium calculation begins.

Dominance is powerful precisely because it asks so little about beliefs. If one option pays less in every relevant case, you need not predict the rival. Yet dominant strategies are uncommon in interesting real settings. A firm may prefer a high price if competitors restrain themselves and a low price if they cut. A government may want to concede if resistance will continue and hold firm if the threat is a bluff. Dependence is what made the problem strategic, so a universally best answer should not be expected.

There is another difficulty. A plan can be optimal against one opponent model and disastrous against another. Calling a move irrational can mean that the critic assumed different beliefs. To evaluate a strategy, ask what it is responding to. If the player expects cooperation, defection may destroy value. If the player expects exploitation, cooperation may invite it. The same action can be generous, reckless or defensive under different strategic environments.

Good modelling therefore separates three claims that everyday argument mixes together. This is what the player did. This is the complete plan implied by that action. This is the belief under which the plan was a best response. Only the third begins to explain the choice.

A strategy is not a motivational speech. It is a map of conditional behaviour. Until the branches are visible, there is no strategy to test, only a preference for how the story should end.

Equilibrium Is Consistency, Not Goodness

John Nash's decisive move was to stop looking for one player who had solved the game and ask whether all players' plans could fit together. A Nash equilibrium is a profile of strategies in which each strategy is a best response to the others. No player can gain by changing alone while everyone else stays with the equilibrium plan.

The phrase changing alone does nearly all the work. Equilibrium does not mean nobody would prefer a different world. It means no represented player can create a better one by a unilateral deviation. If improvement requires several players to move together, the equilibrium may survive while all of them dislike it. Nash stability does not test whether a coalition could coordinate and enforce a profitable departure.

Take a simple coordination problem. Two drivers approach one another on a narrow road. Both keeping left works. Both keeping right works. One choosing left while the other chooses right does not. There are two equilibria because either convention, once expected, gives each driver a reason to follow it. The abstract game does not select between them. Law, history, imitation or a visible cue may do that work.

Now change the payoffs. Two firms would earn more if both kept prices high, but each can gain customers by cutting while the other holds. If both anticipate the temptation, both may cut and earn less. That low-price outcome can be a Nash equilibrium even though joint restraint would pay both more. Stability and collective efficiency have come apart. A Pareto improvement is a change that makes at least one player better off without making another worse off. An equilibrium can fail that test, while a Pareto-efficient outcome can be unstable if one participant can profit by leaving it alone.

The Prisoner's Dilemma gives this structure its famous name. Under the stipulated payoffs, each prisoner has an incentive to defect, so mutual defection is the equilibrium although mutual cooperation would leave both better off. The puzzle is not that people are secretly wicked. Private

incentives can make a jointly preferable outcome unstable. Change the payoffs, information, ability to contract, prospect of repetition or available sanctions, and it may cease to be the same game.

Equilibrium can also be abundant. A coordination game may have several conventions. A bargaining model may permit many divisions once agreement has been reached. Too many equilibria weaken prediction because the solution concept identifies a set while history produces one path. Refinements, learning, focal points and institutional detail can narrow the set or explain selection. Each adds assumptions, giving reality another place to disagree.

A different Nash result addresses cooperative bargaining. Begin with feasible utility outcomes and the disagreement point, meaning what each player receives if agreement fails. Under its axioms, the Nash bargaining solution maximises the product of each player's gain over disagreement. It makes failure explicit. An outside option may shape the disagreement payoff but is not automatically identical to it. The solution is an axiomatic allocation, not a forecast that every negotiation will reach it.

Nor does equilibrium guarantee fairness. A powerful employer and a desperate applicant can settle into an arrangement neither can improve alone under the available options. The worker's inability to leave may stabilise the outcome. Describing that equilibrium is not endorsing it, but formal language can make power look natural when the institutions shaping the options are left unexplained.

Nash equilibrium remains central because it imposes a demanding consistency check. A proposed outcome cannot rest on one participant planning to make a profitable change while everyone else politely ignores it. That eliminates a great deal of wishful thinking.

What remains is not the answer. It is an answer that survives one kind of objection: no player, standing alone inside the model, wants to be the first to leave.

Unpredictability Can Be a Strategy

In matching pennies, two players reveal a coin at the same time. One wins when the faces match; the other wins when they differ. Any fixed choice can be exploited. Once either player's fixed choice is known, the opponent can choose the response that defeats it. There is no equilibrium in pure strategies because every definite plan invites a profitable response.

The solution is to randomise. Each player chooses heads and tails with equal probability. This is a mixed-strategy equilibrium: each uses a probability distribution over complete pure strategies. In this one-move game that is equivalent to mixing over the two actions. Randomness is not noise around the correct plan. It is part of the plan, and it does not require anyone to perform half an action.

The logic appears wherever predictability is exploitable. A goalkeeper who always dives towards the striker's usual side becomes easier to beat. A patrol following one route leaves the others exposed. A bidder whose limits can be read gives rivals room to adapt. The useful frequencies need not be equal, and real players may randomise imperfectly. The principle is that a visible pattern should not give the observer a profitable response.

This connects to the oldest mathematical centre of the field. In a two-player zero-sum game, one player's gain is the other's loss. John von Neumann's minimax theorem established that, for finite games when mixed strategies are permitted, the maximum payoff one side can guarantee equals the minimum upper bound the other can impose. Each player has a distribution that protects the game's value against the worst opposing strategy.

Minimax is defensive. It asks what can be guaranteed against a strategically competent adversary with perfectly opposed interests. That is useful in security, competitive sport and adversarial design. It can be wasteful elsewhere. Treat a supplier, colleague or neighbouring state as a pure enemy and the model removes gains from trade or cooperation before the calculation starts.

A mixed equilibrium has a precise support condition. Every pure strategy

used with positive probability must deliver the same expected payoff against the opponents' mixtures. A pure strategy left unused cannot deliver more. Otherwise the player would shift probability towards the better option. In a two-player game, each player's mixing frequencies make the opponent indifferent among the pure strategies in the opponent's support. The mixture is solving the opponent's incentives as well as concealing one's own action.

Observed frequencies need not reveal conscious randomisation. One person may use a deliberate lottery. A population may contain different fixed types. Players may be learning noisily, or an observer may be uncertain about a pure plan. The same aggregate proportions can arise from different mechanisms, and those mechanisms predict different responses to experience or changed rules. Evolutionary game theory later used population shares rather than private coin tosses, widening the mathematics while changing what a player meant.

Unpredictability is not always valuable. Random pricing may confuse customers, and random promises destroy trust. It matters where opponents can exploit regularity and where the costs of variation do not outweigh the protection.

The larger correction is clean. Rational behaviour need not look determined. When every visible pattern can be attacked, refusing to supply a pattern may be the most disciplined choice available.

Timing Makes Promises and Threats Real or Empty

A simultaneous game hides the sequence of moves. An extensive-form game draws it as a tree. A strategy specifies an action at every information set the player may face, including those never reached. In a perfect-information game, each decision node is its own information set. When a player cannot tell which earlier path occurred, several nodes belong to the same set and the action must be the same at each.

For a finite game of perfect information, the basic method is backward induction. Start at the final decisions and ask what each player would

choose there. Replace those branches with the resulting payoffs, move one step earlier and continue towards the beginning. Today's plan must survive tomorrow's incentives.

Consider an established firm warning a newcomer that entry will trigger a price war. If entry occurs, fighting may impose heavy losses on both firms, while accepting the newcomer leaves the incumbent with lower but positive profit. At that later node, accommodation is the incumbent's best response. A rational entrant anticipates this and enters. The threat fails because it asks the incumbent to hurt itself after the event it hoped to prevent.

This is why a Nash equilibrium in the full strategy table can be too permissive. It may include off-path threats that are never tested because their announcement deters the move that would expose them. Reinhard Selten's subgame-perfect equilibrium requires the strategies restricted to every subgame to form a Nash equilibrium there, reached or not. The threat must remain rational after the bluff is called. In games with imperfect information, some doubtful continuations do not begin proper subgames, so stronger refinements may also specify beliefs at information sets.

Commitment changes the calculation by altering later choice. The incumbent might sign a capacity contract that makes aggressive production cheaper, or delegate pricing to a manager rewarded for market share. A government may place an automatic sanction in law. A negotiator may obtain a public mandate limiting concessions. Each device sacrifices flexibility so that another player believes the future action.

Thomas Schelling saw that weakness can become bargaining power when it is visible and credible. A driver in a game of chicken who cannot swerve may force the other driver to do so. The difficulty is making commitment observable without making disaster unavoidable. Brinkmanship can mean raising the risk that events escape control until the other side yields. The logic influenced nuclear strategy and remains dangerous wherever loss of control is used as leverage.

Timing also determines the value of moving first. A first mover may commit,

occupy scarce capacity or establish a focal point. A second mover may observe and adapt. There is no universal first-mover advantage because advantage comes from the information and commitment created by the sequence. An early but reversible action may reveal intentions while securing nothing.

Backward induction has limits as a behavioural prediction. In centipede experiments, participants often continue beyond the immediate stopping prediction of the standard finite model with narrow monetary payoffs and common knowledge of rationality. Social preferences, doubt about another player's reasoning and learning can alter play. The experiment tests a specified model and population, not the logic of working backwards inside that model.

The test remains indispensable. When someone announces what they will do later, walk to the later information set. Ask whether carrying out the plan will still serve them then. A promise or threat matters only if the future chooser has a reason to honour it, or has been denied the power to escape it.

Information Is Part of the Game

Poker would be trivial if every card were face up. Hiring would change if ability were directly observable. Insurance would be priced differently if risk and behaviour were public. In many games, the strategic problem is not merely what others will do. It is what kind of player they are, what they know and what their action reveals.

Two kinds of ignorance should be separated. Imperfect information means a player does not know the exact history when choosing, as when an earlier action was hidden or simultaneous moves are represented within one information set. Incomplete information means a relevant characteristic, payoff or type is unknown. Ignorance about the path and ignorance about the player create different models.

Knowledge can recurse. Common knowledge means that each player knows a fact, knows that the others know it, and so on. A rule understood

privately by everyone may still fail to coordinate action if nobody knows the understanding is shared. A public announcement can therefore change behaviour without adding a fact any individual lacked.

John Harsanyi supplied the standard route into incomplete information. Nature assigns types according to a probability distribution. A type can represent cost, strength, valuation or another private characteristic relevant to payoffs. Players know their own type, hold beliefs about others and use strategies mapping information into action. A Bayesian Nash equilibrium makes those type-contingent strategies mutual best responses under the beliefs.

Inference has entered play. A low-cost firm may sustain a price that a high-cost rival cannot. An employer may interpret a qualification as evidence about productivity. A bidder seeing rivals remain active may update its estimate of an asset's common value. Actions can change the outcome and the beliefs shaping later actions.

A signal is an observable choice whose meaning depends on the incentives of different types. Cheap talk is costless and non-binding. It may convey nothing when interests conflict completely, yet can transmit useful information when interests overlap enough for some messages to be credible. A costly signal can separate types if imitation burdens them differently. Education, warranties and collateral may carry information, but none is honest by nature. The surrounding payoffs do the work.

Screening reverses the direction. The less-informed side offers a menu intended to make types sort themselves. An insurer combines different premiums and excesses. An employer varies salary and performance pay. A seller offers versions with different features or restrictions. A workable menu must consider which types participate, which option each selects and who is excluded.

More information is not a universal improvement. In a common-value auction, better information can reduce the winner's curse while changing competition and the seller's return. Publishing a patrol schedule helps

legitimate users and potential intruders. Risk disclosure may improve an insurer's estimate while making broad pooling harder for people classified as costly. Full knowledge of bargaining limits can sharpen agreement, extraction or impasse. The effect belongs to the information structure and objective, not to information in the abstract.

These models make demanding assumptions. Beliefs need a source. A common prior is mathematically convenient and sometimes descriptively strained. Types may be multidimensional, manufactured through earlier action or unstable. People can misunderstand their own preferences, while institutions distribute information unevenly through wealth, language and access.

The practical lesson is not to collect everything. Ask what is private, who observes which action, what inference that action supports and whether the rules reward honesty, concealment or imitation. Information changes the game because it changes what each player expects the others to do.

Tomorrow and the Rules Change Today

In a one-shot model with no relevant future interaction, later reward and punishment cannot discipline present play. Repeat the encounter and today's action may change how others treat you tomorrow. The future enters the current payoff, sometimes turning cooperation from an invitation to exploit into a best response.

Return to the Prisoner's Dilemma. In a single round with the standard payoffs, defection dominates. In an indefinitely or uncertainly repeated interaction, cooperation can be sustained when players value the future enough, observe the behaviour that matters and can impose a credible response to defection. The short gain from cheating must be smaller than the discounted loss of future cooperation.

Repeated games permit many strategies. Grim trigger cooperates until the first defection and then defects forever. Tit for Tat begins cooperatively and copies the opponent's previous move. Robert Axelrod's tournaments made Tit for Tat famous because Anatol Rapoport's compact program performed

strongly against the submitted strategies under those rules. Change the noise, horizon, payoffs, population or menu of opponents and the ranking can change.

The folk-theorem family exposes the wider problem. Under stated conditions, patient players can sustain many feasible payoffs that leave each at least as well off as the relevant punishment or security benchmark. Future consequences can enforce cooperation, collusion, exclusion or silence. Repetition creates enforcement capacity. It does not choose the purpose.

An equilibrium must also be reached and selected. Players may learn from feedback, imitate successful rules or inherit a convention. Some adjustment processes converge; others cycle or lock into an inferior convention. Mathematical existence does not guarantee that real players can compute, recognise or learn an equilibrium before the setting changes.

Evolutionary game theory removes conscious calculation. Strategies spread through differential performance in a population. An evolutionarily stable strategy resists invasion by a small share using an alternative under the stated fitness payoffs. John Maynard Smith and George Price used the idea to analyse animal contests without supposing that animals solve equations or act for the good of the species. Ecology still determines the game.

The same logic turns from explanation to design. If participants respond to information, timing and payoffs, rules are active parts of the strategic environment. Mechanism design begins with a property such as efficient allocation, truthful revelation or stable matching, then asks whether rules can make it compatible with participants' incentives.

A second-price sealed-bid auction gives a clean benchmark. Under standard single-item private-value assumptions, bidding one's value is weakly dominant because the winner pays the highest rival bid rather than its own. Independence among values is not needed for that dominance claim, though it matters for other predictions. Common or interdependent values create a different information problem and revive the winner's curse.

Matching markets face different constraints because money may be absent or prohibited. School and medical matching often seek stability: no applicant and institution should be able to abandon their assignments for a mutually preferred match under the model. Kidney exchange instead assembles feasible cycles or chains from compatibility, timing and withdrawal constraints. Either design can aid participation and still conflict with distributional goals.

This closes the causal loop. Interdependence made isolated choice impossible. Once best responses are understood, the same dependence becomes a design lever. Change the action set, sequence, information, enforceability or rights to participate and a different pattern of conduct may become stable.

The danger is technocracy. Designers choose objectives, encode payoffs and decide whose manipulation counts. Participants may collude, create identities, leave the mechanism or contest its legitimacy. A rule that works under one model can fail when enforcement weakens, preferences change or omitted groups acquire power.

Game theory's strongest move is often made before play begins. Do not demand better character from people trapped in a bad equilibrium. Redesign what each person can gain, learn, promise and credibly do. Then check who chose the objective, who has no move and what new game will form around the one designed.

How It Actually Works

Building One Game

Take an illustrative entry problem. A new café can enter a station concourse or stay out. The incumbent can accommodate entry or fight with discounts. If the newcomer stays out, it earns nothing and the incumbent keeps the market. If entry is accommodated, both earn a positive return. If the incumbent fights, both lose money for a period, but the newcomer loses more.

Write those outcomes as payoffs, with the newcomer first. Staying out gives 0 and 5. Entering followed by accommodation gives 2 and 3. Entering followed by a fight gives minus 2 and minus 1. Here the numbers can be read as illustrative profit units. In other games, numerical payoffs must represent preferences strongly enough for any expected-payoff calculation; a bare first-to-last ranking is not always enough.

Now specify timing. The newcomer moves first. The incumbent observes entry and then chooses. Start at the end. After entry, the incumbent prefers 3 from accommodation to minus 1 from fighting. The newcomer anticipates accommodation and prefers 2 from entering to 0 from staying out. Backward induction therefore predicts entry and accommodation.

Change one institutional fact. Suppose the incumbent has already signed an advertising contract that pays only if it meets a low-price target. Fighting now gives it 1 rather than minus 1. Accommodation still gives 3, so nothing changes. Make the contract stronger, raising the incumbent's fighting payoff above 3, and the threat becomes credible. Entry may be deterred. The people have not changed. The action set barely changed. A commitment altered the later best response and therefore the earlier choice.

Now hide the incumbent's cost. The newcomer may face a low-cost type that can fight cheaply or a high-cost type that cannot. Entry depends on beliefs. An early price cut may signal type, but the signal is informative only if imitating it is sufficiently expensive for the high-cost firm. Repeat the market each year and reputation enters. Give a regulator power to restrict predatory pricing and the game changes again.

This is how the machinery works. Define the game, solve the relevant continuation choices, test deviations, then change one assumption at a time. The formal answer belongs to the specified structure, not to cafés in general.

Before the Field Had a Name

In 1838, Augustin Cournot imagined two mineral-water producers choosing output. Each firm treated the other's quantity as given and selected the quantity that maximised its own profit. Put the two best-response calculations together and they met at a stable pair. Cournot had written down the structure of a Nash equilibrium more than a century before Nash, though neither the language nor the general theorem yet existed.

Other fragments appeared around literal games. Ernst Zermelo analysed chess in 1913 and established a determinacy result for finite two-player games of perfect information: one side can force a win, or each can prevent the other from forcing one. The result gives a game-theoretic value in principle; it does not hand a human the strategy or make the game computationally small. Émile Borel studied strategic games and mixed play in the 1920s. These were separate routes towards the same difficulty. Once another calculating mind stands opposite, optimisation becomes circular.

The circle could be described. What it lacked was a general mathematical centre.

The Minimax Breakthrough

John von Neumann supplied one in 1928. His minimax theorem showed that finite two-player zero-sum games have a value once randomisation is allowed: one side's best guaranteed floor meets the other side's least enforceable ceiling. Neither must predict the opponent's exact move. Each can choose a distribution that prevents systematic exploitation.

The theorem gave antagonistic games a value. It also gave strategy a distinctive mathematical shape: optimise against an opponent who is optimising against you. Von Neumann's proof used tools far beyond the recreational puzzles through which the subject is now introduced. The result was narrow in scope and strong within it.

He then joined Oskar Morgenstern, an economist dissatisfied with theories that treated individual choice while ignoring strategic interaction. Their 1944

book, *Theory of Games and Economic Behavior*, announced a new programme. It developed an axiomatic treatment of utility under risk, advanced zero-sum analysis and devoted much of its length to cooperative games in which coalitions could form and divide gains. The book was difficult, ambitious and timed for a world newly interested in formal strategy.

Cooperative game theory starts from what groups can achieve together when binding agreements are available, then asks how the resulting value might be divided. Lloyd Shapley's 1953 value allocates the grand coalition's value by averaging each player's marginal contribution across possible joining orders. It is an axiomatic division rule, not a prediction that real bargainers must accept it. Non-cooperative theory instead specifies the moves, information and enforcement through which any agreement is reached. The branches answer different questions about the same interdependence.

The founding book did not yet provide the solution concept that would dominate non-cooperative games with many players and mixed motives.

RAND, Experiments and the Famous Prisoners

After the Second World War, the RAND Corporation gathered mathematicians, economists, physicists and strategists around problems of military planning and nuclear conflict. Game theory looked made for adversaries whose choices depended on expectations about each other. The fit was never complete. States have internal factions, uncertain aims, accidents and leaders who do not share a payoff table. Yet the language of deterrence, commitment, retaliation and uncertainty offered a way to discipline strategic claims.

In 1950, Merrill Flood and Melvin Dresher devised a non-zero-sum payoff matrix and had two colleagues play it across repeated rounds. Each player had an individual incentive to choose an uncooperative action even though mutual cooperation produced a better joint result. Because the same pair met again, the exercise mixed the one-stage dilemma with learning and reputation rather than cleanly testing a single encounter. Albert Tucker later

supplied the prison story and the name that survived.

The fable's success helped shape the public image of the whole field. Two suspects are separated and offered sentences that reward betrayal. Defection is individually attractive whether the other remains silent or talks. Both therefore defect and receive a worse result than mutual silence. It is clean enough to teach on one page and flexible enough to be applied badly to almost anything.

Its real contribution was structural. It separated individual incentive from joint interest and made institutional questions unavoidable. Communication, enforceable contracts, repeated contact, reputation and changed payoffs are not moral decorations. They alter the game.

Nash Changes the Target

John Nash arrived at Princeton as a young mathematician and treated non-cooperative games directly. His 1950 note in the *Proceedings of the National Academy of Sciences* was two pages long. It proved that every finite normal-form game has at least one Nash equilibrium when mixed strategies are allowed. His 1951 paper developed the framework.

The equilibrium was conceptually spare. Give each player a strategy. Ask whether any one player can improve by changing while the others do not. If nobody can, the profile is an equilibrium. The definition applied to zero-sum conflict, coordination, bargaining, oligopoly and many-player interactions without requiring a central bargain or a coalition theory.

Existence was both triumph and warning. A theorem saying an equilibrium exists does not say it is unique, efficient or easy to calculate. It does not say people will find it. Mixed equilibria may be difficult to interpret, and different refinements can select different predictions. Nash had supplied a common grammar, not a universal forecast.

The 1994 Economics Prize went jointly to Nash, John Harsanyi and Reinhard Selten for pioneering analysis of equilibria in non-cooperative games. The grouping captured what happened after the existence theorem:

the field had to make equilibrium survive time and ignorance.

Schelling Makes Strategy Human

Thomas Schelling worked with less algebra and more attention to the shape of real bargaining. In *The Strategy of Conflict* in 1960, he examined threats, promises, commitment, tacit coordination and the strategic use of risk. His players looked for places to meet without communication, tried to make intentions visible and sometimes gained power by removing options from themselves.

The focal point became one of his best-known ideas. Ask two strangers to choose the same place and time in a city without contacting one another. They may converge on a landmark and noon because each expects the other to expect it. The solution is not contained in the abstract payoffs. Culture, salience and shared imagination select it.

Schelling also showed why a threat's power depends on credibility rather than size. Total retaliation may be terrifying and unbelievable. A limited action that raises the probability of uncontrolled escalation can influence behaviour precisely because nobody claims complete control. His analysis shaped thinking about nuclear deterrence while refusing to pretend that the subject was a clean board game.

The 2005 Economics Prize recognised Schelling and Robert Aumann for improving understanding of conflict and cooperation through game-theory analysis. Aumann's work on repeated games gave mathematical depth to the idea that future interaction can support present cooperation. Schelling showed how institutions, expectations and commitment make those possibilities politically legible.

Time, Types and Refinement

Nash equilibrium could include plans depending on incredible behaviour after an unreachd event. Selten attacked that weakness. In 1965 he introduced subgame perfection, requiring Nash equilibrium in every subgame. Later refinements tested survival under small mistakes. Where imperfect information leaves no proper subgame at a doubtful continuation, sequential equilibrium also disciplines supporting beliefs. The aim was to reject solutions held together by behaviour that would not remain optimal when required.

Harsanyi addressed a different gap. From 1967 to 1968, his three-part treatment of games with incomplete information modelled private characteristics as types drawn by chance. Players formed beliefs about one another and selected type-contingent strategies. Problems that had seemed to contain incompatible private worlds could be represented as one Bayesian game.

This framework became fundamental to auctions, contracts, industrial organisation, political competition and signalling. It also embedded strong assumptions. A common prior distribution and precise type space may be mathematically useful without describing how people form beliefs in a disputed, changing environment.

The field expanded by adding structure to what players knew, when they moved and which plans remained credible. Each advance increased explanatory power and the burden of specification.

The Laboratory Pushes Back

Game theory's theorems are deductive. Its applications to people are empirical. From the mid-twentieth century onward, experiments tested whether participants played dominant strategies, coordinated, bargained, punished and learned as the models predicted.

The ultimatum game became a decisive irritant. One player proposes how to split a sum; the other accepts, making the split happen, or rejects, leaving

both with nothing. With money-only preferences, common knowledge of rationality and a smallest transferable unit, the proposer should offer that minimum and the responder should accept. Experiments beginning with Werner Güth, Rolf Schmittberger and Bernd Schwarze in 1982 found that offers were commonly more generous and low offers were sometimes rejected.

The result did not kill game theory. It forced researchers to reconsider the payoffs and the account of rationality. People may value fairness, reciprocity, status or punishment, and they may doubt what others will accept. Later work found substantial variation across populations and institutions, warning against turning one student laboratory into a universal human type.

Experiments also found learning, framing effects and bounded reasoning. In beauty-contest games, people often choose as if performing only a few rounds of reasoning about others rather than iterating to the formal limit. In centipede games, play commonly continues beyond the immediate stopping prediction. Behavioural game theory grew from the gap between equilibrium as a consistency benchmark and equilibrium as a description of first-time human play.

Cooperation Leaves the Prison

Robert Axelrod invited computer programs to play a repeated Prisoner's Dilemma tournament. Anatol Rapoport submitted Tit for Tat: cooperate first, then copy the opponent's previous move. It was transparent, retaliatory and forgiving after cooperation resumed. It performed strongly in Axelrod's tournaments and became an emblem of reciprocal cooperation.

The emblem outran the result. Change the error rate, population, payoff, number of rounds or menu of strategies and another rule may perform better. A known final round can unravel cooperation by backward induction under restrictive assumptions. An uncertain horizon can preserve it. Networks, reputation and partner choice alter the environment again.

The durable lesson was conditional. Cooperation needs support. The

support may be future trade, punishment, reputation, kinship, shared norms, legal enforcement or the ability to leave. Repetition can sustain good conduct and organised exploitation. A cartel also uses future retaliation to deter defection.

Aumann and others formalised the range of outcomes repeated play can support. The family of folk theorems revealed a paradox: adding the future can make many outcomes consistent, improving the possibility of cooperation while weakening unique prediction. History and institutions return because the mathematics permits several self-enforcing paths.

Animals, Populations and Strategies Without Thought

In 1973, John Maynard Smith and George Price published *The Logic of Animal Conflict*. Animals often settle contests through display or limited fighting rather than escalating every encounter to serious injury. One prominent explanation appealed to restraint for the good of the species. Their game-theoretic account asked whether a behavioural strategy could resist invasion by alternatives through individual reproductive success.

An evolutionarily stable strategy is not a conscious plan. Once common in a population, it cannot be successfully invaded by a small group using an alternative under the stated dynamics and payoffs. Hawk, dove and retaliator strategies became ways to analyse aggression, ownership, mating and signalling without imagining animals calculating matrices.

Evolutionary game theory then fed back into economics and social science. Peter Taylor and Leo Jonker's 1978 model described how strategy shares change with relative performance. Learning models treated adjustment as a process rather than assuming equilibrium appears fully formed. The word player had widened from a deliberating person to a type, organism, rule or behavioural programme.

That gain requires care. Biological fitness is not a moral payoff, and an evolutionarily stable trait is not inevitable, optimal for the group or fixed forever. Ecology, population structure and genetic constraints remain part of

the game.

Computation, Learning and Scale

A small game can be solved by inspection. Large games cannot. A strategy in chess must say what to do after every possible history, and the tree is too large to enumerate. Economic mechanisms may contain thousands of participants, private types and combinatorial choices. Existence does not supply an affordable calculation.

Algorithms therefore became part of the field. Some remove dominated strategies or use backward induction. Others search for equilibria, approximate them, exploit special structure or simulate learning. The Lemke-Howson algorithm, introduced in 1964, is a path-following method for finding an equilibrium in a finite two-player bimatrix game, but the general computational problem can still be difficult. A solution concept may be mathematically elegant and operationally useless if nobody can compute or learn it in the time available.

This pressure encouraged alternatives. A correlated equilibrium allows players to condition actions on signals from a coordinating device, provided nobody benefits by disobeying the recommendation after seeing its own signal. Learning dynamics ask whether repeated adjustment converges towards a solution. Regret minimisation asks whether a player's realised performance approaches what the best fixed alternative would have achieved in hindsight. These ideas connect game theory to online platforms, security allocation and artificial agents without making computation a substitute for modelling.

Scale creates a second warning. An algorithm finds an equilibrium of the encoded game. It does not verify that the encoded payoffs, information and identities match the institution. Faster solution can make a bad model confidently wrong.

Designing Auctions and Matches

Mechanism design reversed the usual question. Instead of accepting a game and solving for behaviour, Leonid Hurwicz, Eric Maskin, Roger Myerson and others asked which rules would make chosen objectives compatible with incentives and private information. The 2007 Economics Prize recognised the foundations of that programme.

Auction theory provided the clearest workshop. William Vickrey showed the strategic properties of second-price auctions. Robert Wilson analysed common-value settings and the winner's curse. Paul Milgrom connected private and common values and helped design formats for selling related items. When the United States needed to allocate radio spectrum licences, economists helped create simultaneous multiple-round auctions so bidders could adjust across related licences instead of buying them one by one without seeing the wider contest. Milgrom and Wilson received the 2020 Economics Prize for advances in auction theory and new formats.

Matching created a different design problem. David Gale and Lloyd Shapley developed an algorithm for stable matching in 1962. Alvin Roth studied labour and school markets in which participants could move early, conceal preferences or bypass assignments. His design work helped reshape medical placements and school choice, while related methods supported kidney exchange. Roth and Shapley received the 2012 Economics Prize.

These applications are game theory at its most consequential and least cinematic. The strategic move is a deadline, priority rule, payment formula or matching algorithm. The aim is not to defeat the participants. It is to make their responses part of the institution's engineering.

How we know

Some claims in game theory are proved from stated assumptions. The minimax theorem, Nash's existence result and properties of particular mechanisms belong to that category. Their certainty is conditional: alter the game, information or solution concept and the theorem may no longer apply.

Claims about people and institutions require different evidence. Laboratory experiments offer control and repetition but can be sensitive to framing, stakes, subject pools and experience. Field data contain real incentives and rules but often make clean causal separation harder. Historical cases add context while making counterfactuals difficult. Cross-cultural experiments show that behaviour in bargaining and cooperation games varies enough to make universal claims unsafe.

Designed institutions provide legible tests because rules, participation and outcomes can be observed, yet success remains objective-specific. An auction may behave differently as bidder composition or values change. A stable match may conflict with equity. A rule designed to make truthful reporting optimal may fail if preferences, enforcement or available messages differ from the formal environment.

The secure method is layered. Use proofs for what follows inside the model, experiments and field evidence for how particular people play, and institutional detail to test whether the model represents the game that exists.

What People Get Wrong

“Game theory finds the winning move”

The name invites the mistake. A theory of games sounds like a manual for beating opponents, and popular puzzles usually arrive with one clever answer hidden under the wording.

Most strategic problems do not contain a move that wins independently of what others do. The field identifies best responses, equilibria, guarantees and trade-offs relative to a model. A minimax strategy can protect a player in a zero-sum game. A dominant strategy, when one exists, performs at least as well against every opposing action. A Nash equilibrium survives unilateral deviation. These are different achievements, and none means that the player defeats everyone.

Some games have several equilibria. Others require mixing. A bargaining

result depends on outside options and timing. A mechanism may aim at efficiency rather than one participant's victory. Chess has a game-theoretic value in principle, but humans do not therefore possess the full winning or drawing strategy from every position.

The correction matters because strategy is conditional. Anyone selling a universal move has removed the other decision-maker from the subject that gave game theory its reason to exist.

“Nash equilibrium means the best outcome”

Equilibrium sounds like balance, and balance sounds desirable. Textbooks reinforce the impression by asking students to find the equilibrium, singular, as though the calculation ends in an endorsed destination.

A Nash equilibrium makes one claim: no player can improve by changing strategy alone, given the others' strategies. It does not test a coordinated deviation by several players. It may be Pareto-inefficient, unequal or sustained by poor outside options. Mutual defection in a Prisoner's Dilemma is an equilibrium although both players prefer mutual cooperation. A traffic pattern can be stable even when coordinated rerouting would shorten many journeys. An oppressive institution can persist because isolated resistance is costly.

There may also be many equilibria. A convention such as driving on one side of the road works because others follow it, and both left and right can be self-consistent. The solution concept does not explain which convention appears without extra history, learning or salience.

The correction separates description from judgement. Stability can explain why an arrangement persists. It cannot establish that the arrangement is efficient, fair or worth preserving.

“The Prisoner's Dilemma proves people are selfish”

The prisoners are memorable because the story turns a payoff structure into a moral drama. Each suspect betrays the other, so the puzzle is often presented as a revelation about human nature.

The result follows from stipulated preferences and options. Defection dominates because the assigned ranking makes it better whether the other cooperates or defects. If a player values loyalty, hates being an informer, expects retaliation, doubts the prosecutor or anticipates future contact, the payoffs change. It may no longer be a Prisoner's Dilemma.

Experiments do not reveal one universal tendency. People cooperate, defect, punish, forgive and learn at rates that vary with framing, repetition, communication, institutions and population. The one-shot model remains useful because it shows how private incentives can undermine a joint interest even among clear-headed players. It does not show that every human relationship has those incentives.

Calling climate policy, marriage or office cleaning a Prisoner's Dilemma without checking the payoff ordering creates a slogan rather than an analysis. First establish the game. The fable cannot do it for you.

“Random play is irrational”

Randomness looks like the absence of thought. Coaches criticise erratic players, managers demand consistency, and a coin toss seems inferior to expertise.

Against a responsive opponent, consistency can be a leak. In matching pennies, any pure plan is exploitable, while the equilibrium mixes evenly. In sport and security, a predictable tendency lets the other side concentrate its response. In equilibrium, each used action must earn the same expected payoff against the opponent's mixture, while unused actions can do no better. Randomisation can protect the value a player can guarantee.

This does not bless chaos. A mixed strategy specifies probabilities chosen

for a reason. The correct mixture depends on payoffs and opposing options. Randomising where customers need reliability or colleagues need trust can destroy value. Real players may also approximate a mixture through varied routines rather than conscious dice.

The misconception persists because people confuse uncertainty in execution with uncertainty in design. A disciplined strategy may deliberately leave the next action uncertain. The rational element lies in the distribution and the incentive it creates for the observer, not in making every moment look planned.

“A severe threat is a credible threat”

Power is often measured by the damage someone can announce. The bigger the punishment, the stronger the deterrent appears.

A threat influences another player only if it is believed, and belief depends on later incentives. An incumbent may threaten a ruinous price war, but after entry it may prefer accommodation. A state may promise overwhelming retaliation for a small provocation, yet hesitate once retaliation would impose catastrophic costs on itself. Severity can weaken credibility by making execution less attractive.

Commitment can repair the gap. Automatic rules, delegated authority, irreversible investments, public mandates and reputational stakes may change the payoff at the moment of action. They also create danger by reducing room to adapt. Schelling's strategic insight was that bargaining power may come from visibly losing control, not possessing limitless control.

The correction matters whenever a plan relies on what someone says they will do later. Walk forward to that decision point. If the threatened action will then be self-defeating and no commitment has changed the incentives, the threat is theatre.

“More information always helps”

In ordinary decisions, information usually improves choice. Strategic settings add a second effect: what you learn, what others learn and what each infers can change behaviour.

A bidder who wins an asset with uncertain common value may discover that victory is bad news, because winning means holding the highest estimate. Public information can reduce that winner's curse while changing bidding competition. Publishing a security patrol pattern aids legitimate users and potential intruders. Finer insurance classification can improve prediction while shrinking the pool available to costly applicants. Revealed bargaining limits may ease agreement, expose no overlap or let one side extract more of the surplus.

Information also has distributional effects. A platform may know far more than its users and design rules around that advantage. A signal may be available only to people who can afford the credential. Disclosure can reward sophisticated participants and expose others.

The right question is not whether information is good. Ask who receives it, when, whether it is reliable, what action it changes and what the other side infers from having or withholding it. Facts enter a game through an information structure, not as neutral light.

“Every strategic problem is zero-sum”

Competition dominates the imagery of strategy: wars, card tables, elections and firms taking customers from rivals. It is easy to assume that one player's gain must be another's loss.

Zero-sum games are a special class. They support clean minimax reasoning because the interests are perfectly opposed. Many important games mix conflict with common interest. Buyers and sellers disagree over price while both may benefit from trade. States bargain over burdens while sharing an interest in avoiding war. Colleagues compete for promotion while depending on the same project. Even opponents may prefer several outcomes to

mutual destruction.

Games can also be positive-sum or negative-sum relative to a baseline. Cooperation may create value; arms races and congestion may consume it. The sum measured for named players can still hide losses imposed on outsiders, so even the label depends on where the modeller draws the boundary.

Treating a mixed-motive game as zero-sum can create the conflict the model assumed. It directs attention towards concealment and defence while hiding coordination, side payments, repeated gains and redesign. The opposite error is equally serious: calling a conflict win-win does not erase incompatible claims or unequal power.

The correction is to inspect payoffs rather than the atmosphere. Some outcomes enlarge the available surplus, some divide it, and some destroy it. Strategy begins by knowing which problem is present.

Use It

Find the Real Players and Choices

When a situation feels strategic, do not begin with motives. Draw the decision-makers and the moves they control.

A company may say that the market forced a price rise, when the relevant players include suppliers, competitors, regulators and customers with different options. A household argument may appear to have two players while employers, school hours and housing costs constrain the feasible choices.

Then look for people who receive consequences without possessing a move. Residents affected by a development, junior staff subject to a rota and future taxpayers carrying a guarantee may sit outside the formal game. Their absence can make an equilibrium look efficient because their losses were never entered.

The useful question is: whose choice can alter whose payoff, and who bears a payoff without being allowed to choose? Until that map is credible, calculation gives precision to the wrong game.

Write the Strategy, Not the Intention

Replace statements such as we will be firm, they will cooperate or I will never concede with contingent plans.

What will you do if the first offer is rejected? What if a rival copies the product, undercuts by a little, undercuts below cost or waits? What will happen after a missed deadline, a partial breach or an honest error? A plan that covers only the hoped-for path is not a strategy. It is a forecast wearing armour.

This exercise exposes inconsistency early. A manager may promise zero tolerance for lateness but know that dismissing a scarce specialist would be costly. A customer may threaten to leave but lack a substitute. An investor may claim a long horizon while planning to sell after a 15 per cent fall. The other side will often see these gaps.

Write enough branches to reveal what the promise means when tested. Then decide whether later incentives support it or whether a commitment device, discretion or a less theatrical claim would work better.

Test One-Person Deviations

When someone proposes an arrangement, ask what happens if one participant quietly changes course while everyone else follows the plan.

A rota may look fair until one worker can avoid undesirable shifts without consequence. A cartel agreement collapses if each firm can expand output while rivals restrain themselves. A shared subscription fails if each member can stop paying while access continues. An environmental pledge weakens if the private saving from defection arrives now and the collective cost is dispersed later.

This is the Nash test used as a practical audit. It does not prove an

arrangement desirable, test coalition departures or explain which equilibrium people will learn. It does identify plans that depend on profitable individual disobedience never occurring.

Where the deviation wins, do not immediately condemn the person. Ask which monitoring, reward, punishment, ownership rule, repeated relationship or outside option would change the comparison. Stable cooperation usually has machinery beneath the sentiment.

Audit Credibility at the Moment of Action

Threats and promises should be judged from the future decision point, not from the confidence with which they are announced.

Imagine that the triggering event has happened. The competitor entered, the tenant missed the date, the state crossed the line, the child ignored the warning. Would carrying out the threatened response now improve the threatener's position? If not, what prevents retreat? A contract, deposit, automatic process, delegated authority, public commitment or reputational cost may do so. Anger alone is unreliable because it may cool exactly when action becomes expensive.

Use the same test for promises. A supplier may promise priority during a shortage, but later have an incentive to serve a higher-paying buyer. A government may promise not to rescue a failing institution, then confront wider damage after failure begins.

Credibility is not a character judgement. It is an incentive that survives arrival. Where flexibility has value, the best design may preserve discretion while replacing grand commitments with smaller, believable steps.

Ask What the Move Reveals

Actions carry information beyond their immediate effect. Before responding, separate what happened from what it may signal.

A low introductory price could reveal low costs, desperation, a temporary subsidy or a plan to lock in customers. Refusing a first offer could signal a

strong outside option or merely a negotiating habit. A long warranty may indicate confidence in quality, but only if poor-quality sellers find imitation expensive. Silence may conceal weakness, protect privacy or show that communication has broken down.

Then reverse the lens. What does your own move teach the other side? A rapid concession may reveal urgency. A punitive response may establish resolve and also reveal sensitivity. Publishing a reserve price informs bidders while limiting discretion.

Do not infer types from one action without comparing alternative explanations and incentives. Signals are reliable when different types face different costs or benefits from sending them. Otherwise a move may be cheap talk, copied theatre or noise.

Change the Game Before Blaming the Players

Repeated bad behaviour is often evidence about rules.

If staff conceal errors, reporting may impose a private cost while correction produces a shared benefit. If drivers crowd one route, each may be responding sensibly to the options visible at departure even though the resulting congestion harms everyone. If online sellers use manipulative listings, the ranking and payment system may reward the behaviour more consistently than a code of conduct discourages it.

Redesign can change the action set, timing, information or payoffs. Make reports confidential. Pay for verified quality rather than raw volume. Require deposits that make cancellation costly. Reveal queue lengths. Separate incompatible roles. Allow participants to exit. Match people through a rule that resists profitable bypass. Sometimes the cleanest intervention is to remove a game by setting a fixed right rather than forcing repeated bargaining.

Every design creates a new strategic response. Test gaming, collusion, exclusion, false identities and burden shifting. The goal is not to make self-interest disappear. It is to stop asking the institution to survive

incentives it created against itself.

The limits

Game theory clarifies interdependence by removing detail. That is its power and its chief danger.

Payoffs may be unknown, unstable or impossible to place on one scale. A person can care about money, duty, identity, fairness and the meaning of the act at once. Organisations contain internal politics. States misperceive one another. People learn, imitate, panic, forget and change preferences. Rules may be disputed rather than fixed, and those with power can rewrite them during play.

Equilibrium adds another limit. It can be a useful benchmark without being a behavioural prediction. Players may not know the game, may reason only a few steps, or may coordinate on history and norms absent from the formal model. Multiple equilibria can leave selection unresolved. A unique equilibrium can still be empirically poor if the assumed preferences are wrong.

Formal elegance can also hide politics. The designer chooses the objective, the participants and the permitted moves. Efficiency for those inside may rest on costs pushed outside. Making the intended action privately optimal does not establish justice, consent or legitimacy.

Use game theory as a disciplined counterfactual machine. State the assumptions, derive what follows, compare it with behaviour and alter the model when the world refuses the prediction. Do not confuse a clean solution with a complete situation.

The one thing to keep

Keep the response inside the choice.

Most weak strategy treats another person as a fixed obstacle. It asks what you want, chooses the strongest-looking action and then adds a paragraph about implementation. Game theory forces the missing step into the centre:

what will the others do because you did that, what will you do because they did, and does the loop settle anywhere?

That question changes how problems look. A discount is also an invitation to competitors. A deadline is also information about urgency. A sanction is also a test of willingness to enforce it. A generous act may begin cooperation or expose a target, depending on the continuation strategy. A rule is also a menu of profitable evasions.

The discipline is neither cynicism nor faith in calculation. It does not assume that everyone is selfish, only that other people have purposes of their own and may respond. It asks you to distinguish a desired outcome from an equilibrium, a speech from a commitment, a visible action from a complete strategy, and an elegant model from the institution it has simplified.

Then it gives you a second chance. When the best responses produce an outcome nobody wanted, stop demanding a heroic move from one player. Change what can be observed, promised, punished, rewarded or chosen. Strategy is not merely selecting a move inside rules. It is seeing that rules create moves, moves reveal information, and every participant is deciding too.

Terms

Game. A model of interdependent choice specifying players, available actions or strategies, payoffs, information and timing. The word covers markets, bargaining and institutions as well as literal games. A game need not be competitive or entertaining.

Player. A decision-making unit inside the model. It may be a person, firm, state, coalition, organism or algorithm, provided its internal conflicts can be ignored safely. The simplification fails when internal bargaining drives action.

Action. A move available at a particular decision point, such as entering a market, setting a price or accepting an offer. An action is one part of a strategy. Different nodes may offer different action sets.

Strategy. A complete contingent plan stating what a player will do at every information set it might reach, including branches that never occur during observed play. This distinction matters most in sequential games.

Payoff. The utility number assigned to a player at an outcome. It may reflect money, safety, status or fairness. Expected-payoff comparisons require numerical gaps carrying more information than a bare rank order.

Utility. A numerical representation of preference. Under expected utility, positive affine changes preserve the same representation for one player. Utility is neither a direct happiness measure nor generally comparable between people.

Rationality. Consistent pursuit of the preferences and beliefs specified by a model. It does not imply intelligence, accurate information, selfishness, emotional coldness or a worthy objective. Different models impose different consistency requirements.

Best response. A strategy giving a player its highest payoff against the strategies chosen by the others. Best responses are conditional and may change when opponents change course. Several best responses may tie.

Dominant strategy. A strategy never worse than any alternative, whatever others do. It is weakly dominant if ties are possible and strictly dominant if it beats every alternative in every case. Such strategies are uncommon.

Dominated strategy. A strategy beaten by another whatever others do. Strict domination means a lower payoff in every case; weak domination allows ties and at least one strict loss. Removing strictly dominated strategies can simplify equilibrium analysis.

Normal form. A representation listing players, their available strategies and the payoff produced by each strategy profile, commonly displayed as a matrix for two-player games.

Extensive form. A game tree showing the order of moves, possible histories, information available at each decision point and payoffs at the terminal branches.

Information set. Decision nodes a player cannot distinguish when choosing. The player must take the same action at all nodes within that set because it does not know which was reached.

Nash equilibrium. A profile from which no player benefits by changing its own strategy while everyone else holds theirs fixed. Each plan is optimal against the plans currently faced.

Pure strategy. A definite contingent plan selecting one action at every information set. Pure-strategy equilibria need not exist even in finite games.

Mixed strategy. A probability distribution over pure strategies. In equilibrium, strategies used with positive probability tie in expected payoff and unused ones do no better. Mixing can prevent exploitation.

Zero-sum game. A game in which one player's gain equals another's loss, so the total payoff is fixed. Most bargaining, trade and coordination problems are not zero-sum.

Minimax. The paired zero-sum rules of maximising the minimum payoff one can guarantee and minimising the opponent's maximum attainable payoff. In finite two-player zero-sum games with mixed strategies, the two values coincide.

Pareto efficiency. A condition in which no feasible change can make someone better off without making someone else worse off. Efficiency neither implies fairness nor guarantees strategic stability.

Cooperative game. A model centred on what coalitions can achieve and how their joint value might be divided when binding agreements are treated as available. It abstracts from the bargaining process that forms them.

Prisoner's Dilemma. A payoff structure in which defection is individually dominant for each player although mutual cooperation would make both better off than mutual defection.

Backward induction. Solving a finite sequential game from its final decisions towards the start, replacing each later branch with the payoff

produced by optimal continuation play.

Subgame-perfect equilibrium. A strategy profile whose restriction to every subgame is a Nash equilibrium. It rules out equilibrium profiles supported by threats that would not be rational in an unreached later subgame.

Commitment. A deliberate restriction or alteration of future choice that makes a promised action credible. Contracts, delegation, irreversible investments and automatic rules can create commitment.

Bayesian game. A game of incomplete information in which privately known types, such as costs or valuations, affect payoffs. Players hold beliefs over others' types and map their own information into actions.

Shapley value. An allocation rule for cooperative games that averages each player's marginal contribution across all possible orders in which the grand coalition could form. It is axiomatic rather than a forecast of bargaining.

Signal. An observable action or attribute that changes beliefs about hidden information. It separates types only when their incentives or costs of imitation differ sufficiently.

Repeated game. A stage game played more than once. Future reward, punishment and reputation can change present incentives, though repetition can support collusion as well as cooperation.

Evolutionarily stable strategy. A strategy that, once common, resists invasion by a rare alternative in the stated fitness and population environment.

Mechanism design. The reverse-engineering branch of game theory: choose rules, messages and outcomes so that strategic behaviour produces a desired property such as truthfulness, efficiency or stability.

Go Deeper

The accessible overview

Avinash K. Dixit and Barry J. Nalebuff, *The Art of Strategy: A Game Theorist's Guide to Success in Business and Life* (W. W. Norton, 2008). This is the friendliest bridge from the subject's formal ideas to recognisable decisions. It explains backward reasoning, commitment, signalling, bargaining and strategic interaction through cases rather than theorem-proof exposition. Read it for breadth and intuition. Its business framing occasionally makes strategic cleverness feel more universal than it is, so keep the modelling limits from this book beside it. It is strongest when it turns an intuition into a decision tree and weakest when an anecdote carries more weight than the institutional setting.

The founding text

John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 60th Anniversary Commemorative Edition (Princeton University Press, 2004; first published 1944). This is the primary document that established game theory as a programme for economics and social science. It is historically indispensable, mathematically demanding and unlike the modern Nash-centred textbook field it helped create. Read the opening argument about economic interaction and the treatment of utility, then sample rather than march through the entire volume unless you want the cooperative theory on its own terms. Notice how much of the founding architecture concerns coalitions and shared gains, rather than the two-prisoner image that later came to represent the field.

The strategic imagination

Thomas C. Schelling, *The Strategy of Conflict* (Harvard University Press, 1960; paperback edition with a new preface, 1980). Schelling shows what formal strategy can see when it takes communication, focal points, commitment, bargaining and limited control seriously. The nuclear setting is unmistakably Cold War, but the analysis of threats and promises has aged better than many later formulas. It requires little mathematics and rewards slow reading because the examples often reverse the obvious interpretation of strength. Use it to understand why fewer options can increase leverage and why imperfect control can become a bargaining instrument.

The empirical correction

Colin F. Camerer, *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press and Russell Sage Foundation, 2003). This is the substantial next step after learning equilibrium concepts. Camerer compares formal predictions with laboratory evidence on bargaining, coordination, learning, mixed strategies and bounded reasoning. It is denser than the other recommendations and some evidence has developed since publication, but its method remains exemplary: treat equilibrium as a benchmark, examine how people depart from it, and build better models rather than declaring either mathematics or behaviour defeated. Pay particular attention to how experience, feedback and strategic sophistication change play across rounds, because first decisions and learned decisions answer different questions.

Notes and Sources

Scope, terminology and the reader model

The definitions of games, actions, strategies, payoffs, information sets, best responses, dominance and equilibrium follow the standard non-cooperative framework presented in Martin Osborne's *An Introduction to Game Theory* and Osborne and Ariel Rubinstein's *A Course in Game Theory*. The manuscript distinguishes an ordinal ranking of certain outcomes from the cardinal structure needed for expected-utility comparisons. Positive affine transformations preserve a player's expected-utility representation; the numbers are not direct psychological measurements or interpersonal units. A strategy is a complete contingent plan assigning an action at every information set, which is why unreached branches matter in sequential games.

The distinction between a decision problem and a strategic problem is organisational rather than a claim that uncertainty disappears outside game theory. Decision theory can include uncertain states and other people's behaviour as probabilities. Game theory becomes useful when those other decision-makers are modelled as purposeful responders whose choices depend on their expectations of one another.

The kidney-exchange opening is based on Alvin Roth, Tayfun Sönmez and M. Utku Ünver's 2004 account, Michael Rees and colleagues' 2009 account of a non-simultaneous donor chain, and the Royal Swedish Academy's 2012 materials on stable allocations and market design. The text avoids claiming that one algorithm, legal rule or operational design applies across jurisdictions. It uses incompatible donor-patient pairs to show why preferences, participation, timing and withdrawal risks belong inside the mechanism.

The Core Ideas

Nash equilibrium is defined against unilateral deviations, following Nash's 1950 existence note and 1951 paper. The existence claim is restricted to finite games with mixed strategies. The manuscript separates existence from uniqueness, efficiency, computation, learning and empirical prediction. Pareto efficiency is treated as a different property: an outcome can be stable without being Pareto-efficient, and an efficient outcome can be unstable.

The Nash bargaining solution follows his 1950 *Econometrica* paper. The manuscript states the two-player product form relative to a disagreement point and presents it as an axiomatic cooperative solution, not an empirical law of negotiation. An outside option may help determine the disagreement payoff but is not treated as definitionally identical to it. Practical tactics, communication and deal process remain with *Negotiation in a Hurry*.

The zero-sum and minimax account follows John von Neumann's 1928 theorem and the later treatment in von Neumann and Oskar Morgenstern's *Theory of Games and Economic Behavior*. The equality applies under the finite two-player zero-sum mixed-strategy formulation used here. The matching-pennies example illustrates why no pure-strategy equilibrium exists while equal mixing does. The support condition is stated precisely: every pure strategy used with positive probability must maximise expected payoff against the opponents' mixtures, and unused strategies cannot do better.

Cooperative game theory is represented through von Neumann and Morgenstern's coalition analysis and Lloyd Shapley's 1953 value. The Shapley value is described as an axiomatic allocation based on average marginal contribution across coalition-entry orders. The text does not claim that it predicts bargaining outcomes, guarantees fairness under every moral theory or resolves enforcement.

Sequential-game analysis follows Reinhard Selten's 1965 introduction of subgame perfection and Schelling's treatment of commitment, threats and bargaining. Subgame perfection is described as requiring Nash equilibrium

in every subgame. David Kreps and Robert Wilson's 1982 sequential-equilibrium paper supports the brief warning that games with imperfect information may require beliefs at information sets as well as subgame tests. The claims about centipede play are grounded in Richard McKelvey and Thomas Palfrey's 1992 experiment and are presented as a setting-specific empirical departure, not a universal rule or a refutation of backward induction inside the stipulated model.

Incomplete-information games follow John Harsanyi's three papers published in 1967 and 1968. The distinction between imperfect information and incomplete information is kept visible. Signalling examples draw on Michael Spence's job-market model. The cheap-talk claim follows Vincent Crawford and Joel Sobel's 1982 analysis: costless, non-binding communication can transmit information when interests are sufficiently aligned, but it need not do so. No credential, warranty or price is treated as inherently informative. Screening and insurance examples are illustrative applications of standard adverse-selection models rather than claims about one live market.

The winner's-curse discussion follows common-value and interdependent-value auction theory associated with Robert Wilson, Paul Milgrom and Robert Weber. It is limited to settings where the asset has an important common component and bidders possess noisy information. It is not applied to all auctions. The second-price statement follows William Vickrey's 1961 analysis and is bounded to the standard single-item private-value setting. Truthful bidding is weakly dominant there without requiring bidders' values to be statistically independent; independence matters for other conclusions, including common benchmark revenue comparisons.

Repeated-game claims follow Robert Aumann's work, the folk-theorem literature and Drew Fudenberg and Eric Maskin's 1986 result. The text limits the broad support claim to feasible, individually rational payoffs under stated patience, monitoring and punishment conditions. It does not suggest that repetition automatically produces cooperation. Axelrod's tournaments and

Anatol Rapoport's Tit for Tat are based on Axelrod's 1984 synthesis. Performance is confined to the submitted strategies and tournament environment; no universal optimality claim is retained.

Evolutionary game theory is anchored in John Maynard Smith and George Price's 1973 *Nature* paper and Maynard Smith's later book. Peter Taylor and Leo Jonker's 1978 paper supports the historical sentence on replicator dynamics. The manuscript distinguishes population stability from conscious calculation, group benefit, moral desirability and inevitability.

Mechanism design is framed through the Royal Swedish Academy's 2007 scientific account and the foundational work of Leonid Hurwicz, Eric Maskin and Roger Myerson. Desired properties such as efficiency, truthfulness, budget balance and stability can conflict. The text therefore treats design as objective-specific and assumption-dependent rather than as a route to a neutral optimum.

Historical development

The Cournot discussion refers to his 1838 duopoly model, in which each producer chooses quantity against the other's quantity. Calling the resulting intersection a Nash equilibrium uses modern terminology to describe the mutual-best-response structure; it does not attribute Nash's general concept to Cournot.

The treatment of early mathematical games follows Robert Leonard's history. Zermelo's 1913 chess result is compressed to a determinacy statement for finite two-player perfect-information games and does not imply computational tractability. Borel is credited with early study of strategic games and mixed play, not with von Neumann's general minimax proof.

Von Neumann's theorem appeared in 1928. *Theory of Games and Economic Behavior* first appeared in 1944. The manuscript notes that the book devoted extensive attention to cooperative games because the later public image of game theory as almost entirely Nash equilibrium and Prisoner's Dilemma is historically misleading.

The RAND account is based on Merrill Flood's 1952 research memorandum *Some Experimental Games*, later published in *Management Science*, and on Leonard's history. Flood and Melvin Dresher devised the relevant payoff matrix and repeated experimental setting in 1950; Albert Tucker supplied the prison-sentence story and name. The manuscript makes the repetition visible so the observations are not presented as a clean one-shot test. It does not claim that the game was invented as a complete model of nuclear deterrence. RAND's interest in strategic analysis forms the institutional setting.

Nash's 1950 note occupies pages 48 to 49 of volume 36 of the *Proceedings of the National Academy of Sciences*. His fuller 1951 paper appeared in the *Annals of Mathematics*. The 1994 Economics Prize materials were used to verify the joint recognition of Nash, Harsanyi and Selten and the Academy's description of their contribution.

Schelling's focal points, credible commitments, bargaining and strategic risk are drawn from *The Strategy of Conflict*. The manuscript does not turn brinkmanship into a recommendation or reduce nuclear history to a payoff matrix. The 2005 prize citation and background materials were checked for the joint recognition of Schelling and Aumann's work on conflict and cooperation.

Selten's first subgame-perfection treatment appeared in 1965. Harsanyi's incomplete-information series appeared across 1967 and 1968. The wording avoids implying that either refinement removes all equilibrium multiplicity or supplies beliefs automatically.

Experiments and behavioural limits

The ultimatum-game account is anchored in Werner Güth, Rolf Schmittberger and Bernd Schwarze's 1982 experiment. The narrow money-maximising prediction is stated with its assumptions. The text does not supply a universal offer or rejection percentage because results depend on stakes, protocol, experience and population.

Variation beyond standard student laboratories is supported by Joseph

Henrich and colleagues' 2001 and 2010 cross-cultural work. Those studies cover particular small-scale populations and experimental designs. They justify rejecting one universal behavioural response, not ranking cultures or claiming that market integration or religion alone causes fairness.

The beauty-contest reference follows Rosemarie Nagel's 1995 experiment and later cognitive-hierarchy work. The centipede reference follows McKelvey and Palfrey. These examples show bounded depth of reasoning and departures from the simplest equilibrium prediction in particular settings. They do not establish that participants are generally irrational.

Colin Camerer's *Behavioral Game Theory* supplies the wider empirical synthesis. The manuscript treats changed preferences, learning and bounded reasoning as alternative modelling responses to observed behaviour. A rejection in the ultimatum game can be rational under preferences that include punishment or fairness; changing the payoff model is not an after-the-fact rescue unless the revised preferences are independently tested.

Cooperation, computation and design

Axelrod's tournaments are used as a historical turning point in the study of repeated cooperation. Tit for Tat's properties are described without claiming that it dominates under noise, evolutionary change or every payoff structure. Repetition can enforce cartels and exclusion as well as reciprocal cooperation.

The Lemke-Howson algorithm dates to 1964 and concerns finite two-player bimatrix games. Robert Aumann introduced correlated equilibrium in 1974. Regret-based learning is represented by Sergiu Hart and Andreu Mas-Colell's 2000 procedure. The computation section makes no claim that these methods solve every large game efficiently or that every learning process converges. Its purpose is to separate existence, computation, selection and learning.

Auction design draws on Vickrey, Wilson, Milgrom and Robert Weber, plus the Royal Swedish Academy's 2020 materials. The simultaneous

multiple-round spectrum-auction example is described at the level needed to show why linked licences and sequential sale can create exposure problems. No revenue figure or claim of universal format superiority is retained.

Stable matching begins with David Gale and Lloyd Shapley's 1962 paper. Roth's empirical and design work on the American medical labour market supplies a stable-matching application. Kidney exchange with Sönmez and Ünver is a related market-design application built around compatibility cycles and chains rather than Gale-Shapley stability. The text states that either design can conflict with distributional objectives and does not claim that formal success makes an allocation fair.

Current verification and evidence limits

The official Royal Swedish Academy and Nobel Prize pages for the 1994, 2005, 2007, 2012 and 2020 Economics Prizes were rechecked on 3 September 2026. RAND's catalogue record for Flood's 1952 memorandum and publisher or library metadata for the four *Go Deeper* works were checked on the same date. No current market statistic, regulation, software ranking or live institutional performance measure is used in the body.

The strongest empirical limitation is transportability. Laboratory and field results depend on incentives, framing, experience, culture, repeated contact and institutional rules. The most setting-specific evidence retained is the early ultimatum, centipede and tournament evidence, each identified by setting and used to demonstrate a mechanism or limit rather than a universal rate. The most contestable causal inference is the extent to which repeated-game structure explains observed cooperation, because norms, selection, enforcement and shared identity often change alongside repetition. Designed mechanisms are also evaluated only against declared objectives and within their legal, operational and participation conditions.

Bibliography

Primary and original research

Aumann, Robert J. "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics* 1, no. 1 (1974): 67-96.

Axelrod, Robert. *The Evolution of Cooperation*. New York: Basic Books, 1984.

Cournot, Augustin. *Researches into the Mathematical Principles of the Theory of Wealth*. Translated by Nathaniel T. Bacon. New York: Macmillan, 1897. First published in French in 1838.

Crawford, Vincent P., and Joel Sobel. "Strategic Information Transmission." *Econometrica* 50, no. 6 (1982): 1431-1451.

Flood, Merrill M. *Some Experimental Games*. Research Memorandum RM-789-1-PR. Santa Monica, CA: RAND Corporation, 1952.

Fudenberg, Drew, and Eric Maskin. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica* 54, no. 3 (1986): 533-554.

Gale, David, and Lloyd S. Shapley. "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69, no. 1 (1962): 9-15.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior & Organization* 3, no. 4 (1982): 367-388.

Harsanyi, John C. "Games with Incomplete Information Played by 'Bayesian' Players, I: The Basic Model." *Management Science* 14, no. 3 (1967): 159-182.

Harsanyi, John C. "Games with Incomplete Information Played by 'Bayesian' Players, II: Bayesian Equilibrium Points." *Management Science* 14, no. 5 (1968): 320-334.

- Harsanyi, John C. "Games with Incomplete Information Played by 'Bayesian' Players, III: The Basic Probability Distribution of the Game." *Management Science* 14, no. 7 (1968): 486-502.
- Hart, Sergiu, and Andreu Mas-Colell. "A Simple Adaptive Procedure Leading to Correlated Equilibrium." *Econometrica* 68, no. 5 (2000): 1127-1150.
- Henrich, Joseph, et al. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review* 91, no. 2 (2001): 73-78.
- Henrich, Joseph, et al. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment." *Science* 327, no. 5972 (2010): 1480-1484.
- Kreps, David M., and Robert Wilson. "Sequential Equilibria." *Econometrica* 50, no. 4 (1982): 863-894.
- Lemke, C. E., and J. T. Howson Jr. "Equilibrium Points of Bimatrix Games." *SIAM Journal on Applied Mathematics* 12, no. 2 (1964): 413-423.
- Maynard Smith, John, and George R. Price. "The Logic of Animal Conflict." *Nature* 246 (1973): 15-18.
- McKelvey, Richard D., and Thomas R. Palfrey. "An Experimental Study of the Centipede Game." *Econometrica* 60, no. 4 (1992): 803-836.
- Milgrom, Paul R., and Robert J. Weber. "A Theory of Auctions and Competitive Bidding." *Econometrica* 50, no. 5 (1982): 1089-1122.
- Nagel, Rosemarie. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review* 85, no. 5 (1995): 1313-1326.
- Nash, John F. "Equilibrium Points in N-Person Games." *Proceedings of the National Academy of Sciences* 36, no. 1 (1950): 48-49.
- Nash, John F. "The Bargaining Problem." *Econometrica* 18, no. 2 (1950): 155-162.

Nash, John F. "Non-Cooperative Games." *Annals of Mathematics* 54, no. 2 (1951): 286-295.

Rees, Michael A., et al. "A Nonsimultaneous, Extended, Altruistic-Donor Chain." *New England Journal of Medicine* 360, no. 11 (2009): 1096-1101.

Roth, Alvin E., and Elliott Peranson. "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." *American Economic Review* 89, no. 4 (1999): 748-780.

Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver. "Kidney Exchange." *Quarterly Journal of Economics* 119, no. 2 (2004): 457-488.

Shapley, Lloyd S. "A Value for n-Person Games." In *Contributions to the Theory of Games II*, edited by H. W. Kuhn and A. W. Tucker, 307-317. *Annals of Mathematics Studies* 28. Princeton, NJ: Princeton University Press, 1953.

Schelling, Thomas C. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press, 1980. First published 1960.

Selten, Reinhard. "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit." *Zeitschrift für die gesamte Staatswissenschaft* 121 (1965): 301-324 and 667-689.

Spence, Michael. "Job Market Signaling." *Quarterly Journal of Economics* 87, no. 3 (1973): 355-374.

Taylor, Peter D., and Leo B. Jonker. "Evolutionarily Stable Strategies and Game Dynamics." *Mathematical Biosciences* 40, nos. 1-2 (1978): 145-156.

Vickrey, William. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance* 16, no. 1 (1961): 8-37.

Von Neumann, John. "Zur Theorie der Gesellschaftsspiele." *Mathematische Annalen* 100 (1928): 295-320.

Von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. 60th anniversary ed. Princeton, NJ: Princeton

University Press, 2004. First published 1944.

Zermelo, Ernst. "Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels." In *Proceedings of the Fifth International Congress of Mathematicians*, vol. 2, 501-504. Cambridge: Cambridge University Press, 1913.

Modern works

Camerer, Colin F. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press and Russell Sage Foundation, 2003.

Dixit, Avinash K., and Barry J. Nalebuff. *The Art of Strategy: A Game Theorist's Guide to Success in Business and Life*. New York: W. W. Norton, 2008.

Leonard, Robert. *Von Neumann, Morgenstern, and the Creation of Game Theory: From Chess to Social Science, 1900-1960*. Cambridge: Cambridge University Press, 2010.

Maynard Smith, John. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press, 1982.

Osborne, Martin J. *An Introduction to Game Theory*. New York: Oxford University Press, 2004.

Osborne, Martin J., and Ariel Rubinstein. *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

Institutional sources

Royal Swedish Academy of Sciences. Official prize announcement and scientific background for the 1994 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, awarded to John C. Harsanyi, John F. Nash Jr. and Reinhard Selten. Rechecked 3 September 2026.

Royal Swedish Academy of Sciences. Official prize announcement and scientific background for the 2005 Sveriges Riksbank Prize in Economic

Sciences in Memory of Alfred Nobel, awarded to Robert J. Aumann and Thomas C. Schelling. Rechecked 3 September 2026.

Royal Swedish Academy of Sciences. “Mechanism Design Theory.” Scientific background for the 2007 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Rechecked 3 September 2026.

Royal Swedish Academy of Sciences. “Stable Allocations and the Practice of Market Design.” Scientific background for the 2012 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Rechecked 3 September 2026.

Royal Swedish Academy of Sciences. “Improvements to Auction Theory and Inventions of New Auction Formats.” Scientific background for the 2020 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Rechecked 3 September 2026.