Colman, A. M., & Stirk, J. A. (1998). Stackelberg reasoning in mixed-motive games: An experimental investigation. *Journal of Economic Psychology, 19*, 279-293. [Awarded the Citation of Excellence by ANBAR Electronic Intelligence]

STACKELBERG REASONING IN MIXED-MOTIVE GAMES: AN EXPERIMENTAL INVESTIGATION

Andrew M. Colman and Jonathan A. Stirk
Department of Psychology, University of Leicester, Leicester LE1 7RH, UK

Abstract

The Stackelberg heuristic is a simulation heuristic in which a player optimizes against best-reply counterstrategies, and a game is Stackelberg-soluble if the resulting Stackelberg strategies are in equilibrium. To test the hypothesis that players use this heuristic in Stackelberg-soluble games, 100 subjects played all 12 ordinally nonequivalent 2 × 2 games, nine of which (including Prisoner's Dilemma and Stag Hunt) were Stackelberg-soluble and three (including Battle of the Sexes and Chicken) were non-Stackelberg-soluble. Subjects significantly preferred Stackelberg strategies in Stackelberg-soluble games, and a protocol analysis of stated reasons for choice showed that joint payoff maximization and strategic dominance were given as reasons significantly more frequently in Stackelberg-soluble than in non-Stackelberg-soluble games.

Keywords: cognitive hypothesis testing, decision making, game theory, heuristic modeling, non zero sum games
PsycINFO classification 2340 Cognitive Processes
JEL classification C72 Noncooperative Games

1. Introduction

There are grounds for believing that, in certain classes of games at least, players tend to use a form of reasoning that Colman and Bacharach (1997) called the *Stackelberg heuristic*. This is a generalization of the "minorant" and "majorant" models that von Neumann and Morgenstern (1944, section 14.4.1, pp. 100-104) used to rationalize their solution of strictly competitive games. The assumption underlying the Stackelberg heuristic is that players reason with a type of simulation heuristic (Kahneman and Tversky 1982) in which both players choose strategies that maximize their own payoffs on the assumption that any choice will invariably be met by a counterstrategy that maximizes the co-player's payoff, as if the co-player could choose second in a perfect-information version of the game with foreknowledge of the first player's choice. For example, a player in a simultaneous-choice game such as the well-known Prisoner's Dilemma game may use the following type of Stackelberg reasoning:

> Suppose my fellow prisoner were able to move second, with the benefit of knowing what I had chosen to do. In that case, if I were to cooperate, then my fellow prisoner's selfish interest would be best served by responding with a defecting choice, and I would end up with the worst possible payoff. On the other hand, if I were to defect, then my fellow prisoner's self-interest would again be served by a defecting choice, but in this case I would get a slightly better payoff. It follows that to do the best for myself in the

sequential version of the game I should defect; therefore I'll defect in the actual simultaneous-choice version of the game.

Formally, in any two-person game $\Gamma = \langle S_i, H_i \rangle$, where S_i, $i \in \{1, 2\}$ is Player i's strategy set and H_i is a real-valued payoff function defined on the set $S = S_1 \times S_2$, Player i applies the Stackelberg heuristic as follows. First, Player i assumes that Player j can anticipate Player i's thinking about the game as if choosing second with perfect information of Player i's choice. This implies that for every strategy s_i in Player i's strategy set S_i, Player j, who is assumedly rational, responds with a strategy f such that $f(s_i)$ is invariably a best reply of Player j to s_i, so that

$$H_j(s_i, f(s_i)) \geq H_j(s_i, s_j) \quad \forall\, s_j \in S_j.$$

Second, assuming that both players are fully informed about the rules of the game and the payoff functions, Player i expects this to happen and chooses a payoff-maximizing strategy based on the assumption that Player j's response will invariably be a best reply $f(s_i)$. In other words, Player i chooses a counterstrategy for which $\max_i H_i(s_i, f(s_i))$ is attained. This is called Player i's Stackelberg strategy or h strategy. If in any game Γ the players' h strategies constitute a Nash equilibrium, then Γ is Stackelberg-soluble (or h-soluble) and the corresponding equilibrium is called a Stackelberg solution (or h solution). In a two-person game, a (Nash) equilibrium is an outcome in which each strategy is a payoff-maximizing best reply to the co-player's strategy.

There are several grounds for believing that the Stackelberg heuristic may influence the thinking of human decision makers. First, the Stackelberg heuristic relies on evidential reasoning (Eells 1985; Jeffrey 1983; Nozick 1969, 1993), a form of reasoning in which decisions are made by maximizing the conditional expected utilities of possible acts rather than the pure expected utilities as in classical decision theory, the utilities of outcomes being weighted by their probabilities given the specified acts, irrespective of whether there is any *causal* connection between the acts and the outcomes. There is experimental evidence to show that people do, in certain circumstances, use evidential forms of reasoning. For example, Quattrone and Tversky (1984) reported that subjects expressed a greater willingness to vote in an election when they believed that the outcome would depend on the proportion of like-minded voters who voted rather than on the behaviour of non-aligned voters, even though the effect of an individual's vote would be negligible (and equal) in both cases, and they predicted that their preferred candidate would be more likely to win the election if they themselves voted. This is a clear example of evidential reasoning.

Second, the Stackelberg heuristic is a form of simulation heuristic, by which people reason about events through an operation resembling the running of a simulation model, and Kahneman and Tversky (1982) have provided empirical evidence that people use simulation heuristics to analyse counterfactual propositions (e.g., "If only I had driven home by my regular route, I wouldn't have collided with the truck at the intersection") by mentally "undoing" events that have occurred and then running mental simulations of the events with the corresponding parameters of the model altered.

The third ground for believing that people use Stackelberg reasoning is the powerful psychological appeal of *payoff dominance* as a criterion for choosing between Nash equilibria in games, and the ability of experimental subjects to exploit payoff-dominant equilibrium points by focusing successfully on them in certain pure coordination games (Schelling 1960, chap. 3; Mehta et al. 1994a, 1994b) and in other games of common interests (Cooper *et al.* 1990, 1992a, 1992b). If e and f are any two distinct Nash equilibria in a two-person game, then e (strictly) payoff-dominates f iff $H_i(e) > H_i(f)$ for both players $i \in \{1, 2\}$. According to the payoff-dominance principle, in any game, if one equilibrium point e payoff-dominates all others, rational players will choose strategies corresponding to *e*. Harsanyi and Selten's (1988) general theory of equilibrium selection in games is based on this assumption (together with a slightly different principle of risk dominance), and most game theorists have accepted its intuitive force (e.g., Crawford and Haller 1990; Farrell 1987, 1988; Gauthier 1975; Lewis 1969; Sugden 1995).

Games with unique payoff-dominant Nash equilibria that are Pareto-optimal are called *games of common interests* (Aumann and Sorin 1989), and although in any such game the payoff-dominant equilibrium is the "obvious" choice, it has been argued that the payoff-dominance principle cannot be justified by conventional game-theoretic reasoning (Gilbert 1989, 1990). In a two-person common-interest game, for example, Player I has a reason to choose the strategy corresponding to the payoff-dominant Nash equilibrium only if there is a reason to expect Player II to do likewise; but Player II has a reason to choose it only if there is a reason to expect Player I to choose it, and we have an infinite regress that provides neither player with any reason for choosing the payoff-dominant equilibrium. Colman and Bacharach (1997) argued that the Stackelberg heuristic provides a clarification of the psychological appeal of payoff dominance and an explanation for the otherwise inexplicable ability of experimental subjects to solve certain types of pure coordination games in practice. Colman (1997) has shown that all pure coordination games are soluble by the Stackelberg heuristic with the help of Bacharach's (1993) variable-frame theory.

Colman and Bacharach (1997) proved that every game of common interests is Stackelberg-soluble or *h*-soluble, that its *h* solution is its payoff-dominant Nash equilibrium, and that in every game with Pareto-rankable Nash equilibria, a Stackelberg solution is a payoff-dominant Nash equilibrium. They also proved that there are *h*-soluble games that are not games of common interests and games with payoff-dominant Nash equilibria that are not *h* solutions. These last two theorems raise the empirical question of which classes of games elicit Stackelberg reasoning, assuming that people do indeed use Stackelberg reasoning to choose strategies in certain games.

The aim of the experiment described below is to examine the strategy choices of subjects in a wide range of simple two-person mixed-motive games and to determine whether Stackelberg strategies tend to be chosen. Each subject made a single strategy choice in each of twelve 2×2 mixed-motive games, nine of which were Stackelberg-soluble or *h*-soluble. We hypothesized that players would tend to use Stackelberg strategies in the *h*-soluble games.

2. Method

2.1. Subjects

The subjects were 100 undergraduate students at the University of Leicester, 53 male and 47 female. Their ages ranged from 18 to 43 years with a mean of 21.64 ($SD = 4.18$). They were recruited via posters on campus notice-boards and a mail message broadcast via the Internet to all computer users at the university. They were paid £1.00 for completing the experiment plus a variable additional sum ranging from a minimum of £1.20 to a maximum of £4.80 according to the payoffs that they accumulated in the twelve games during the half-hour testing session.

2.2. Payoff matrices

The strategic structures of games can be distinguished according to their ordinal payoff structures. In the case of 2×2 games, if each player's preferences among the four possible outcomes are rank-ordered from most preferred (with a payoff of 4) to least preferred (with a payoff of 1), then there are exactly 78 ordinally nonequivalent 2 × 2 games (Rapoport and Guyer 1966). Of these 78, twelve are symmetric in the sense of presenting both players with the same payoff function, so that each would face the same choice if the players' roles were reversed. The twelve ordinally nonequivalent 2 × 2 games used in this study (see Figure 1) were generated by inserting the numbers 4, 3, 2, and 1 into the four cells of a 2 × 2 matrix in every possible order or permutation, constructing a symmetric game for each permutation, and deleting games that turned out to be identical to others another apart from the (strategically irrelevant) labelling of their rows or columns.

Figure 1 about here

Each payoff matrix was printed on a separate sheet, with the row-player's strategy labels and payoffs printed in red and the column-player's in blue, and the twelve matrices were presented to the subjects in random order. Of the twelve games used in the experiment, nine (Games 1, 4, 5, 7, 8, 9, 10, 11, and 12 in Figure 1) turned out to be Stackelberg-soluble, and the remaining three (Games 2, 3, and 6) were non-Stackelberg-soluble.

2.3. Design and procedure

A within-subjects experimental design was used, with each subject making a single strategy choice in each of the twelve games. Subjects were tested in small even-numbered groups, and throughout the testing session the word "game" was avoided in favour of "decision-making task". After being split into pairs, the subjects were given the following written instructions:

> You have just been split up into pairs and you are each going to have to choose between two alternatives, *A* or *B*. You have each been given 12 grids. Your objective is to maximize the number of points that you win for yourself, and therefore the amount of money that you will receive. The numbers in the grid are the points you win. You will be given £1 for completing all 12 grids and, on top of that, 10 pence per point scored. Each person must decide whether to choose *A* or *B*. You will do this without knowing what your partner has chosen. After each choice, explain clearly but briefly (a sentence or two at most) why you made the choice that you made. Enter your choice and explanation into the table on the next sheet.

Each member of a pair was arbitrarily assigned the role of row-player (labelled Red) or column-player (labelled Blue). The pair was then told that Red would choose between rows *A* and *B* and Blue between columns *A* and *B*. It was explained to them that their payoffs would depend on the choices made by both members of the pair, who would write down their choices out of sight of each other. It was further explained that the pair of numbers in each cell of the matrix represented the participants' payoffs — the first number the payoff to Red and the second the payoff to Blue. They were reminded that each unit of payoff represented 10 pence, which would be added to their payment after completion of all twelve decisions.

The subjects were asked to supply demographic information (age and sex) on the front page of their answer booklets and were then shown a sample game based on a payoff matrix that was not used in the actual experiment. Their comprehension of the rules and payoffs was tested by asking each of them to write down in their answer booklets what would happen in a specified hypothetical outcome of this game. They were invited to ask for clarification of anything they did not understand, and the experiment proper was not begun until the experimenter was satisfied that the subjects understood what they had to do and precisely how their joint strategy choices would affect their payoffs.

The subjects in each pair were then seated back to back and were invited to examine the first payoff matrix (which was labelled Grid 1). They were asked to record their choices (*A* or *B*) opposite the label of Grid 1 in their answer booklets, where they were also asked to write their brief explanations for their choices. After 60 seconds, they were told that they had 30 seconds remaining on that grid, and after 90 seconds they were asked to examine the second game (Grid 2). The procedure was repeated for each of the remaining games. No feedback regarding the co-player's choices or points accumulated was given until all twelve games had been played.

3. Results

3.1. Quantitative results

The frequencies of Stackelberg (*h*) strategy choices in the twelve games are shown in Table 1, together with the results of chi-square tests. The purpose of the chi-square tests was to evaluate the significance of departures from chance in the relative frequencies of Stackelberg (*h*) and non-Stackelberg (non-*h*) strategy choices, the main hypothesis being that there would be a significant bias towards *h* strategies in each of the *h*-soluble games.

Table 1 about here

It is clear from Table 1 that there was a highly significant bias towards Stackelberg (*h*) strategies in all nine of the Stackelberg-soluble games. The relative frequencies with which Stackelberg strategies were chosen in these games were strikingly high, usually over 90 of the 100 subjects and never fewer than 78/100 choosing the *h* strategies. The effect sizes were determined using Cohen's (1988, 1992) index $w = \sqrt{(\chi^2/N)}$, which is designed to provide an indication of the magnitude of any departure from a chance baseline. Following Cohen's suggestion, we classified $w \geq .50$ as a large effect size, $w \geq .30$ as medium, and $w \geq .10$ as small. Using this index, the effect sizes in the Stackelberg-soluble games were all large, ranging from $w = .56$ in Game 12 to $w = .96$ in Game 1. These findings strongly confirm the main hypothesis of the experiment.

The non-Stackelberg-soluble games yielded much smaller effect sizes, and the direction of bias was mixed: in one case (Game 2) there was a significant bias (72/100) towards the Stackelberg strategy, with a medium effect size; in another (Game 3) there was a significant bias (70/100) towards the non-Stackelberg strategy, with a medium effect size; and in the third (Game 6) there was no significant bias in either direction, with 47/100 of subjects choosing the Stackelberg strategy and a negligible effect size.

3.2. Qualitative results

A simple protocol analysis was performed on the subjects' stated reasons for their choices, which were recorded immediately after each decision. An examination of these reasons by two raters resulted in a rough classification into the following nine categories (illustrative verbatim examples from the subjects' protocols are shown in parentheses):

1. Expected utility maximization: choosing a strategy that maximizes the average or expected payoff ("The average of the two blue B numbers is higher than the A"; "Because 2 + 4 is better than 1 + 3").

2. Joint payoff maximization: choosing a strategy that maximizes the total payoff of the pair ("Most points for both"; "Mutually beneficial").

3. Strategic dominance: choosing a strategy that yields a better payoff than the other

strategy irrespective of the co-player's choice ("Because I'll get more points for B whether he chooses A or B"; "Because I get more points whatever she chooses").

4. Sequential reasoning or mind-reading: choosing a strategy on the basis of a guess or inference about the co-player's likely choice ("Because it is the highest payoff for me if player II chooses B which I believe she will as A is too risky for her"; "Red [the other player] will choose B as their minimum score will be 2. Blue will score 2 or 3").

5. Relative payoff maximization: choosing a strategy with the aim of beating the co-player ("Cannot lose, can only be equal or better [than the other player]"; "The points are either equal to or higher than the blue points").

6. Minimax: choosing a strategy that offers the best of the worst possible payoffs ("To avoid scoring 1"; "Prefer to guarantee two points than risk getting only one if Red chooses A").

7. Maximax: choosing a strategy that provides the possibility of receiving the highest possible payoff in the game ("4 pts is available"; "I might get 4 points").

8. Individual payoff maximization: choosing a strategy with the aim of maximizing personal payoff irrespective of other considerations (e.g. "Carries the highest possible scores for me"; "Maximum personal benefit. No point in competing").

9. Ambiguous or unclassifiable: obscure, indefinite, enigmatic, and irrelevant ("Balanced, equal in that square"; "With B I can get either 3 or 4 points, with A it's 1 or 2").

Table 2 about here

In the relatively few cases in which subjects gave two or more distinct reasons for a choice, the reason that was judged to be most strongly emphasized was counted, and where two or more reasons seemed to be equally emphasized, only the first was counted. The distribution of reasons that emerged from the protocol analysis is shown in Table 2.

Across games, the most frequent reason for choice, by a slim margin, was joint payoff maximization (Reason 2). However, an examination of Table 2 shows that this reason was almost never given for choices in non-*h*-soluble games. This is to be expected, because it is only in games of common interests, which Colman and Bacharach (1997) proved are necessarily Stackelberg-soluble, that the notion of joint payoff maximization makes sense. Some of the other reasons are also differentially applicable across games for theoretical reasons. In particular, expected utility maximization (Reason 1) is indecisive in games 3, 6, 8, and 12, because in those games both strategies yield the same expected utility if equal probabilities are assigned to the co-player's choices (on the principle of insufficient reason); and strategic dominance does not apply to games 2, 3, 6, 8, 11, and 12, because those games do not have dominant strategies.

Our main concern was with the differences between *h*-soluble and non-*h*-soluble games. We therefore performed a series of chi-square goodness-of-fit tests to evaluate the significance of any differences in the frequencies of reasons for choice between *h*-soluble and non-*h*-soluble games, correcting the expected frequencies for the fact that three-quarters of the games were *h*-soluble. The results showed, first of all, that there were no significant differences between *h*-soluble and non-*h*-soluble games in the frequencies of relative payoff maximization (Reason 5), individual payoff maximization (Reason 8), and ambiguous or unclassifiable responses (Reason 9). Two reasons were given significantly more frequently in *h*-soluble than non-*h*-soluble games, namely joint payoff maximization (Reason 2), $\chi^2(1) = 46.49$, $p < .0001$, and strategic dominance (Reason 3), $\chi^2(1) = 33.31$, $p < .0001$. Four reasons were given significantly more frequently in non-*h*-soluble than *h*-soluble games: expected utility maximization (Reason 1), $\chi^2(1) = 12.37$, $p < .0004$; sequential reasoning or mind-reading (Reason 4), $\chi^2(1) = 15.62$, $p < .0001$; minimax (Reason 6), $\chi^2(1) = 35.70$, $p < .0001$; and maximax (Reason 7), $\chi^2(1) = 24.35$, $p < .0001$.

4. Discussion

The results of this experiment confirm the hypothesis that players tend to choose Stackelberg strategies in Stackelberg-soluble games. In all nine of the Stackelberg-soluble games, the subjects manifested large and highly significant biases towards the Stackelberg strategies.

Six of the twelve games used in the experiment (Games 1, 4, 7, 8, 11, and 12) were games of common interests (Aumann and Sorin 1989), and all of them were Stackelberg-soluble. Partly overlapping with the common-interest games were six (Games 1, 4, 5, 7, 9, 10) in which both players had strictly dominant strategies (a strictly dominant strategy being one that yields a higher payoff than the other strategy irrespective of the response of the co-player), and all of these were Stackelberg-soluble. Included in this dominant-strategy group was Game 10, the familiar Prisoner's Dilemma game, in which 79 per cent of subjects chose Stackelberg strategies, Game 4, Harmony (90 per cent), and Game 9, Deadlock (97 per cent). In all of the games with dominant strategies, the Stackelberg strategies and dominant strategies coincided, so it is unclear whether they were chosen because of their Stackelberg property or their dominance property. In these games the subjects' choices, though consistent with the hypothesis of Stackelberg reasoning, provide inconclusive evidence of it.

The evidence is more persuasive in the remaining Stackelberg-soluble games (Games 8, 11, and 12) in which choices were not influenced by considerations of strategic dominance. In Game 8, the Stag Hunt game, 79 per cent subjects chose Stackelberg strategies, and in Games 11 and 12 the proportions were 97 per cent and 78 per cent respectively. All of the corresponding effect sizes were large, and these results suggest that the subjects were indeed influenced by Stackelberg reasoning. Further indirect evidence for this emerges from the protocol analysis of the subjects' stated reasons for choice (Table 2). In Games 8, 11, and 12 the predominant reasons for choice were joint payoff maximization and sequential reasoning or mind-reading. However, chi-square analysis revealed that, although joint payoff maximization was mentioned as a reason significantly more frequently in *h*-soluble than non-*h*-soluble games, the reverse was true for sequential reasoning or mind-reading, a puzzle that we shall discuss below.

The remaining three games were non-Stackelberg-soluble, and had no direct bearing on the main hypothesis, but the behaviour of the subjects in these games was of some interest none the less. All three were characterized by double, asymmetric Nash equilibria (see Colman 1995, chap. 6 for a detailed analysis of their strategic properties). In Game 2, Battle of the Sexes, the subjects showed a significant tendency to avoid the altruistic *B* strategies and to choose the Stackelberg *A* strategies, which do not intersect in a Nash equilibrium point. In Game 3, the game of Chicken, subjects showed a significant avoidance of the Stackelberg *B* strategies, which are both dangerous and out of equilibrium, and a preference for the cautious *A* strategies, which are also out of equilibrium. Finally, in Game 6, the game of Leader, almost equal numbers of subjects chose the non-equilibrium Stackelberg *B* strategies and the non-equilibrium *A* strategies.

There is no particular reason to expect players to choose Stackelberg strategies that do not intersect in Nash equilibria. The famous indirect argument of von Neumann and Morgenstern (1944, section 17.3, pp. 146-148), developed further by Luce and Riaffa (1957, pp. 63-65), implies that if there is a unique solution to a game, then it must be what is nowadays called a Nash equilibrium. The gist of the indirect argument is that, if the players are rational payoff-maximizers and are fully informed about the game, and if they know that their co-players are also rational and fully informed, then they can anticipate each other's thinking through what Bacharach (1987) has called the "transparency of reason", and each will choose a best reply to the other's (rightly) anticipated strategy, and because by definition a Nash equilibrium consists of strategies that are best replies to each other, it follows that a unique rational solution must invariably be a Nash equilibrium. This suggests that if rational players are influenced by Stackelberg reasoning, such reasoning should lead them to choose Stackelberg strategies in games that are Stackelberg-soluble. The findings reported in this article corroborate that hypothesis strongly for simple 2×2 games, but further evidence from a wider range of games is required before a more general conclusion can be drawn about Stackelberg reasoning.

The protocol analysis of the subjects' stated reasons for their choices showed that joint payoff maximization and strategic dominance were given as reasons significantly more frequently in *h*-soluble than in non-*h*-soluble games, and that expected utility maximization, sequential reasoning or mind-reading, minimax, and maximax were given significantly more frequently in non-*h*-soluble than in *h*-soluble games. The disproportionate frequency of joint payoff maximization and dominance as reasons for choice in *h*-soluble games is explained by the fact that, in contrast to the non-*h*-soluble games, all of the *h*-soluble games were either games of common interest or games with dominant strategies. Joint payoff maximization and dominance are powerful and persuasive reasons for choice in games in which they are applicable. It may seem surprising, at first, that sequential reasoning or mind-reading were given as reasons significantly more frequently in non-*h*-soluble than in *h*-soluble games, because these forms of reasoning are characteristic of Stackelberg reasoning, which is most obviously applicable to *h*-soluble games. However, this category included all reasons based on guessing or inferring what the other player was likely to choose, and in hindsight it seems obvious that players were likely to be driven by default to these forms of speculation in games in which more cogent considerations of common interests and strategic dominance were lacking. This may also explain why expected utility maximization, minimax, and maximax were given as reasons for choice significantly more frequently in non-*h*-soluble games.

Acknowledgements

The research reported in this article was supported by Grant No. L122251002 from the Economic and Social Research Council as part of the Economic Beliefs and Behaviour programme. A brief version was presented at the 21st annual conference of the International Association for Research in Economic Psychology (IAREP) in Paris on 13 September 1996. We are grateful to Michael Bacharach, Colin Camerer, and George Lowenstein for suggestions and advice during the planning of the experiment, to Robin Cubitt for helpful comments on an earlier draft of the article, to two anonymous referees for useful suggestions, to Zazie Todd for suggesting the protocol analysis, and to Florence Gaunet and Bénédicte Nanty for technical assistance.

References

Aumann, R. J. and Sorin, S., 1989. Cooperation and bounded recall. Games and Economic Behavior 1, 5-39.

Bacharach, M., 1987. A theory of rational decision in games. Erkenntnis 27, 17-55.

Bacharach, M., 1993. Variable universe games. In: K. Binmore, A. Kirman, and P. Tani (eds), Frontiers of Game Theory (pp. 255-275). Cambridge, MA: MIT Press.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences (2nd edn). Hillsdale, NJ: Erlbaum.

Cohen, J., 1992. A power primer. American Psychologist 112, 155-159.
Colman, A. M., 1995. Game Theory and its Applications in the Social and Biological Sciences (2nd edn). Oxford: Butterworth-Heinemann.
Colman, A. M., 1997. Salience and focusing in pure coordination games. Journal of Economic Methodology, 4, 61-81.
Colman, A. M. and Bacharach, M., 1997. Payoff dominance and the Stackelberg heuristic. Theory and Decision, 43, 1-19.
Cooper, R. W., DeJong, D. V., Forsythe, R., and Ross, T. W., 1990. Selection criteria in coordination games: Some experimental results. American Economic Review 80, 218-233.
Cooper, R. W., DeJong, D. V., Forsythe, R., and Ross, T. W., 1992a. Communication in coordination games. Quarterly Journal of Economics 107, 739-771.
Cooper, R. W., DeJong, D. V., Forsythe, R., and Ross, T. W., 1992b. Forward induction in coordination games. Economics Letters 40, 167-172.
Crawford, V. P. and Haller, H., 1990. Learning how to cooperate: Optimal play in repeated coordination games. Econometrica 58, 571-595.
Eells, E., 1985. Causality, decision, and Newcomb's paradox. In: R. Campbell and L. Sowden (eds), Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem (pp. 183-213). Vancouver: University of British Columbia Press.
Farrell, J., 1987. Cheap talk, coordination, and entry. Rand Journal of Economics 18, 34-39.
Farrell, J., 1988. Communication, coordination and Nash equilibrium. Economics Letters 27, 209-214.
Gauthier, D., 1975. Coordination. Dialogue 14, 195-221.
Gilbert, M., 1989. Rationality and salience. Philosophical Studies 57, 61-77.
Gilbert, M., 1990. Rationality, coordination and convention. Synthese 84, 1-21.
Harsanyi, J. C. and Selten, R., 1988. A General Theory of Equilibrium Selection in Games. Cambridge, MA: MIT Press.
Jeffrey, R., 1983. The Logic of Decision (2nd ed.). Chicago: University of Chicago Press.
Kahneman, D. and Tversky, A., 1982. The simulation heuristic. In: D. Kahneman, P. Slovic, and A. Tversky (eds), Judgment Under Uncertainty: Heuristics and Biases (pp. 201-208). Cambridge: Cambridge University Press.
Lewis, D. K., 1969. Convention: A Philosophical Study. Cambridge, MA: Harvard University Press.
Luce, R. D. and Raiffa, H., 1957. Games and Decisions: Introduction and Critical Survey. New York: Wiley.
Mehta, J., Starmer, C., and Sugden, R., 1994a. Focal points in pure coordination games: An experimental investigation. Theory and Decision 36, 163-185.
Mehta, J., Starmer, C., Sugden, R., 1994b. The nature of salience: An experimental investigation of pure coordination games. American Economic Review 84, 658-673.
Nozick, R. 1969. Newcomb's problem and two principles of choice. In: N. Rescher (ed.), Essays in Honor of Carl G. Hempel: A Tribute on His Sixty-Fifth Birthday (pp. 114-146). Dordrecht-Holland: D. Reidel.
Nozick, R. 1993. The Nature of Rationality. Princeton, NJ: Princeton University Press.
Quattrone, G. A. and Tversky, A., 1984. Causal versus diagnostic contingencies: On self-deception and the voter's illusion. Journal of Personality and Social Psychology 46, 237-248.
Rapoport, A. and Guyer, M., 1966. A taxonomy of 2×2 games. General Systems 11, 203-214.
Schelling, T. C., 1960. The Strategy of Conflict. Cambridge, MA: Harvard University Press.

Sugden, R., 1995. A theory of focal points. Economic Journal 105, 533-550.
von Neumann, J. and Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton, NJ: Princeton University Press.

Table 1
Stackelberg (*h*) strategy choices in 2 × 2 games ($N = 100$)

Game	*h*	χ^2	*p* <	Effect size $w^\dagger$	Bias towards
1*	98	92.17	.001	.96	*h* strategy
2	72	19.37	.001	.44	*h* strategy
3	30	16.01	.001	.40	non-*h* strategy
4*	90	64.01	.001	.80	*h* strategy
5*	95	81.01	.001	.90	*h* strategy
6	47	0.37	.548	.06	*ns*
7*	98	92.17	.001	.96	*h* strategy
8*	79	33.65	.001	.58	*h* strategy
9*	97	88.37	.001	.94	*h* strategy
10*	79	33.65	.001	.58	*h* strategy
11*	97	88.37	.001	.94	*h* strategy
12*	78	31.37	.001	.56	*h* strategy

**h*-soluble games
$^\dagger w \geq .50$ large, $w \geq .30$ medium, $w \geq .10$ small

Table 2
Distribution of reasons for choice across games

	Reasons for choice†								
Game	1	2	3	4	5	6	7	8	9
1*	1	29	30	13	4	2	8	7	6
2	11	4	1	39	6	17	12	2	8
3	5	14	0	26	4	22	10	1	18
4*	5	46	3	18	4	7	4	1	12
5*	0	12	25	21	3	10	1	1	27
6	1	4	0	30	4	23	13	0	25
7*	1	32	11	20	4	5	1	1	25
8*	0	28	0	33	4	6	5	0	24
9*	0	8	29	9	4	8	1	1	40
10*	3	10	13	23	5	16	5	0	25
11*	4	43	1	16	6	7	3	2	18
12*	2	39	0	19	7	7	6	2	18
Total	33	269	113	267	55	130	69	18	246

**h*-soluble games
† See text for descriptions of the reasons

Fig. 1. Ordinally nonequivalent *h*-soluble and non-*h*-soluble symmetric 2 × 2 games, showing Nash equilibria and *h* strategies.

	A*	B
A*	4, 4	3, 2
B	2, 3	1, 1

Game 1†

	A*	B
A*	2, 2	4, 3
B	3, 4	1, 1

Game 2

	A	B*
A	3, 3	2, 4
B*	4, 2	1, 1

Game 3

	A*	B
A*	4, 4	2, 3
B	3, 2	1, 1

Game 4†

	A*	B
A*	3, 3	4, 2
B	2, 4	1, 1

Game 5†

	A	B*
A	2, 2	3, 4
B*	4, 3	1, 1

Game 6

	A*	B
A*	4, 4	3, 1
B	1, 3	2, 2

Game 7†

	A*	B
A*	4, 4	1, 3

B	3, 1	2, 2

Game 8†

	A*	B
A*	3, 3	4, 1
B	1, 4	2, 2

Game 9†

	A	B*
A	3, 3	1, 4
B*	4, 1	2, 2

Game 10

	A*	B
A*	4 4	2, 1
B	1, 2	3, 3

Game 11†

	A*	B
A*	4, 4	1, 2
B	2, 1	3, 3

Game 12†

†*h*-soluble games
**h*-strategies
The first number in each cell is the payoff to the row player; the second is the payoff to the column player. Shaded cells are Nash equilibria.