

A696

ENCYCLOPEDIA OF COGNITIVE SCIENCE

2000

©Macmillan Reference Ltd

## Coordination Games

Coordination problems, experiments, communication, battle -of-sexes game, stag-hunt game, risk dominance, potential

Goeree, Jacob K

Jacob K Goeree

University of Virginia, Charlottesville, Virginia USA

Holt, Charles A

Charles A Holt

University of Virginia, Charlottesville, Virginia USA

**Coordination games typically possess multiple Nash equilibria, some of which are preferred by one or more players. Coordination on desired equilibria can be facilitated by communication, repetition, and introspection.**

### Introduction

Despite the well-known efficiency properties of competitive markets, economists have long been concerned that there may be multiple equilibria in complex social systems. The possibility of becoming mired in a bad equilibrium dates back at least to Thomas Malthus' notion of a "general glut." The problem is one of coordination: e.g. it may not make sense for me to go to the market to trade if you are going to stay home on the farm. This situation is one of common interest in that everyone prefers the equilibrium with high market activity to the one without any trade. Coordination problems can also arise when there is a conflicting interest, i.e. when each person prefers a different equilibrium outcome. The classic example is the "battle-of-the-sexes" game, where the man prefers to attend a baseball game and the women prefers to attend an opera, but both would rather do something together than go to separate events. In all of these examples there are multiple outcomes that are equilibria in the sense that no single individual has an incentive to deviate if others are conforming to that outcome. For instance, in the battle-of-the-sexes, the man would attend the opera

if he thinks the woman will be there even though he prefers the other equilibrium outcome in which both attend the baseball game.

Society has developed a number of ways to solve such coordination problems. The most obvious cases involve social norms and rules, e.g. driving on the left side of the road in the UK and Japan. In the absence of explicit rules, individuals may rely on focalness, e.g. Schelling (1980) points out that two people that are supposed to meet each other somewhere in New York city, are likely to go to Times Square. Historical accidents and precedents can also be used as coordination devices since past experience often affects beliefs about others' future behaviour. For example, prior discrimination against workers of a particular racial background may produce a situation in which the workers don't invest in skills because they anticipate unfavourable job assignments. These beliefs can be self-fulfilling since the workers' choices make it rational for the employer to expect low performance from those workers. In this case, a period of "reversed" discrimination, e.g. affirmative action, may be needed to change expectations and allow coordination on the preferable merit-based equilibrium. Coordination can also be achieved through explicit communication and reciprocity, e.g. baseball this weekend and opera next.

Economists and psychologists are also interested in strategic settings where opportunities for communication and reciprocity are limited perhaps due to the large number of people involved. Rapoport et al. (1998), for example, consider a large group of sellers who must decide whether or not to enter a particular market that is profitable only in the absence of excessive entry. Even without communication, low profits will drive some sellers away while high profits will attract more sellers, and observed behaviour in laboratory experiments shows that the costs of uncoordinated outcomes can be negligible. Laboratory experiments such as these have uncovered interesting patterns of behaviour that have led to theoretical models of equilibrium selection. These models of learning and introspection provide natural extensions of classical equilibrium concepts, which are unable to predict which outcome has the strongest drawing power.

### **Types of Coordination Games**

A typical coordination game with conflicting interests is the battle-of-the-sexes game where the preferred equilibrium provides a player with a high payoff of 600, while the less preferred equilibrium only yields 200. This game shown in Table 1, was used by Cooper et al. (1989) in a laboratory experiment that was presented in neutral terms. There are two obvious Nash equilibria of this game: (Baseball, Baseball) and (Opera, Opera). In addition, there is an equilibrium involving randomisation, where both the man and woman choose their preferred activity with probability 0.75. In this manner, each person is indifferent between the two options since each yield a payoff of 150 on average ( $0.75 * 200 = 0.25 * 600 = 150$ ).

		Man	
		Opera	Baseball
Woman	Baseball	0, 0	200, 600
	Opera	600, 200	0, 0

Table 1. A Battle-of-the Sexes Game (Woman’s payoff , Man’s payoff)

Another coordination game with opposite interests is the market entry game shown in Table 2. Two firms have to decide whether or not to incur a cost of 50 to enter a market. The entrants profits will be 150 if it is alone, but will be zero if both firms enter. The profits from staying out are zero independent of the rival’s choice. Again there are two obvious Nash equilibria and each firm prefers the outcome in which it is the sole entrant. Notice that the case of excess entry produces losses for each firm, which makes this a “game of chicken.” This terminology is inspired by the movie *Rebel without a Cause*, where James Dean and his competitor drive toward a cliff and the first to stop is considered to be “chicken.” Each prefers the outcome where the other one stops first, but if neither stops both incur a severe cost (possibly worse than -50).

		Firm 2	
		Stay Out	Enter
Firm 1	Stay Out	0, 0	0, 100
	Enter	100, 0	-50, -50

Table 2. A Market Entry Game (Firm 1’s payoff , Firm 2’s payoff)

Rousseau's classic “stag hunt” game describes a situation in which two hunters decide to hunt for stag or hare. Each hunter alone could be sure of bagging a hare, but both hunters are needed to corner the stag, which is the preferred game. If only one hunts stag, that person is left empty handed, so there is a low-payoff equilibrium in which both hunt hare and a high-payoff equilibrium in which both hunt stag. Unlike the two previous examples, this game is one of common interests since both prefer the “better” outcome. Rousseau’s game has a “weakest link” property since the stag will escape through any sector left unguarded by a hunter. This is analogous to a situation where workers provide different parts to be assembled, and the final product requires all parts. In this case, the total amount produced is determined by the slowest worker, i.e. the weakest link in the production chain.

One way to model a weakest-link, or minimum-effort coordination game is to let players choose effort levels, 1 or 2, where each unit of effort results in a cost of  $c < 1$ . The output per person is determined by the minimum of the effort levels chosen. For example, if both choose an effort level of 1, then each player receives a payoff of  $1-c$ , as shown in the upper-left box of Table 3. Similarly, when both choose a high effort level of 2, they each obtain  $2-2c$ . But if they fail to coordinate with effort choices of 1 and 2, then the minimum is 1 and payoffs are  $1-c$  for the low-effort individual and  $1-2c$  for the high-effort individual. Notice that the low effort outcome is an equilibrium since a costly increase in effort by only one person will not raise the amount produced. The high effort outcome is also an equilibrium, since a reduction in effort by only one person will lower the minimum by more than the cost savings  $c$ .

		Player 2's Effort	
		1	2
Player 1's Effort	1	1 - c, 1 - c	1 - c, 1 - 2c
	2	1 - 2c, 1 - c	2 - 2c, 2 - 2c

Table 3. A 2H2 Coordination Game (Player 1's payoff, Player 2's payoff)

The game in Table 3 can be generalized to allow for more than two players and multiple effort levels, with payoffs determined by the minimum effort level chosen. If a player's effort is denoted by  $e_i, i = 1, \dots, n$ , payoffs are:

$$p_i(e_1, \dots, e_n) = \min\{e_1, \dots, e_n\} - c e_i, \quad i = 1, \dots, n, \tag{1}$$

where  $c$  is the per-unit effort cost. As long as  $c$  is less than 1, payoffs are maximized when all players choose the highest possible effort. Note, however, that *any* common effort level constitutes a Nash equilibrium, since a costly unilateral increase in effort will not raise the minimum, and a unilateral decrease will reduce the minimum by more than the cost when  $c < 1$ . This argument does not depend on the number of players, so non-critical changes in  $c$  and  $n$  will not alter the set of Nash equilibria, despite the reasonable expectation that efforts should be high for sufficiently low effort costs and low numbers of participants.

### Experimental Evidence

Van Huyck et al. (1989) used a minimum-effort structure in one of the most widely cited and influential game theory experiments. In their experiment subjects could choose seven possible effort levels, 1 through 7, and payoffs were a linear function of the difference between the minimum effort and one's own divided by two ( $c = 1/2$ ). Thus there are seven equilibria, in which all players choose the same effort level but the equilibrium with the highest payoffs is the one where all players choose an effort of 7. The experiment involved 14-16 players who made independent effort choices. After choices were collected and the minimum was announced, payoffs were calculated. This whole process was repeated for ten rounds. Van Huyck et al. report that efforts declined dramatically, with the final-period efforts clustered at the equilibrium that is *worst* for all. This result surprised many game theorists, who were comfortable assuming that rational individuals would be able to coordinate on an outcome that is best for all. Van Huyck et al. also find that an extreme reduction in the cost of effort ( $c = 0$ ) results in an overwhelming number of high effort decisions, which is not surprising since raising effort is costless.

Goeree and Holt (2000) explore the effects of changes in effort cost and the number of players more systematically. They report two and three-player minimum-effort coordination game experiments with varying effort costs. Individuals could choose continuous effort levels in the range from 110 to 170 with payoffs determined as in (1). With two players, average effort levels are initially around the midpoint 140 in both the low-cost ( $c = 1/4$ ) and high-cost ( $c = 3/4$ ) treatments. By the third period,

however, there is a strong separation, with higher efforts for the low-cost treatment. In the final rounds, the average effort was 159 in the low-cost treatment and 126 in the high-cost treatment. Even though *any* common effort in the range from 110 to 170 is a Nash equilibrium independent of the effort cost, the observed behaviour is affected by the magnitude of the effort cost in an intuitive manner. Likewise, Goeree and Holt (2000) find that an increase in the number of players tends to decrease the final-period effort levels.

A different line of experimentation involves factors that facilitate coordination. Cooper et al. (1989) investigate the effects of “cheap talk” forms of communication in a battle-of-the-sexes game in Table 1. Subjects first submitted a message about which choice they intended to make. These messages were non-binding in the sense that a player could deviate from their reported intent after seeing the other’s message. When both messages matched one of the equilibria, i.e. (Baseball, Baseball) or (Opera, Opera), then actual decisions corresponded to stated intentions more than 80 percent of the time. When messages differed, however, individuals tended to deviate from their original message and choose their preferred activity about 71 percent of the time. Communication can therefore facilitate coordination, but when communication fails the behaviour corresponds more closely to the equilibrium in randomised strategies (choosing the preferred activity with probability 0.75).

The Cooper et al. (1989) experiments involved random pairings of subjects in each period. In contrast, fixed pairings permit coordination that is based on the history of past decisions. Prisbrey (1991) finds a common pattern of alternating choices, with outcomes (Baseball, Baseball) and (Opera, Opera) in successive rounds. (As in the other papers mentioned above, Prisbrey used neutral terminology in presenting the payoffs to subjects.) This alternation can be interpreted as a form of reciprocity, where “nice” behaviour in one round is rewarded by a nice response in the next. A number of other coordination experiments are surveyed in Ochs (1995).

## Theoretical Explanations

One of the most commonly suggested criteria for the analysis of games with multiple equilibria is to select the one with the highest payoffs for all, if such a “Pareto-dominant” outcome exists. The Van Huyck et al. (1989) experiment showed that this method is incorrect, and the Goeree and Holt (2000) experiment showed that any explanation must take into account the effort cost and the number of players.

Harsanyi and Selten's (1988) notion of risk dominance is sensitive to the effort cost that determines the losses associated with deviations from best responses to others' decisions. To illustrate the concept of risk dominance, consider the two-person minimum-effort game shown in Table 3. When both players are choosing efforts of 1, the cost of a unilateral deviation to 2 is just the cost of the extra effort,  $c$ , which will be referred to as the “deviation loss.” Similarly, the deviation loss at the (2,2) equilibrium is  $1-c$ , since a unilateral reduction in effort reduces the minimum by 1 but saves the marginal effort cost  $c$ . The deviation loss from the low-effort equilibrium is greater than that for the high-effort equilibrium if  $c > 1 - c$ , or equivalently, if  $c > 1/2$ , in which case we say that the low-effort equilibrium is risk dominant. Risk dominance, therefore, has the desirable property that it selects the low-effort outcome if the cost of effort is sufficiently high.

There is, however, no consensus on how to generalize risk dominance for games with more players, a continuum of decisions, etc. A related concept that does generalize is the notion of maximization of a “potential” of a game. Loosely speaking, the idea behind potential is to find a function for a game that is maximized by a Nash equilibrium for that game. Stated differently, a potential function is a mathematical formula that is positively related to individual players' payoffs: when a change in a player's own decision raises that player's payoff, then this change necessarily raises the value of the potential function by the same amount, and vice versa for decreases. If such a potential function exists for the game, then each person trying to increase their own payoff may produce a group result that maximizes the potential function for the game as a whole. Think of two people holding adjacent sides of a treasure box, with one pulling uphill along the East-West direction and the other pulling uphill along the North-South axis. Even though each person is only pulling in one direction, the net effect will be to take the box to the top of the hill, where there is no tendency to change (a Nash equilibrium that maximizes potential).

For instance, for the  $n$ -player minimum effort game given in (1), the potential function is simply the common production function that determines a *single* player's payoff, minus the sum of *all* players' effort costs:

$$V(e_1, \dots, e_n) = \min\{e_1, \dots, e_n\} - c \sum_{i=1}^n e_i \quad (2)$$

The maximization of potential will obviously require equal effort levels. At any common effort,  $e$ , the potential in (2) becomes:  $V = e - nce$ , which is maximized at the lowest effort when  $nc > 1$ , and is maximized at the highest effort when  $nc < 1$ . In two-person games, this condition reduces to the risk dominance comparison of  $c$  with  $1/2$ .

The notion of potential can be used to evaluate results from previous laboratory experiments. Van Huyck et al. (1990) conducted games with 14 to 16 players and an effort cost of either 0 or  $1/2$ , so  $nc$  was either zero or about seven. Compared to the critical  $nc$  value of 1, these parameter choices appear rather extreme, which may explain why their data exhibit a huge shift in effort decisions. By the last round in the experiments in which  $nc = 0$ , almost all (96%) participants chose the highest possible effort, while over three-quarters chose the lowest possible effort when  $nc$  was around seven. One purpose of Van Huyck et al. (1990) was to show that a Pareto-inferior outcome may arise in coordination games, presumably because it is harder for large numbers of participants to coordinate on good outcomes. Other experiments were conducted with 2 players, but the payoff parameters were such that  $nc$  exactly equaled the critical value 1, and, with a random matching protocol, the data showed a lot of variability. The Goeree and Holt (2000) experiment also implements two-person random pairings in order to avoid serious possibility of tacit collusion in repeated games, which may drive efforts to maximal levels in sufficiently long series of repeated two-person coordination games. Given the knife-edge properties of  $c = 1/2$  for two-person coordination games, we conducted one treatment with  $nc = 1/2$  and another with  $nc = 3/2$ . As noted above, this change has a large effect on observed final-period effort choices even though the set of Nash equilibria is unaffected.

Risk dominance (or some generalization) predicts well in the long run, but the patterns of adjustment suggest that theories of learning and adaptation play an important role in the analysis of coordination. Each round of an experiment provides

some information about others' decisions and the resulting payoffs, and individuals may adapt their behaviour in the direction of a best response to past decisions. This is the approach taken by Crawford (1995), who specifies and estimates a model in which a person's decision is a weighted average of the person's past decision and the decision that would have been a best response to others' decisions in the past. This model incorporates some type of inertia, as well as an adaptive response. Another low-cognitive model is that players respond to payoffs received. These "reinforcement" learning models have been successfully applied by Erev and Roth (1997). In contrast, "belief" learning models assume that players use the history of past decisions to predict their opponents' next choice. Camerer and Ho (1999) have developed a hybrid model that combines elements of both reinforcement and belief learning models. Learning models such as these are especially useful in predicting how final-period behaviour may be explained by the history of past decisions.

Finally, many games in real life are played only once, e.g. elections, military contests, and legal disputes. In such cases, there is no relevant past history and players must rely on introspection about what others might do. This problem is particularly interesting for coordination games with multiple equilibria:

It should be emphasized that coordination is not a matter of guessing what the "average man" will do. One is not, in tacit coordination, trying what another will do in an objective situation; one is trying to guess what the other will guess one's self to guess the other to guess, and so on ad infinitum. (Schelling, 1980, p. 92-3)

Formal models of introspection often have such an iterative structure, incorporating responses to others' conjectured responses to others' conjectures, etc. Although these models are less well developed, they have shown some promise for explaining behaviour in games played only once (Goeree and Holt, 2001).

## References

- Camerer C and Ho TH (1999) Experience Weighted Attraction Learning in Normal-Form Games. *Econometrica* **67**: 827-874.
- Cooper R, DeJong DV, Forsythe, R and Rust TW (1989) Communication in the Battle of the Sexes Game: Some Experimental Results. *Rand Journal of Economics* **20**: 568-587.
- Crawford VP (1995) Adaptive Dynamics in Coordination Games. *Econometrica* **63**: 103-144.
- Erev I and Roth AE (1997) Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review* **88**: 848-881.
- Goeree JK and Holt CA (2000) An Experimental Study of Costly Coordination. Working Paper, University of Virginia.
- Goeree JK and Holt CA (2001) Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. *American Economic Review* **91**
- Guyer MJ and Rapoport A (1972) 2 x 2 Games Played Once. *Journal of Conflict Resolution* **16**: 409-431.

- Ochs J (1995) Coordination Problems. In: Kagel J and Roth A (eds.) *Handbook of Experimental Economics*, pp. 195-249. Princeton: Princeton University Press.
- Prisbrey, J (1991) An Experimental Analysis of the Two-Person Reciprocity Game. Working Paper, California Institute of Technology.
- Rapoport A, Seale DA, Erev I and Sundali JA (1998) Equilibrium Play in Large Group Market Entry Games. *Management Science* **44**: 129-141.
- Schelling TC (1980) *The Strategy of Conflict* Cambridge, Massachusetts: Harvard University Press.
- Van Huyck JB, Battalio RC and Beil RO (1990) Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. *American Economic Review* **80**: 234-248.

### Further Reading

- Anderson SP, Goeree JK and Holt CA (2001) Minimum Effort Coordination Games: Stochastic Potential and Logit Equilibrium. *Games and Economic Behavior* **34**: 177-199.
- Cachon G and Camerer C (1996) Loss-Avoidance and Forward Induction in Experimental Coordination Games. *Quarterly Journal of Economics* **111**: 165-194.
- Coate S and Loury G (1993) Will Affirmative Action Eliminate Negative Stereotypes? *American Economic Review* **83**:1220-1240.
- Cooper R and John A (1988) Coordinating Coordination Failures in Keynesian Models. *Quarterly Journal of Economics* **103**: 441-464.
- Goeree JK and Holt CA (1999) Stochastic Game Theory: For Playing Games, Not Just for Doing Theory. *Proceedings of the National Academy of Sciences* **96**: 10564-10567.
- Sefton M (1999) A Model of Behavior in Coordination Game Experiments. *Experimental Economics* **2**: 151-164.
- Straub PG (1995) Risk Dominance and Coordination Failures in Static Games. *Quarterly Review of Economics and Finance* **35**: 339-363.
- Van Huyck JB, Battalio RC and Rankin FW (1997) On the Origin of Convention: Evidence of Coordination Games. *Economic Journal* **107**: 576-596.