

CAN WE RATIONALLY LEARN TO COORDINATE?¹

ABSTRACT. In this paper we examine the issue whether individual rationality considerations are sufficient to guarantee that individuals will learn to coordinate. This question is central in any discussion of whether social phenomena (read: conventions) can be explained in terms of a purely individualistic approach. We argue that the positive answers to this general question that have been obtained in some recent work require assumptions which incorporate some convention. This conclusion may be seen as supporting the viewpoint of ‘institutional individualism’ in contrast to ‘psychological individualism’.

KEY WORDS: Individual rationality, collective rationality, learning, coordination, conventions.

“We must remember that a pattern whether of the past or of the future is always arbitrary or partial in that there could always be a different one or a further elaboration of the same one.”

(C. Palliser, *Quincunx*, 1989, Book V: The Maliphants, Part V).

1. INTRODUCTION

Many situations in every day life can be understood in the framework of coordination problems. Familiar examples include which side of the road to drive on, whether or not to call back once the connection of a telephone conversation has suddenly broken and, in the context of economics, what type of commodity to carry in your pocket as a medium of exchange. The typical feature of these and other coordination problems is that the optimal action of every individual agent crucially depends on the action that is chosen by the other agent(s). Moreover, there are two or more (Nash equilibrium) combinations at which each agent’s action is optimal given the actions of the other agents. Finally, the agents are often indifferent between the different equilibrium combinations. Accordingly, there is a strong sense in which the best course of action is undetermined.

Experience suggests that people in 'real' coordination problems are, on the whole, very well able to coordinate their actions. It seems that for every particular problem some actions 'stand out' compared to the other possible actions and that people tend to use the actions that stand out. At the theoretical level, this has led to investigations into the nature of 'salience' (Schelling, 1960) and 'conventions' (Lewis, 1969). The integration of these concepts into the individualistic language of game theory in which the coordination problems are cast has turned out to be rather problematic, however. Two questions have attracted some attention in this respect. First, how do the notions of salience and convention relate to the view that individual agents choose their actions in a rational way? (See, e.g., Gauthier, 1975 and Gilbert, 1989.) Second, can the notions of salience and convention themselves be grounded on rationality principles? (See, e.g., Bacharach, 1991 and Sugden, 1992.)

One might argue that these questions are inherently difficult to address in a *static* context and in this paper we take a slightly different perspective by investigating whether individual agents can learn, in a rational way, to coordinate their actions if the coordination problem that they face is a repeated one. If this were the case, the second question above could be answered in the affirmative, *i.e.*, a convention could be said to emerge from a rational learning process. Another rationale for adopting a dynamic perspective is that it is generally believed that although individual agents might not be able to solve a particular coordination problem at once, they will learn to coordinate actions after some time.

This paper will argue that individual rationality considerations are *not sufficient* to ensure that individual agents learn to coordinate. On the contrary, some form of common background (convention) has to be assumed for rational learning to coordinate to take place. The conclusion that we draw from this analysis for the individualistic program underlying much of the social sciences is in line with the position defended by Agassi (1960); see also Janssen (1993). He argues, in our view convincingly, against a psychological form of individualism and in favor of an institutional form of individualism. According to the latter view, the aim of the individualistic program is "neither to assume the existence of all co-

ordination nor to explain all of them, but rather to assume the existence of some co-ordination in order to explain the existence of some other co-ordination” (Agassi, 1960, p. 263). We will argue that some conventions have to exist in order to explain the emergence of other conventions. In recent papers (Goyal and Janssen, 1993a,b) we have followed this approach in developing formal models to study questions of inertia and dynamic stability of conventions.

The above point of view might appear to be in conflict with the general idea behind recent papers by Crawford and Haller (1990) and Kalai and Lehrer (1993). We will discuss these papers in some detail while developing our argument. Crawford and Haller argue that there are optimal rules of learning to coordinate. These optimal rules themselves are, however, in general not unique and, in our view, this non-uniqueness problem can only be resolved by positing the existence of a coordinating convention at a higher level. Kalai and Lehrer, on the other hand, show that under some restrictions on prior beliefs, Bayesian (‘rational’) learning leads in the long run to Nash equilibrium. Applying this result to our context might suggest that Bayesian learning leads to individual agents playing one of the pure Nash equilibria of the coordination game. We will argue that the restrictions on prior beliefs that they require implicitly assume a conventional selection of priors.

Another conclusion one might draw from our analysis is that the perfect rationality paradigm is not very powerful when it comes to the study of interactive learning processes. Instead, one could resort to models of boundedly rational behavior or evolutionary models. We think recent developments (cf., Kandori *et al.*, 1993; Young, 1993) within these approaches are interesting, but they fall outside the scope of the present paper. Here, we focus on the issue to what extent learning by perfectly rational agents can ensure coordination.

In order to be able to state our argument with some care, we have to give definitions of the basic terms that are used. This will be done in Section 2. Section 3 contains the discussion of the problems rational individual agents face while learning to coordinate. Sections 4 and 5 examine the papers by Crawford and Haller (1990) and Kalai and Lehrer (1993a), respectively. Section 6 concludes with a brief comparison of the two approaches.

2. NOTATION AND DEFINITIONS

In order to concentrate on issues of coordination we will consider pure coordination games only. In a pure coordination game there is no conflict of interest; the only aim of the players is to coordinate. For notional simplicity we will restrict ourselves to the two players case. Most of the time we also restrict attention to the simplest coordination game, namely the following two action game:

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (1, 1) & (0, 0) \\ (0, 0) & (1, 1) \end{pmatrix} \end{array}$$

This game has two pure strategy Nash equilibria: (T, L) and (B, R) . In the discussion of Crawford and Haller (1990) in Section 4 we will also consider a k action coordination game, where the pay-off matrix is the $k \times k$ identity matrix. Denote the row player and the column player as 1 and 2, respectively. The two players are playing this game an infinite number of times. Let $\delta \in [0, 1)$ denote the common discount factor. The players observe all actions and the resulting pay-offs, *i.e.*, history is completely observable. We also assume that players have perfect recall. The players cannot communicate except via their actions.² These assumptions and the rationality of the players are assumed to be common knowledge.

It is possible that the two players might choose one of the two coordination equilibria by chance. If, for example, in period t player 1 chooses T with probability $\frac{1}{2}$ and player 2 chooses L with probability $\frac{1}{2}$, then the chance of coordination in period t will be equal to $\frac{1}{2}$. Moreover, if the two players keep playing these mixed strategies, then for any finite number of times the game is actually played, the chance that they have coordinated at every instance is strictly positive. This type of ‘coordination’ is unsatisfactory, however. What one wants is that coordination is *ensured*, not that there is a (small) probability that it might result. Also, it is not clear what the players learn when they constantly play this mixed strategy.

In order to clarify what we mean by ‘learning to coordinate’, we introduce the following (standard) notation and definitions. A (behavior) strategy³ of player i in period t is a function $s_i : H \rightarrow \Delta(A_i)$, where H is the set of *all possible* histories, $\Delta(A_i)$ denotes

the set of probability distributions over A_i and $A_1 = \{T, B\}$ and $A_2 = \{L, R\}$. Thus, a strategy specifies how a player randomizes over his actions after all possible histories. We denote by $\pi_i(s_1, s_2)$ the discounted sum of expected pay-offs of player i , if strategies s_1 and s_2 are played, *i.e.*,

$$\pi_i(s_1, s_2) = (1 - \delta) \sum_{t=1}^{\infty} E_{(s_1, s_2)} \delta^t \pi_i(a_1(t), a_2(t)) \quad i = 1, 2,$$

where $a_i(t) \in A_i$ is the action of player i in period t and $\pi_i(a_1(t), a_2(t))$ is the stage-game pay-off of player i in period t .

Player i 's belief about player $-i$'s strategy is denoted by μ_i , where $\mu_i : H \rightarrow \Delta(A_i)$. Such a belief specifies how player i expects player $-i$ to randomize over his actions after any possible history. Note that the possibility for learning about the other player's strategy is incorporated in the belief.

Given the above definitions of strategies and beliefs, the notion of individual rational behavior that is employed in this paper is a traditional one. A strategy s_i is said to be *individually rational* if there is a belief μ_i such that (i) s_i maximizes the discounted sum of expected pay-offs given μ_i and (ii) μ_i is 'consistent' with all information gathered throughout history. For the moment, we restrict ourselves to a very general notion of consistency according to which a belief should be such that at any point in time the observed history up to that time should have been possible given a player's own strategy and this belief. Other notions of consistency, like Bayesian updating, will be introduced later.

We are now ready to define the basic concept of 'coordination'. Coordination is regarded as a property of a pair of strategies as follows.

DEFINITION 1. A pair of strategies $\{s_1, s_2\}$ is *coordinated* if for any $\varepsilon > 0$ there is a finite $t^*(\varepsilon)$ such that for all $t > t^*(\varepsilon)$,

$$\begin{aligned} P[(a_1(t), a_2(t)) = (T, L) | (s_1, s_2)] &> 1 - \varepsilon \quad \text{or} \\ P[(a_1(t), a_2(t)) = (B, R) | (s_1, s_2)] &> 1 - \varepsilon. \end{aligned}$$

The expression $P[(a_1(t), a_2(t)) = (T, L) | (s_1, s_2)]$ in the above definition has to be read as the probability that the action combination (T, L) is played in period t given that the players have chosen the strategy combination (s_1, s_2) .

Two elements of this definition are worth emphasizing. First, we say that a pair of strategies is coordinated only if one of the two pure strategy equilibria is played from some finite time period onwards with very high probability. A pair of strategies which results, for example, in alternating between (T, L) and (B, R) therefore does not satisfy this requirement. We have opted for this formulation, because we want to investigate to what extent a convention to play the *stage game* can emerge out of a process in which rational players learn while playing the game. Second, we do not require that there is a finite period after which the players always play the same pure strategy equilibrium. This is done in order to take account of the possibility that players randomize until a certain, somehow desired, outcome (for example, (T, L)) in the stage game is obtained and play pure strategies ever after (cf., Section 4). For every finite period there is always a (small) chance that they have not reached the desired outcome if they choose to play such a strategy.

3. THE PROBLEM STATED

In the previous section we have not discussed the learning process itself: how do rational individuals learn? It is clear that the strategies that players can rationally choose do not change over time, if the beliefs μ_i about future plays of the opponent that are consistent with past observations do not change. This suggests that the problem of learning in games is of the following type: what can the players infer about the future on the basis of a finite number of observations? Can they become more certain about future plays of the opponent the more observations they have about the opponent's past play? This is, of course, very similar to the well-known Problem of Induction. In the present context, however, there is an additional problem. An agent should take into account the possibility that the opponent's play is not stationary, because she (the opponent) is also learning about the strategies he is playing and will try to adjust optimally.

Two notions of learning that are frequently employed in game theory and in economics are adaptive (Cournot) learning and 'stationary' (Bayesian) learning (cf., Eichberger *et al.*, 1993).⁴ For the present purposes, both notions are unsatisfactory, because they assume in one way or another that an individual player takes the other player's stage game strategy as stationary. This is something

they know to be false, since they know their opponent is learning at the same time.

A notion of learning that is more satisfactory in this respect is the notion of *sophisticated learning* introduced by Milgrom and Roberts (1991). They show that the stage game strategies of what they regard as *sophisticated learners* converge to the set of iteratively undominated strategies. Applying their results (especially Theorem 5) to our context yields the conclusion that sophisticated learning *cannot* ensure coordination. The idea, basically, is that an uncoordinated pair of strategies as, for example, $\{s_1, s_2\} = \{BBB \dots, LLL \dots\}$, where *BBB* stands for 'always *B*' is consistent with sophisticated learning, because at any time each player can consistently forecast that the other player will see that it is in his or her interest to switch actions the next time period, *i.e.*, nothing prevents player 1 from expecting player 2 to play *R* the next period after having observed a finite number of *L*'s. It is then rational to continue to play *B*.

One might wish to argue that the notion of sophisticated learning is not an appropriate notion of rational learning in the present context if $\{BBB \dots, LLL \dots\}$ is a pair of strategies that is consistent with sophisticated learning. Indeed, it seems reasonable to assume that in an infinitely repeated game the players will reach one of the coordination equilibria at least once. So, let us consider the question of what the players are able to infer after having reached a pure strategy equilibrium of the stage game. We argue that the strategic uncertainty the individual players face in deciding upon their strategy at period $t + 1$ is basically the same whether or not they have reached a coordination equilibrium at period t . Suppose the players have played (T, L) in period t . For the players individually there is no way of knowing whether they have coordinated on a pair of strategies $\{s_1, s_2\} = \{TBTB, \dots, LRLR \dots\}$ or on $\{s_1, s_2\} = \{TTTT, \dots, LLLL \dots\}$ or on still other pairs of strategies that could possibly have resulted in the action combination (T, L) in period t . Accordingly, it is possible that one of the players thinks that they will alternate deterministically between (T, L) and (B, R) , while the other player thinks that they will continue playing (T, L) . This means that after (T, L) has occurred in period t , the players may find themselves at (B, L) in period $t + 1$. The same is true for any number of times they have reached a coordination equilibrium. In other words, on the basis of only a finite number of observations

the players cannot ascertain what the other player will do in the (near) future, i.e., the players face the well-known problem that at any point in time there is an infinite number of hypotheses that is consistent with past observations (each of these hypotheses having different implications for expected future play). We will call the problem that the players do not know which hypothesis to use for predicting future opponent's play the Projection problem (after Goodman, 1973).

The argument we are putting forward is very much like an argument recently made by Gilbert (1990). Gilbert (p. 10) argues that "common knowledge of precedent as such will not by itself automatically generate expectation of conformity". The reason she offers is that after having coordinated before, a rational agent will infer that it is best to conform to precedent if *and only if* his opponent will do likewise. He will realize that his opponent does not have an independent reason for conforming to precedent either. Therefore, there is nothing inherently rational in conforming to precedent.

In this section we have argued that the essential problem in rational learning to coordinate lies in the difficulty of making inferences about future play on the basis of past play. The fact that people often seem to think they can learn from past experience might suggest that our argument is too agnostic to provide an accurate account of how rational agents behave in repeated coordination situations. This motivates our discussion of two other approaches which try to ensure coordination by imposing more structure on the interactive learning process.

4. APPROACH I: EXTENDED RATIONALITY AND NO COMMON LANGUAGE

The approach considered in this section is based on a paper by Crawford and Haller (1990), henceforth C&H, who extend the notion of rationality in coordination games and introduce a 'no common language' assumption. We will introduce their ideas below and look at the consequences of their approach for the problems we have indicated in the previous section.

Extended Rationality

C&H "maintain the working hypothesis that players play an optimal ... attainable strategy combination" (p. 580). They define an *attain-*

able strategy combination as a pair of strategies in which each player's 'undistinguished' actions enter his overall strategy in a symmetric way and in which players whose positions are 'undistinguished' play identical strategies. An *optimal* attainable strategy combination is an attainable strategy combination that maximizes both player's repeated-game pay-offs. For the moment, we will assume that players have a common language in which they describe the game so that the players' actions are distinguished and all strategy combinations are attainable. In this case, C&H's working hypothesis reduces to the players choosing an optimal strategy combination, *i.e.*, they coordinate immediately.

The idea of players choosing an optimal strategy *combination* is similar to a suggestion by Sugden (1991) who proposes that in coordination games, players be regarded as a team in a game against nature. As a member of a team, each rational player should *individually* think of a strategy *combination* which, if both players did their part, would optimize their individual (and collective) pay-offs. We will call this notion 'individual rationality in the extended sense', because this notion defines rational behavior in such a way as to include a belief of a 'good' strategy of the other player. This extended notion of individual rationality is, we think, quite natural in pure coordination games, because players know that it is in their mutual interest to coordinate. Hence, it is reasonable to expect 'good will' on the part of the opponent. The notion of extended rationality is a generalization of the Principle of Coordination to the context of repeated games. This principle says that in coordination games with a strict Pareto-best equilibrium, each player should do his or her part in it (see, e.g., Gauthier, 1975 and Bacharach, 1991).

We now formally state the definition of 'extended individual rationality' and distinguish it from 'collective rationality'. We will make this distinction in order to be able to express what it means for players to choose a strategy combination.

DEFINITION 2. A strategy s_i is *individual rational in the extended sense* if it is part of a pair of strategies $\{s_i, s_{-i}\}$ such that for $i = 1, 2$,

$$\pi_i(s_i, s_{-i}) \geq \pi_i(s'_i, s'_{-i}),$$

for all strategies s'_i, s'_{-i} of players i and $-i$.

DEFINITION 3. A pair of strategies $\{s_1, s_2\}$ is *collectively rational* if

$$\pi_i(s_1, s_2) \geq \pi_i(s'_1, s'_2), \quad i = 1, 2,$$

for all strategies s'_1, s'_2 of players 1 and 2.

The difference between the two definitions is that only the second definition requires some collective action to be undertaken. The first definition requires the players individually to come up with a ‘team’ solution (pair of strategies) that would be optimal for both. Each player should then perform his part of the team solution. The two definitions yield the same outcomes if and only if there is a *unique* pair of strategies that is *collectively* rational. In this case, both players will necessarily find the same solution to the ‘team problem’ and the definition of extended *individual* rationality prescribes that the players perform their part of this pair of strategies. Since, typically, there are many pairs of strategies that are collectively rational, every individual player of ‘good will’ can choose his or her part of a different pair of such strategies. Accordingly, coordination is not ensured.

The fact that in the case of a common language C&H’s approach yields a coordinated outcome from the first period onwards is due to their implicit assumption that players choose a pair of strategies that is collectively rational, *i.e.*, our definition of collective rationality coincides with their hypothesis that players play an optimal attainable strategy combination. This is exemplified by their saying that players can use “their perfect recall and knowledge of the structure of the game to maintain coordination forever once they locate a pair of coordinated actions. They can do this either by repeating those actions, or by alternating deterministically between them and the other coordinated pair” (p. 575). As Crawford and Haller (and we) assume that communication is not possible except by playing the game, we think that players *individually* cannot maintain coordination once they have reached a coordination equilibrium of the stage game. It is precisely because both ‘repeating the same action as in the last period’ and ‘choose the action that you didn’t choose in the last period’ are collectively rational that player 1 might think that player 2 will repeat her action, while player 2 might think that player 1 alternates his action deterministically. Thus, it is not certain that

rational players will maintain coordination once they have located a pair of coordinated actions (see also Section 3).

Of course, it might be that coordination results even in cases where there is no unique solution that is collectively rational. This can be the case, for example, when one ‘team solution’ is more prominent than the others. This would, however, introduce a focal point solution at the level of which pair of strategies to select from the set of strategies that is collectively rational. This means that C&H’s approach only seems to take the coordination problem one stage backwards.

No Common Language and No Common Labeling

A second aspect of C&H’s paper is that they allow for situations in which players do not have a common labeling to describe the game. This assumption has the powerful implication that not all strategy combinations are attainable. As an example of no common labeling, one can think of the game described in Section 2 with the modification that player 1 does not know what L and what R is for player 2 and player 2 does not know what T and what B is for player 1. In this case, it makes no sense for a player to try to coordinate on say (T, L) , because he does not know whether T and L have the same meaning for the other player. Hence, Definitions 2 and 3 have to be modified in such a way that the set of strategy combinations is restricted to the set of attainable ones. When there are multiple attainable optimal strategy combinations, the same difficulties as the ones described above arise. However, due to the restriction to attainable strategies there are cases (see below) in which there is a unique optimal strategy combination and, under the assumption of extended rationality, individuals will do their part of this strategy combination.

It will turn out that a distinction between ‘no common *language*’ and ‘no common *labeling*’ is useful.⁵ We will say that players do not have a common language if there is no one-to-one correspondence from one language to the other; they have no common labeling of terms if there exists such a one-to-one correspondence between every pair of terms.⁶ We will argue that a new set of problems arise if the assumption of no common language is taken seriously. On the other hand, for the no common labeling assumption C&H obtain a nice result for the 3×3 action game.

To illustrate the nature of this problem, we consider a reformulation of Goodman's 'grue' and 'bleen' example (Goodman, 1973). Suppose, that player 1 describes the coordination game in terms of the terms *TOP* and *BOTTOM*, while player 2's language describes the game in terms of *TOTTOM* and *BOP*. The peculiar feature of the two languages is that what *TOP* is for player 1 is, for player 2, *TOTTOM* until t and *BOP* after t . Similarly, what *BOTTOM* is for player 1 is for player 2 *BOP* until time t and *TOTTOM* after t . Reciprocally, what *TOTTOM* (*BOP*) is for the second player is for the first player *TOP* (*BOTTOM*) until t and *BOTTOM* (*TOP*) after time t .⁷

Let us now suppose that players have attained a stage game equilibrium for the first time in period t and that the pair of actions that have produced this equilibrium is (*TOP*, *TOP*) in player 1's language and (*TOTTOM*, *TOTTOM*) in player 2's language. In this case the focal point 'stick to the same action' does not work, because player 1 will persist with *TOP* and the second player will persist with *TOTTOM* and they will *not* coordinate in period $t + 1$. It is interesting to note that *both* players might accuse the other of changing actions!⁸

Let us now briefly consider the case of no common labeling. For the 3×3 action game, C&H propose the following strategy: "play each of your actions with equal probability in the first stage. If coordination results, maintain it by repeating your first-stage action. If not, rule out your first stage action and the action that would have yielded coordination given your partner's first action; then play the action not ruled out from the second stage onwards" (p. 585). It is relatively easy to see that this strategy combination is the *unique* attainable strategy combination that is collectively rational. Hence, individual rational players in the extended sense should play their part of this strategy combination.⁹ For the 2×2 case, C&H propose that the players randomize over their two actions with probabilities $(\frac{1}{2}, \frac{1}{2})$ until a stage game equilibrium is reached and then to continue to play these actions or to alternate deterministically. This pair of strategies ensures coordination in the way it has been defined in Section 2. However, as the proposal of how to continue after an equilibrium is reached suffers from non-uniqueness, there is, again, no guarantee that individual players who are rational in the extended sense, play their part of the *same* pair of strategies. Therefore, indi-

vidual rationality in the extended sense does not ensure coordination in the 2×2 case. The non-uniqueness argument applies to all $k \times k$ action games, where $k = 2, 4$ and $k \geq 6$ (cf., C&H, p. 585).

Concluding Observations

C&H assume that players choose optimal attainable strategy combinations. In order to evaluate whether their analysis is based on purely individualistic assumptions we have distinguished between ‘individual rationality in the extended sense’ and ‘collective rationality’. In the discussion above we have seen that for the 3×3 action case (and for the 5×5 action case with a low value of the discount factor), when players do not have a common labeling of their actions, individual rationality in the extended sense suffices to ensure that a convention can emerge out of a rational learning process. In all other cases, however, extended individual rationality does not ensure coordination (due to non-uniqueness). To resolve the non-uniqueness problem, C&H implicitly resort to the non-individualistic assumption of ‘collective rationality’. Another way out of the non-uniqueness problem is to assume the pre-existence of a convention (or focal point) at a higher level (the level corresponding to which pair of strategies to select from the set of strategies that is collectively rational) to explain the endogenous emergence of a new convention.

5. APPROACH II: BAYESIAN LEARNING IN THE INFINITE STAGE GAME

In a sequence of papers, Kalai and Lehrer (1993(a), (b)) – K&L for short – study the nature of Bayesian learning and its implications for equilibrium play in an infinitely repeated game context. They are concerned with learning repeated game strategies and as these strategies are given and do not change over time, they overcome the above mentioned difficulty of the possible non-stationarity of individual agents’ stage game strategies. We will restrict the discussion of the results K&L obtain to the two-player coordination games under consideration.

The basic idea of K&L’s paper is the following. At the start of the repeated game, players have prior beliefs over the possible strategies that the opponent plays in the *repeated game*. Given these prior

beliefs, each player chooses the repeated game strategy that maximizes his discounted sum of pay-offs. The beliefs are updated using Bayes' rule. The crucial assumption K&L use is that the combination of strategies that is actually played is *absolutely continuous* with respect to the players' prior beliefs.¹⁰ Roughly speaking, they show that players will eventually correctly predict the future play of their opponent if this assumption is satisfied. As players maximize discounted pay-offs, this implies that their strategies will eventually form a Nash equilibrium of the repeated game. The results they obtain are more easily explained if a condition stronger than absolute continuity, namely what they call the 'grain of truth' assumption, holds.¹¹ Therefore, we will consider and comment on this case first.

Suppose player 1's beliefs about player 2's strategy are as follows:

$$\begin{aligned} P(\text{always } R) &= \alpha; & P(\text{always } L) &= \frac{1}{2} - \alpha \quad (0 < \alpha < \frac{1}{2}); \\ P(RLLL\dots) &= \frac{1}{3}; & P(RRLL\dots) &= \frac{1}{9}; \\ P(RRRL\dots) &= \frac{1}{27} & \text{and so on,} \end{aligned}$$

where, for example, $P(RLLL\dots) = \frac{1}{3}$ is to be read as saying that the prior probability of observing R in the first period and L always after is equal to $\frac{1}{3}$. Suppose, moreover, that player 2 plays always R . The 'grain of truth' assumption says that each player should attach a strictly positive prior probability to the actual strategy that is chosen by the opponent. Observe that the assumption is satisfied in the example considered here. K&L's argument works as follows. After observing R in period 1, player 1 knows that player 2 does not always play L . Bayesian updating requires that the probability mass of $\frac{1}{2} - \alpha$ is distributed proportionally to its weight over the strategies that are possibly played. After observing R in the first two periods, $5/6 - \alpha$ is distributed in this way over the remaining possible strategies. It is clear that for any value of $\alpha > 0$, after sufficiently long time t , player 1 is almost sure that player 2 plays strategy 'always R ' and the chance of observing R in period t after having observed a finite number of R s up to time $t - 1$, $P(R_t | R_1, \dots, R_{t-1})$, will be arbitrarily close to 1 for large t . The optimal response to such a belief is for player 1 to play B in all subsequent (infinite) periods.

To see more clearly the importance of the ‘grain of truth’ assumption, suppose that, instead of the above beliefs, player 1’s beliefs are slightly different:

$$\begin{aligned} P(\text{always } L) &= \frac{1}{2}; \\ P(RLLL\dots) &= \frac{1}{3}; \quad P(RRLL\dots) = \frac{1}{9}; \\ P(RRRL\dots) &= \frac{1}{27} \quad \text{and so on.} \end{aligned}$$

In this case the grain of truth assumption is not satisfied and $P(R_t|R_1, \dots, R_{t-1}) = \frac{1}{3}$ for any value of t . Accordingly, at the beginning of every period player 1 thinks (wrongly) that the chance that player 2 plays l in this period is larger than $\frac{1}{2}$ and, consequently, he will play T . Thus, the players will never coordinate given these prior beliefs.

The grain of truth assumption is, however, a very strong condition to impose on the prior beliefs. This is because a positive prior probability can be attached to a countably infinite number of strategies only, whereas there is an uncountable number of strategies from which the true strategy can be chosen. Any countable set of strategies is very small compared to the uncountable strategy set from which the true strategy is chosen. Thus, there are in principle a very large number of priors, each of which assigns a positive probability to a countable number of strategies. From this set of possible priors there is very little reason for individuals to choose a prior of the right sort, *i.e.*, one that assigns strictly positive probability to the true strategy of the opponent.

The ‘grain of truth’ assumption can also be examined in relation to the Problem of Induction and the Projection Problem. In Section 3 we argued that the main difficulty for players in learning to coordinate is the problem of what to infer about future play of the opponent, given some observations about past play. In general, there is an uncountable number of hypotheses consistent with a finite number of observations. One can only hope to make some useful inference about future play if one can argue that one of these hypotheses is more likely to be true than the others. In the Bayesian approach, the prior is responsible for the selection of the most reasonable hypothesis given the available evidence. In the first example above, player 1 argues that a particular hypothesis, namely ‘always R ’ after having observed $R_1 \dots R_{t-1}$, is more likely to be true than any other

hypothesis (for large t), because of the form of the prior. We have seen that this conclusion no longer holds in the second example. But what reasons can a player have for choosing a particular prior?

Recall that the aim of the present paper is to examine whether a convention can emerge out of an interactive learning process between rational players. The critical point then seems to be how players choose their priors. One common answer here is that priors are subjective. However, the second example shows that the priors in K&L's approach cannot be just *any* subjective priors. This is because subjective priors will, in general, not satisfy the grain of truth requirement. One can interpret the drastic restriction implied by the grain of truth assumption as saying that priors are of a *conventional nature*. This conventional nature of priors could be defended by arguing that players will not choose 'unusual' strategies and the players 'know' that the other player will not do this, and so on. This means that many of the possible strategies will not be chosen. The point here is that the players should have this 'knowledge' in common, *i.e.*, not only should a player not choose 'unusual' strategies, the other player should also know that he will not do so. In other words, there should be a convention saying that certain observations are indicative of particular strategies. Thus, K&L implicitly introduce a conventional aspect in their analysis.

Can the above criticism be overcome by replacing the 'grain of truth' assumption by the 'absolute continuity' assumption? After all, it is the latter assumption that K&L really employ in their paper. The above criticism still goes through for all strategies with the property that some action combinations in the repeated game occur with strictly positive probability and it is in such strategies that we are interested in (cf., Definition 1). Let us consider, as an example, the strategy that Crawford and Haller propose for the 2×2 coordination game, *i.e.*, 'randomize over your possible actions with probability $\frac{1}{2}$ until you reach a coordination equilibrium, and continue to play your part of this equilibrium ever after'. This strategy combination¹² results with probability $\frac{1}{4}$ in $\{(T, L), (T, L), \dots\}$, with probability $\frac{1}{4}$ in $\{(B, R), (B, R), \dots\}$, with probability $1/8$ in $\{(T, R), (T, L), (T, L), \dots\}$, with probability $1/8$ in $\{(B, L), (B, R), (B, R), \dots\}$, and so on. Two things should be noted. First, there is a countably infinite number of action combinations that occurs with

positive probability and, second, there is an uncountable number of action combinations that will not occur given this strategy combination. The absolute continuity assumption requires that the prior beliefs assign positive probability to all action combinations that occur with strictly positive probability. The simplest way to guarantee this, of course, is 'the grain of truth' assumption, but there are other ways as well. The crucial point is, however, that one can choose the priors over strategies only such that a countable number of action combinations receive a strictly positive probability. The same argument as above can now be applied replacing 'strategies' by action combinations: as, at the start of play, there are an uncountable number of potentially true action combinations that are consistent with the available evidence and as any countable set of action combinations is very small compared to an uncountable set, it is very unlikely that a player's belief assigns strictly positive probability to the true actions that are played by the other player in the repeated game. So, we conclude that the 'absolute continuity' assumption cannot overcome the above criticism.

More recent contributions to Bayesian learning, in particular Kalai and Lehrer (1994) and Lehrer and Smorodinsky (1993), try to overcome the restrictions imposed by the absolute continuity assumption. They show that a 'weak' learning may take place when the beliefs are dispersed around the true strategy being played (in the sense that the beliefs assign a positive probability to any 'neighbourhood' of the true strategy being chosen). 'Weak learning' in this context means that a player is able to predict opponent's play in the near future only. As a careful technical analysis of the topologies and notions of learning needs to be considered the reader is referred to the original papers for a precise statement of the results and the conditions under which they hold (see also Section 7.2 of Kalai and Lehrer, 1993a). For the purpose of this paper these recent contributions are interesting in that they show that the conventions needed to ensure coordination may be more vague than the conventions implied by the absolute continuity assumption.¹³

Our conclusion on the conventional nature of the selection of priors is strengthened by the following observation. K&L show that play will converge to the Nash equilibria of the repeated game, *not* to the pure strategy Nash equilibria of the stage game. There are very many equilibria of the repeated game and the equilibrium to

which one converges depends on the priors. Recall that we have defined coordination as a property of a pair of strategies in such a way that a convention to play the *stage game* emerges out of a rational learning process (see Definition 1 above). Thus, even if the ‘absolute continuity’ assumption holds, there is no guarantee that coordination is ensured. On the contrary, it may very well be that the Bayesian players of K&L converge to an equilibrium in which they, for example, alternate between (T, L) and (B, R) , or worse, in which they play the mixed strategy Nash equilibrium of the stage game for ever. In order to obtain convergence to one of the pure strategy Nash equilibria of the stage game, additional restrictions on the priors may have to be imposed. This makes it even more unlikely that players have the ‘right’ prior.

In concluding this section, we can say that K&L’s approach circumvents the problem of the non-stationarity of stage game strategies by applying Bayesian learning to the repeated game strategies, instead of to the stage game strategies. However, their approach has its own problems, two of which are intimately related to the issues we have discussed in the present paper. First, the Projection Problem is ‘solved’ by attributing a selective prior to the players. The set of opponent’s strategies to which a positive prior probability can be given is, however, very small with respect to the set of all possible strategies. The above discussion has suggested that this implicitly implies a conventional selection of priors. Second, K&L show convergence to one of the Nash equilibria of the *repeated* game, not to one of the pure strategy stage game equilibria. Thus, their results do not satisfy the requirements of ‘learning implies coordination’ which were developed in Definition 1 above.

6. CONCLUSION

In this paper we have examined the issue whether individual rationality considerations are sufficient to guarantee that individuals will learn to coordinate. This question is central in any discussion of whether social phenomena (read: conventions) can be explained in terms of a purely individualistic approach.

In the analysis we have paid special attention to recent work by Crawford and Haller (1990) and Kalai and Lehrer (1993a,b). With

respect to C&H we have argued that the hypothesis that players employ optimal attainable strategy combinations is, in general, non-individualistic. On the other hand, the K&L result on learning to play Nash equilibrium is based on the crucial assumption of absolute continuity, which we have argued implicitly assumes that some strategies are salient. This salience is not explained in individualistic terms.

To summarize: we have argued that the positive answers obtained to the general question posed above require assumptions which incorporate some convention. This conclusion may be seen as supporting the viewpoint of 'institutional individualism' in contrast to 'psychological individualism'.

NOTES

¹ Earlier versions of this paper have been presented at a workshop on 'The Emergence and Stability of Institutions' (Louvain-la-Neuve, June 1991), at a workshop on Game Theory and Philosophy (Amsterdam, December 1991, at a seminar at Carnegie-Mellon University (November 1993) and at a conference on Epistemic Logic and The Theory of Games and Decisions (Marseille, January 1994). Comments by workshop and seminar participants, especially by André Orléan and Robert Sugden, and by Vincent Crawford, Ehud Kalai, Theo Kuipers, Nick Vriend and two anonymous referees are gratefully acknowledged.

² This is a reasonable assumption in some settings. Examples include telephone connections (see above); the paratroopers problem, or the electronic mail game (Rubinstein, 1989). Moreover, in the philosophical literature it has been argued that communication itself (language and truth telling) is a coordination problem (cf., Hodgson, 1969; Lewis, 1972; den Hartog, 1985).

³ For simplicity we will restrict ourselves to behavior strategies. It is well-known that, under the assumption of perfect recall, every mixed (hence, pure) strategy can be replaced by an equivalent single behavior strategy (cf., Kuhn, 1953; Kalai and Lehrer, 1993).

⁴ This latter notion is not to be confused with the Bayesian learning that takes place in Kalai and Lehrer (1993a), to which we return in Section 5.

⁵ C&H do not make this distinction. In private correspondence, Professor Crawford has pointed out that they introduced the term 'no common language' as a convenient synonym for what we call 'no common labeling'. Thus, the discussion on 'no common language' below should not be regarded as a criticism of their paper.

⁶ It is well-known that for almost any pair of languages there are many words for which no one-to-one correspondence exists. The 'proper' translation often depends on the context in which the term is used.

⁷ It is interesting to observe that the coordination problem is so closely related to the problem of induction the players are confronted with. On p. 120 of the

monograph, Goodman (1973) argues that “the roots of inductive validity are to be found in our use of language”. This indicates that it is even more difficult for the players to solve the coordination problem once the common language assumption is dropped.

⁸ The language problem is not resolved by assuming that there is mutual or common knowledge of the two languages. In this case the players do not know which language to use: they should use the language that the opponent is also using, but know that their opponent uses the language that she thinks he will choose, and so on.

⁹ Note that the proposed pair of strategies does not ensure coordination from the second stage onwards if the players do not have a common language.

¹⁰ Recall that the actual strategies chosen induce a probability distribution over the set of possible outcomes (play paths) of the game. These strategies are absolutely continuous with respect to the beliefs, if the players’ prior beliefs are such that they assign a strictly positive probability to every outcome that occurs with strictly positive probability given the actual strategies chosen.

¹¹ In independent work, Nyarko (1990) has used a similar assumption on the support of agents’ beliefs to demonstrate convergence to Nash equilibrium in a repeated game with quantity setting firms.

¹² Recall that both players are advised to choose the same strategy.

¹³ We are grateful to one of the referees for this interpretation of the recent work on Bayesian learning.

REFERENCES

- Agassi, A.: 1960, ‘Methodological Individualism’, *British Journal for Sociology* **11**, 144–170.
- Bacharach, M.: 1991, ‘Conceptual Strategy Spaces’, Oxford discussion paper.
- Boland, L.: 1982, *Foundations of Economic Method*, London: Allen and Unwin.
- Caldwell, B.: 1984, *Beyond Positivism*, London: Allen and Unwin.
- Crawford, V. and Haller, H.: 1990, ‘Learning How to Cooperate: Optimal Play in Repeated Coordination Games’, *Econometrica* **58**, 571–595.
- Eichberger, J., Haller, H. and Milne, F.: 1993, ‘Naive Bayesian Learning in 2×2 Games’, *Journal of Economic Behavior and Organization* **22**, 69–90.
- 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.
- Goodman, N.: 1973, *Fact, Fiction and Forecast* (3rd Ed.), New York: Bobbs-Merrill Company.
- Goyal, S. and Janssen, M.: 1993a, ‘Dynamic Coordination Failures and the Efficiency of the Firm’, Tinbergen Institute Research paper No. 93–92. *Journal of Economic Behavior and Organization* (forthcoming).
- Goyal, S. and Janssen, M.: 1993b, ‘Non-exclusive Conventions and Social Coordination’, Tinbergen Institute Research paper No. 93–233.
- Hartog, G. den.: 1985, *Wederkerige Verwachtingen* [Mutual Expectations], Unpublished Ph.D. Thesis, University of Amsterdam (in Dutch).

- Hodgson, D.H.: 1967, *Consequences of Utilitarianism*, Oxford: Oxford University Press.
- Janssen, M.: 1993, *Microfoundations: A Critical Inquiry*, London: Routledge.
- Kalai, E. and Lehrer, E.: 1993a, 'Rational Learning Leads to Nash Equilibrium', *Econometrica* **61**, 1019–1045.
- Kalai, E. and Lehrer, E.: 1993b, 'Private-Beliefs Equilibrium', *Econometrica* **61**, 1231–1240.
- Kalai, E. and Lehrer, E.: 1994, 'Weak and Strong Merging of Opinions', *Journal of Mathematical Economics* **23**, 73–86.
- Kandori, M., Malaith, G., and Rob, R.: 1993, 'Learning, Mutation and Long Run Equilibria in Games', *Econometrica* **61**, 29–56.
- Kuhn, H.W.: 1953, 'Extensive Games and the Problem of Interaction', in H.W. Kuhn and W.W. Tucker (Eds.), *Contributions to the Theory of Games*, Vol. II, pp. 193–216. Princeton: Princeton University Press.
- Lehrer, E. and Smorodinsky, R.: 1993, 'Compatible Measures and Merging', Tel Aviv University, Mimeo.
- Lewis, D.: 1969, *Conventions*, Cambridge (Mass.): Harvard University Press.
- Lewis, D.: 1972, 'Utilitarianism and Truthfulness', *Australasian Journal of Philosophy* **50**, 17–19.
- Milgrom, P. and Roberts, J.: 1991, 'Adaptive and Sophisticated Learning in Normal Form Games', *Games and Economic Behaviour* **3**, 82–100.
- Nyarko, Y.: 1990, 'Bayesian Rationality and Learning without Common Priors', New York University, mimeo.
- Rubinstein, A.: 1989, 'The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge', *American Economic Review* **79**, 385–389.
- Schelling, T.: 1960, *The Strategy of Conflict*, Cambridge (Mass.): Harvard University Press.
- Sugden, R.: 1991, 'Rational Choice: A Survey of Contributions from Economics and Philosophy', *Economic Journal* **101**, 751–786.
- Sugden, R.: 1992, 'Towards a Theory of Focal Points', University of East Anglia, discussion paper.
- Young, H.P.: 1993, 'The Evolution of Conventions', *Econometrica* **61**, 57–84.

*Econometric Institute (S.G.) and
Department of Microeconomics, (M.J.),
Erasmus University,
P.O. Box 1738,
3000 DR Rotterdam, The Netherlands*