An "Evolutionary" Interpretation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination

VINCENT P. CRAWFORD*

*Department of Economics, University of California, San Diego,
La Jolla, California 92093*

Received January 10, 1990

This paper proposes an adaptive interpretation of the results of some recent experiments with repeated tacit coordination games. These experiments revealed several behavioral regularities, including a systematic discrimination between strict Nash equilibria in certain games, that appear to be driven by strategic uncertainty, and are not explained by traditional equilibrium refinements. The observed patterns of discrimination correspond closely to predictions based on Maynard Smith's notion of evolutionary stability. An adaptive model, in the spirit of the evolutionary dynamics but recognizing the important differences between learning in human populations and evolution, promises to yield a unified explanation of the results. *Journal of Economic Literature* Classification Numbers: 020, 026. © 1991 Academic Press, Inc.

1. INTRODUCTION

In recent years there has been an upsurge of interest in coordination. In economics, coordination problems are normally modeled as noncooperative games whose players' preferences are coherent enough to make any Pareto–efficient strategy combination a Nash equilibrium, but in which there are two or more (equilibrium) combinations at which each player's

strategy is uniquely optimal if the others correctly anticipate it, but not in general otherwise. Coordination games involve no incentive problems, as these are normally characterized, because their efficient outcomes are supportable as equilibria. Nevertheless, playing them often involves real difficulties. Similar difficulties lie at the heart of many important questions usually analyzed under the assumption that players can coordinate their strategy choices on any desired equilibrium. These include the design of optimal incentive schemes; the characterization of which outcomes can be supported by implicit contract in a long-term relationship; and the determination of whether, and how, bargainers share the surplus from making an agreement.

Convincing answers to these questions must go beyond the observation that if rational players have commonly known, identical beliefs about how a game will be played, those beliefs must be consistent with some equilibrium in the game. But the traditional approach to analyzing games with multiple equilibria relies on "refining" Nash's notion of equilibrium until (ideally) only one survives; and the most extensively studied and widely accepted refinements are designed mainly to ensure that, if a particular equilibrium is already expected to govern play, it is actually in each player's interest to play his equilibrium strategy.[1] Coordination games, by definition, have more than one equilibrium that passes this kind of test. The usual refinement techniques are therefore not much help in playing them. Players in a coordination game can therefore be expected to be uncertain about how other players will respond to its multiplicity of equilibria, even when they have complete information. This *strategic uncertainty* undermines the standard argument that rationality requires players to play according to a particular equilibrium, and even calls into question the rationality of playing an equilibrium strategy.

Because standard techniques do not yield convincing characterizations of behavior in coordination games, information from other sources about how they are played is crucial in answering questions that turn on coordination. Van Huyck, Battalio, and Beil (henceforth "VHBB") (1990, 1991) have recently conducted a series of experiments whose designs were especially well chosen for learning about behavior under strategic uncertainty. The remarkable results they obtained may be the cleanest experimental evidence on coordination now available. In one set of experiments, their subjects discriminated sharply and systematically between *strict* Nash equilibria—equilibria in which each player's strategy is a unique best reply to the other players'. This is one of several behavioral regularities VHBB found that are not explained by traditional game the-

[1] Harsanyi and Selten's (1988) "general theory of equilibrium selection" is a notable exception. Kohlberg and Mertens (1986) explain why the optimality of playing equilibrium strategies is not an immediate consequence of the definition of equilibrium, and provide a good introduction to the literature on equilibrium refinements.

ory, but which make perfect sense (as I shall argue) when strategic uncertainty is not assumed away.

VHBB took considerable care in interpreting their results, and the interpretations they offered add a great deal to the value of their experiments. However, an adaptive analysis in the spirit of evolutionary game theory (see, for example, Maynard Smith, 1982), although not at all inconsistent with VHBB's interpretations, seems to sharpen them significantly. This paper outlines such an analysis, and discusses the behavioral assumptions required for an adaptive explanation of VHBB's results. The adaptive analysis reflects the differences between learning in the finite human populations of the experiments and evolution in the large animal populations usually studied in evolutionary game theory. Indeed, although an uncritical application of Maynard Smith's notion of evolutionary stability predicts behavior much like that VHBB observed, a full explanation appears to depend on recognizing some important differences between learning and evolution.

The rest of this paper is organized as follows. Section 2 summarizes VHBB's experimental designs and results, and reviews their comparisons of their results with the predictions of traditional game theory. Section 3 provides an introduction to the relevant parts of evolutionary game theory, discusses the relationship between evolutionary stability and Nash equilibrium, and compares predictions based on evolutionary stability with VHBB's results, finding a surprisingly close correspondence. Section 4 discusses the differences between evolution and learning in VHBB's experimental environments and outlines an adaptive model that reflects these differences and suggests a simple, unified explanation of their results. This explanation assigns a central role to strategic uncertainty, and suggests that VHBB's subjects had a clearer understanding of how it affected their incentives than is consistent with a traditional analysis of the games they played.

2. VAN HUYCK, BATTALIO, AND BEIL'S EXPERIMENTAL DESIGNS AND RESULTS

This section summarizes VHBB's (1990, 1991) experimental designs and results and reviews their comparisons of the behavior they observed with the predictions of traditional game theory.[2] In the experiments, groups of 2 to 27 subjects played series of one–stage simultaneous–move coordination games, each symmetric across players, with no communciation al-

---

[2] VHBB (1990, 1991) provide more complete descriptions of the experiments and results. VHBB (1989) report the results of some interesting additional experiments (not discussed here) in which the right to participate in VHBB's (1991) games was auctioned off before play, with surprising consequences for observed behavior.

lowed before or during play. Most of these games had multiple strict Pareto–ranked pure–strategy Nash equilibria; most also had multiple mixed–strategy equilibria, which of course were not strict. (In what follows, the term "game" is reserved for stage games, and "repeated game" is used for series of stage games. "Treatment" refers to a series of plays of a given stage game by a group of subjects.) Each experiment involved one or more treatments in sequence, each typically lasting between 5 and 10 stages. With one exception, noted below, subjects played with the same group in every stage of a given treatment. A summary statistic of subjects' strategy choices, described below, was publicly announced after each stage. The structure of each treatment was publicly announced at the start, with two apparently unimportant exceptions: The number of stages in a treatment was not announced in advance in some experiments, and the structures of future treatments were never announced until they were about to begin. At the end of each experiment, subjects were paid the (undiscounted) sum of their payoffs in the games they played. There was ample evidence that they understood the rules, and that they were paid well enough to induce the intended preferences.

In the descriptions that follow, I use VHBB's notation whenever possible. In each of the games used in the experiments, each player i chose a pure strategy, denoted e_i, from the set $\{1, \ldots, 7\}$. (Mixed strategies were presumably also feasible, but allowing them adds nothing to—and takes nothing away from—the explanations offered here.) The value of e_i is called player i's "effort" for reasons made clear below. In each stage, each player's payoff was determined by his own effort and a simple summary statistic of those of the players in his group. In the experiments reported in VHBB (1990) this statistic was the group minimum effort, denoted $\min(e_i, e_{-i})$, where $e_{-i} \equiv \min(e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n)$; in those reported in VHBB (1991) (in which the number of players was always odd) it was the group median effort, denoted M. With minor exceptions, only these statistics were announced after each stage. The announcements in each case contained all directly payoff–relevant information, and VHBB's (1990) comparison of the results with those of the exceptional treatments, in which entire effort profiles were announced, suggests that announcing entire profiles would have made little difference.

In VHBB's "minimum" experiments, subject i's stage-game payoff was given by $\pi(e_i, e_{-i}) \equiv a[\min(e_i, e_{-i})] - be_i + c$, $i = 1, \ldots, n$, with $a > b \geq 0$. VHBB motivated this specification by noting that this kind of game was suggested by Bryant (1983) (see also Cooper and John, 1988) as a model of Keynesian effective demand failures. This game has a long history, which can be traced to the Stag Hunt example Rousseau (1973, p. 78; originally published in 1755) used in his discussion of the origins of the social contract. Imagine (adding a bit of game–theoretic detail to Rousseau's discussion) that each of a group of hunters must independently decide how much of his effort to devote to hunting a stag with the others.

Hunting a stag requires the cooperation of each member of the group, with the minimum of members' efforts determining the probability of success. The opportunity cost of effort devoted to the stag hunt is determined by the expected benefit of hunting rabbits, which requires no cooperation. The effort e_i in VHBB's minimum game then corresponds to the amount of effort hunter i allocates to the stag hunt, the parameter a reflects its expected benefit, given the e_i, and the parameter b reflects the expected benefit of hunting rabbits. In more prosaic terms, the minimum game is a model of team production with equal output shares and costly effort. Output more than repays players' effort costs if all work equally hard, but the minimum effort in the team creates a bottleneck, so that if anyone shirks, the balance of the others' efforts is wasted.

There were four treatments in the minimum experiments, called A, A', B, and C. A given set of 14–16 subjects normally participated in a sequence of treatments, either $\{A\}$, $\{A, B, A'\}$ or $\{A, B, A', C\}$. Treatments A, A', and B used "large" groups consisting of the entire set of 14–16 subjects; treatment C used "small" groups of 2 (or in one case, 3) subjects, randomly selected from that set. The members of a pair were told only the pair's minimum effort after each stage. There were in fact two versions of treatment C. In one, here called "C_d," subjects' pairings were different in each stage; in the other, here called "C_f," they were fixed.

The payoff parameter c was set at $0.60 in each treatment to avoid negative payoffs. In treatments A and C, the parameter a was set at $0.20, and the parameter b was set at $0.10.[3] In treatment B, b was lowered to $0.00, making $e_i = 7$, the highest feasible effort (and the only one consistent with efficiency), a weakly dominant strategy. Treatment A' restored the payoffs of treatment A following a switch to those of treatment B. For any number of players, the payoffs for treatments A, A', and C were as follows:

		Group minimum effort						
		7	6	5	4	3	2	1
	7	1.30	1.10	0.90	0.70	0.50	0.30	0.10
	6		1.20	1.00	0.80	0.60	0.40	0.20
	5			1.10	0.90	0.70	0.50	0.30
Subject's effort	4				1.00	0.80	0.60	0.40
	3					0.90	0.70	0.50
	2						0.80	0.60
	1							0.70

[3] Despite these low stage-game payoffs, the range of possible payoffs over the course of an experiment gave the subjects (undergraduates at Texas A&M University) fairly strong incentives to maximize their payoffs: A subject's possible total earnings in a typical 2-hr experiment ranged from $5.74 to $35.30.

In VHBB's "median" experiments, subject i's stage-game payoff was normally given by $\rho(e_i, M) \equiv \alpha M - \beta[M - e_i]^2 + \gamma$, with $a > 0$ and $\beta \geq 0$. (I depart from their notation here to avoid confusion with the function $\pi(\cdot)$ and the parameters a, b, and c in VHBB (1990).) Thus, other things equal, the closer subjects' efforts were to the group median effort, the higher their payoffs, with higher common effort choices yielding higher payoffs. As VHBB suggest, this game resembles the newspaper competitions to which Keynes (1936, Chap. 12) likened professional investment,

> . . . in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view.

There were four different median treatments, called Γ, Γdm, Ω, and Φ. A given set of subjects normally participated in a series of treatments, either $\{\Gamma, \Omega, \Gamma\}$, $\{\Gamma$dm$, \Omega\}$, $\{\Omega, \Gamma\}$, or $\{\Phi, \Gamma\}$. These treatments used groups of 9 subjects, with a fixed group playing in each stage of a given treatment, except that in the "dual-market" treatment Γdm each subject played two games simultaneously (each with the same payoff table as treatment Γ), one in a group of 9 and one in a group of 27. VHBB combined the results of treatments Γ and Γdm because they found no statistically significant differences between them; I follow their lead in this.

In each median treatment, γ was set at \$0.60 or \$0.65 to avoid negative payoffs. In treatments Γ and Γdm, the payoff parameter α was normally set at \$0.10 and β was normally set at \$0.05. In treatment Φ, α was reduced to \$0.00. In treatment Ω, subject i's payoff was \$0.00 unless $e_i = M$, in which case his payoff was given by $\rho(e_i, M)$ with $\beta = \$0.00$ and α either reduced to \$0.05 or left at \$0.10. Thus, treatments Γ and Γdm combined a preference for a higher median effort, other things equal, with increasingly severe penalties for being further and further away from the

<table>
<tr><td></td><td></td><td colspan="7">Group median effort</td></tr>
<tr><td></td><td></td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td></tr>
<tr><td></td><td>7</td><td>1.30</td><td>1.15</td><td>0.90</td><td>0.55</td><td>0.10</td><td>−0.45</td><td>−1.10</td></tr>
<tr><td></td><td>6</td><td>1.25</td><td>1.20</td><td>1.05</td><td>0.80</td><td>0.45</td><td>0.00</td><td>−0.55</td></tr>
<tr><td></td><td>5</td><td>1.10</td><td>1.15</td><td>1.10</td><td>0.95</td><td>0.70</td><td>0.35</td><td>−0.10</td></tr>
<tr><td>Subject's effort</td><td>4</td><td>0.85</td><td>1.00</td><td>1.05</td><td>1.00</td><td>0.85</td><td>0.60</td><td>0.25</td></tr>
<tr><td></td><td>3</td><td>0.50</td><td>0.75</td><td>0.90</td><td>0.95</td><td>0.90</td><td>0.75</td><td>0.50</td></tr>
<tr><td></td><td>2</td><td>0.05</td><td>0.40</td><td>0.65</td><td>0.80</td><td>0.85</td><td>0.80</td><td>0.65</td></tr>
<tr><td></td><td>1</td><td>−0.50</td><td>−0.05</td><td>0.30</td><td>0.55</td><td>0.70</td><td>0.75</td><td>0.70</td></tr>
</table>

median. Treatment Φ eliminated the preference for a higher median, and treatment Ω imposed different penalties for being away from the median, generally higher than in treatments Γ and Γdm but independent of the distance. The payoffs for treatments Γ and Γdm were as shown in the table at the bottom of p. 30.

Before summarizing VHBB's experimental results, it is useful to review the implications of traditional game theory for behavior in the games they studied. I restrict attention to pure strategies; it can be shown that the conclusions remain valid when mixed strategies are allowed. I also restrict attention to stage-game strategies. Repeated-game strategies are theoretically potentially relevant in most of VHBB's treatments. But allowing them seems to help in understanding the results only in minimum treatment C_f.[4]

Recall that in my terminology, "games" are stage games. In every game used in VHBB's experiments but that of minimum treatment B, the strategy combinations in which each player chooses the same effort are the only pure-strategy Nash equilibria, and each of these is strict, and therefore trembling-hand perfect, proper, and strategically stable (taken by itself).[5] In the game used in treatment B, any strategy combination in which more than one player's effort equals the group minimum is a Nash equilibrium, but the combination in which each player chooses the highest effort, $e_i = 7$, is the only one that does not involve weakly dominated strategies. Because this equilibrium is strict, it (unlike the others in this game) also satisfies the above refinements.

The only equilibrium refinement that discriminates between the strict equilbria in VHBB's games is Harsanyi and Selten's (1988) general theory of equilibrium selection. In every game but that of median treatment Φ (in which all of the pure-strategy equilibria yield the same payoffs), the "payoff-dominance" principle they incorporated into their theory selects the equilibrium in which each player chooses the highest effort, because only that equilibrium is Pareto-efficient in the set of equilibria.[6]

---

[4] The learning models proposed below of course resemble repeated-game strategies, in that they make players' current choices history-dependent. Except in treatment C_f, however, VHBB's subjects appeared to use history mainly to predict how the other subjects would play, not because they hoped thereby to influence their future choices.

[5] Trembling-hand perfectness follows immediately for strict equilibria, because small trembles in players' strategies do not alter their strict preferences for the equilibrium strategies; properness follows by a similar argument. It can be shown that each pure-strategy equilibrium is a strategically stable set, taken by itself, by "pruning" pure strategies that are never weak best responses.

[6] Harsanyi and Selten's theory gives payoff-dominance first priority, subject to the qualification that perfect symmetries and the absence of a common-knowledge labeling of strategies may make it impossible for players to be sure of coordinating their strategy choices (see also Crawford and Haller, 1990). In such cases, they apply payoff-dominance after an

These conclusions, like the rules of the games are used in the experiments, are readily accessible to common sense. (The simplicity of the games VHBB studied probably increased the clarity of their results, by allowing subjects to rely on intuitions gleaned from everyday experience more than they could in experiments with games that cannot be described in simple common-sense terms.) In each case the efficient outcome is an equilibrium, so there are no incentive problems as they are normally characterized in economic theory. Moreover, the efficient equilibrium is the best possible outcome for all players, and the argument in favor of playing it does not depend on a deep understanding of game theory: it is obviously the "correct" coordinating principle.

To realize the efficient outcome, however, players must overcome a more subtle kind of incentive problem. In the large-group minimum game of treatments A and A', for instance, the highest effort's higher payoff when all players choose it must be traded off against its greater risk of lower payoffs if they do not. (This risk is entirely due to strategic uncertainty, hence not reflected in a traditional analysis.) For a player to find it rational to choose the highest effort, treating the influence of his choice on future developments as negligible, he must believe that the correctness of this choice is sufficiently obvious that it is likely that *all* of the other players in his group will believe that its correctness is sufficiently obvious to all. Although the payoff-dominance principle that dictates this choice is not at all subtle and its rationale does not depend on the number of players, when playing games like the minimum game most people appear to consider the highest effort a good bet in small groups, but not in large groups.[7]

The strategic uncertainty that underlies this incentive problem can profoundly affect behavior. Despite payoff-dominance, in VHBB's large-group minimum treatments A and A' subjects initially chose widely dispersed efforts and then rapidly approached the lowest effort, $e_i = 1$: In treatment A, 72% of the subjects reached that effort within 10 stages and it was the modal 10th-stage effort in every experiment; in treatment A', 84% of the subjects reached that effort within 5 stages and it was the modal

---

expected-payoff calculation that takes this uncertainty into account. This qualification is relevant here only for median treatment Φ; its relevance is limited even there, because in that treatment VHBB's subjects appeared to use the labeling VHBB announced to coordinate their strategy choices.

[7] Note that these beliefs are self-confirming! They are plausible because if players choose independently, with probabilities independent of the number of players, the obviousness of the principle is less likely to be sufficient, the larger the group. But this intuition concerns a choice between strict equilibria, hence is not captured by traditional refinements (with the partial exception of the "risk-dominance" part of Harsanyi and Selten's theory, discussed further in Section 4).

5th-stage effort in every experiment. By contrast, in the intervening minimum treatment B, in which the payoff-dominant highest effort was also a weakly dominant strategy, 84% of the subjects initially chose that effort and 96% of them reached it by the 5th stage. Thus, treatments A and A' revealed a remarkable behavioral regularity that is not explained by traditional equilibrium refinements, and is at variance with the only such refinement (payoff-dominance) that discriminates between the strict equilibria in these treatments. The contrasting regularity observed in treatment B was very close to a unique prediction based on weak dominance (and therefore reinforced by strategic uncertainty). This suggests that the very different results observed in treatments A and A' were due not to subjects' failure to understand their incentives, but to strategic uncertainty.

The dynamics were very different in the small groups of minimum treatments C_d and C_f, which had the same payoff table as treatments A and A', and in the various median treatments, which had comparable numbers of players. In each case subjects' initial effort choices varied substantially. In treatment C_d subjects' choices drifted over time with no clearly discernible trend, while in treatment C_f 90% of the subjects reached the highest effort within seven stages. By contrast, in every median experiment subjects converged completely, within 5 or 10 stages (with three minor exceptions), to the Nash equilibrium determined by the historical accident of their initial treatment median, despite considerable variation in the initial median across treatments.

Two features of these results are particularly noteworthy. There is a striking contrast between subjects' behavior in large- and small-group minimum games, with large groups discriminating very sharply between strict equilibria and small groups discriminating either much less sharply or in the reverse direction, depending on whether or not their pairings were fixed. There is also a striking contrast between subjects' behavior in large-group minimum games and median games, with no history-dependence whatsoever in the former but very strong history-dependence in the latter. There can be little doubt that these results are replicable. Similar results have been obtained in formal experiments with games like the large-group minimum game by Isaac *et al.* (1989), and with games like the median game by VHBB (1989); these are easily reproduced in informal experiments with simplified versions of VHBB's games. Moreover, informal experiments and the results of VHBB's treatment A' (which was preceded by an A treatment with the same structure and subjects) suggest that the patterns of behavior VHBB observed do not disappear when subjects learn to anticipate them; if anything, they are reinforced. Thus, although VHBB's subjects' behavior is not fully consistent with any traditional notion of rationality, it passes what is arguably the most reasonable test of rationality in coordination games.

3. EVOLUTIONARY GAME THEORY

Explaining VHBB's experimental results promises to help in under-
standing how the difficulty of coordination influences strategic behavior.
It may also shed some light on how group size affects cooperation, an-
other important question that has long resisted analysis along traditional
lines. In this section and the next, I outline a unified adaptive explanation
of the behavior they observed.

My basic premise is that in environments with strategic uncertainty,
players' expectations are likely to be heavily influenced by what they can
learn about other players' likely strategy choices from their experience
with analogous games. To learn in this way, players must assume, at least
implicitly, that some aspect of the stochastic process that describes other
players' behavior remains constant over time. In environments with the
repeated-game structure of VHBB's experiments, previous plays of the
stage game provide the closest analogies. It is therefore natural (except
when repeated-game strategies are plainly essential, as in treatment C_f) to
begin by considering models in which stage-game strategies are the ob-
jects of choice, with players adjusting them over time on the assumption
that the distribution of the other players' strategy choices will remain
approximately constant from one stage to the next.

Models with these features, whose assumptions correspond closely to
the structures of VHBB's experimental environments, were introduced
by Schelling (1973; 1978, Chap. 7) and Maynard Smith (1974, 1982). Pre-
dictions based on Maynard Smith's notion of evolutionary stability are
surprisingly close to the patterns of discrimination between Nash equilib-
ria in VHBB's experiments. Understanding why this is so is helpful in
constructing an adaptive explanation of VHBB's results, even though the
true explanation cannot be literally evolutionary. This section provides an
introduction to evolutionary game theory and shows how it can be applied
to VHBB's games; my discussion draws on Crawford (1990b), which
provides more detail and further references. Section 4 then discusses the
differences between learning in human populations and evolution, and the
changes in the model required for a full explanation of VHBB's results.

Unless otherwise noted, "game" and "strategy" refer to the stage
game. In the simplest versions of Schelling's and Maynard Smith's
models, a population of identical players repeatedly play a symmetric
game. (This specification includes, but is not limited to, the familiar case
in which pairs of players are randomly selected from the population to
play a two-person game.) Players' strategies are identified, so that it is
meaningful to say that different players choose the same strategy, or that
a player chooses the same strategy in different stages. Individual players
play only pure strategies, with each player's payoff determined by his own

strategy and the population frequencies of the other players' strategies, so that it make no difference "who did what" among the others. The population is *large*, in the sense that each player's strategy has a negligible effect on the population strategy frequencies, the population frequencies can be treated as continuous variables, and the average payoff of those using each strategy can be identified, via the law of large numbers, with its mathematical expectation. Finally, the dynamics have the property that strategies with higher current payoffs increase in frequency over time. In Maynard Smith's model, this follows from the assumptions that players inherit their parents' strategies and reproduce asexually, at rates determined by their payoffs. In Schelling's model, it is an aggregate consequence of a plausible hypothesis about individual strategy adjustment.[8]

It is natural to hope that the dynamics of the population strategy frequencies are convergent, so that their limiting behavior can serve as a simple summary of the model's implications. Because global stability is usually too much to hope for in coordination games, Schelling focused on frequencies that are locally stable, in the standard sense that the population returns to them, at least asymptotically, after any sufficiently small departure. This idea of local stability of the population strategy frequencies is also the motivation for Maynard Smith's notion of evolutionary stability. He defined an *evolutionarily stable strategy* (or "ESS") as a strategy (pure or mixed) that, if played by all members of a large population, has strictly higher payoff than any mutant strategy that enters the population with sufficiently low frequency.[9]

Maynard Smith's definition rests on the intuition that if the members of a population all play an ESS, any mutants that enter the population with low frequency reproduce more slowly than players who play the ESS. As a result, the relative frequency of the mutants approaches zero over time, restoring the population strategy frequencies of the ESS. This intuition has since been formalized by Taylor and Jonker (1978), Zeeman (1979), and Hines (1980a, 1980b), who showed (with varying degrees of generality) that evolutionary stability corresponds closely to local stability of the

---

[8] This hypothesis finds considerable support in VHBB's observations of their subjects' behavior and predictions of each other's behavior, and in other experimental work; see Banks *et al.* (1988), Isaac *et al.* (1989), and Selten and Stoecker (1986).

[9] A formal definition is given below. Although Maynard Smith's definition refers to a *monomorphic* population, in which all players play the same strategy in equilibrium, a mixed-strategy ESS can also be used (as illustrated below) to characterize the limiting behavior of a *polymorphic* population, in which players play different strategies in equilibrium. In a polymorphic equilibrium, evolutionary stability requires that on average, the members of the population have higher expected payoff than any mutant. I allow mixed-strategy ESS's, despite my simplifying assumption that players play only pure strategies, to avoid arbitrarily ruling out such equilibria.

evolutionary dynamics; Crawford (1989, Sect. 1) provides a brief survey and further references.

It is well known (see, for example, Maynard Smith, 1982, p. 14) that there is also a close correspondence between evolutionary stability and Nash equilibrium. For instance, in the random-pairing model referred to above, with a large population, a strategy that is not a best response to itself can clearly be invaded by a low-frequency mutation that is a best response to that strategy. It follows that in this model, an ESS must correspond to a Nash equilibrium in the two-person game played by paired players; the definition requires that this equilibrium be symmetric across players. It can also be shown that any strict, symmetric Nash equilibrium corresponds to an ESS. It is shown below that this correspondence extends to a much wider class of models. The results of Taylor and Jonker, Zeeman, and Hines can thus be viewed as an adaptive justification for Nash equilibrium.

Schelling's and Maynard Smith's models differ from traditional game theory in two important ways. First, it is evident that their characterization of behavior is not "rational" in any of the traditional senses. Strategies are not even chosen in Maynard Smith's model. Although they are chosen in Schelling's model, they are not in general best responses to rationality-based predictions about their current and future consequences. Instead, the population in effect comes to know the equilibrium strategy frequencies as they evolve: players' experience with analogous games is substituted for deductions based on rationality.

Some such departure from traditional notions of rationality appears to be essential for a full understanding of how strategic uncertainty affects behavior.[10] It is clear from Section 2's arguments that ignoring the dynamics of learning and applying traditional equilibrium refinements to the stage game leaves far too much unexplained. And although there is no obstacle in principle to a traditional analysis of the entire learning process, in which players choose decision rules that are in (suitably refined) equilibrium in the repeated game, a repeated game normally inherits the multiple-equilibrium problems of the stage games that make it up. Such an approach compounds coordination problems rather than resolves them, and requires the unrealistic, question-begging assumption that players' expectations are coordinated at the start of play—usually, in practice, at an equilibrium selected by applying refinements that are largely insensitive to strategic uncertainty.

---

[10] Such a departure would be difficult to defend, for applications involving human players, if it persisted indefinitely. In the games considered here, however, the dynamics normally converge to a Nash equilibrium in the stage game.

The second important difference is that, in conjunction with evolutionary stability, the labels that identify players' strategies have substantive implications: they are the language in which conventions that emerge to solve coordination problems must be expressed. Their role is best illustrated by considering some examples from the theory of traffic. Imagine a large population of identical players, repeatedly paired at random to play one of the following two-person games:

	G	S
G	0 0	1 1
S	1 1	0 0

Intersection

	L	R
L	1 1	0 0
R	0 0	1 1

Confrontation

These games are stylized versions of the situations drivers face when meeting on different roads at an intersection, and when confronting each other head-on on a single road; the labels G and S refer to Go and Stop, and L and R refer to driving on the Left and on the Right. In Intersection, but not in Confrontation, players must choose strategies with different labels to achieve efficient coordination, even though they have no way to distinguish their roles in the game. Because this is "only" a difference in labeling, these games would be treated as equivalent in a traditional analysis (see, however, Harsanyi and Selten, 1988). But it is clear that this kind of difference can matter a great deal in practice.

Evolutionary game theory models the effects of this difference by ruling out systematic differences between players' strategy choices when they have no way to distinguish their roles. In Intersection, for instance, it is easy to verify from the definition, or directly from the population dynamics, that the only ESS is "G with probability $\frac{1}{2}$"; the resulting strategy frequencies can be realized as a pure-strategy polymorphism, with half the players choosing G and half choosing S.[11] Although individual players play only pure strategies in this configuration, the symmetry across players imposed by evolutionary stability rules out both of the pure-strategy Nash equilibria in the two-person game.

---

[11] A strictly evolutionary analysis would allow these frequencies to be realized in many other ways, including in particular each player playing the mixed-strategy ESS. It is argued in Crawford (1989), however, that sensible learning dynamics must converge, if at all, to a configuration in which individual players play only pure strategies.

In Confrontation, by contrast, there are two ESS's: "L with probability one" and "R with probability one"; the associated strategy frequencies can be realized only by players all choosing L or all choosing R. These outcomes of course correspond to the symmetric pure-strategy Nash equilibria in the two-person game. In this case, however, there is no ESS that corresponds to the symmetric mixed-strategy equilibrium in that game.

It is intuitively clear that players might be able to coordinate their decisions better in Intersection if they could find some way to distinguish their roles when they meet. Evolutionary game theory models the effect of such a coordination device by building it into the game and then requiring symmetry across players in the larger strategy space that results. Installing a stoplight in an intersection, for instance, allows a player to make his decisions contingent on what color light he sees. Assuming that the light is green for one player at the intersection if and only if it is red for the other, and that it is red and green with equal probabilities for both, yields an expanded Intersection game with payoff matrix:

	G,G	S,S	G,S	S,G
G,G	0	1	1/2	1/2
S,S	1	0	1/2	1/2
G,S	1/2	1/2	1	0
S,G	1/2	1/2	0	1

Intersection with stoplight

("G,S" stands for the strategy "Go on green, Stop on red"; "S,G" stands for "Stop on green, Go on red"; and so on. The single number in each cell represents both players payoffs for the associated strategy combination.) In the expanded strategy space, symmetry across players is consistent with efficient coordination. It is easy to show that evolutionary stability actually requires players to use the stoplight to solve the coordination problem in the familiar way—though not necessarily with our conventional meanings of red and green.

To see how these ideas can be applied to VHBB's experimental environments, it is best to begin with a simplified version of the minimum game used in their treatments A, A' and C, which I shall call the "Stag Hunt." Consider a group of players, each of whom must choose simultaneously between two efforts, 1 and 2. As before, the minimum of their chosen efforts determines total output, which they share equally. Each

player's effort cost increases with effort, but effort is sufficiently productive that if all players in the group choose the same effort, the resulting output shares more than repay the cost. Assume, for definiteness, that output *per capita* is twice the minimum effort and that each player's unit cost of effort is one. Then each player's payoffs, for any finite number of players, each player's payoffs are as follows:

Group minimum effort

		2	1
Player's effort	2	2	0
	1	1	1

Stag Hunt

This game has two pure-strategy Nash equilibria, one in which all players choose effort 2 and one in which all choose effort 1; all players prefer the former to the latter.

It is important in what follows to distinguish between two population models. The *random-pairing* model, in which players are randomly selected from the population and paired to play a symmetric two-person game, has already been discussed. The *simultaneous-play* model allows players to interact in any way that can be described as simultaneous play, by all members of the population, of a single symmetric game.[12]

The random-pairing Stag Hunt is a simplified model of VHBB's treatment C_d and the simultaneous-play Stag Hunt is a simplified model of their treatments A and A'. When players have only two pure strategies and the population is large, a graphical stability analysis of these models is possible. (The same conclusions can be obtained directly from the ESS definition.) Figure 1 graphs the expected payoffs of efforts 1 and 2 for a representative player in the random-pairing Stag Hunt against the population frequency of effort 1. With random pairing, a player's expected payoff is a population frequency-weighted average of his effort's payoffs

---

[12] The simultaneous-play model was introduced in economics by Schelling (1973; 1978, Chap. 7), and in biology (where it is called the "playing the field" model) by Maynard Smith (1982, pp. 23–27), who attributed it to Hammerstein. Random pairing can be seen as a special case of simultaneous play by thinking of players' interactions as a single symmetric game, in which players choose their strategies before the uncertainty of pairing is resolved. Both models can also be viewed as special cases of a still more general model, in which randomly selected n-person subsets of the population play an n-person game, as in Schaffer (1988). (Curiously, this model can also be viewed as simultaneous play, by the above device.)
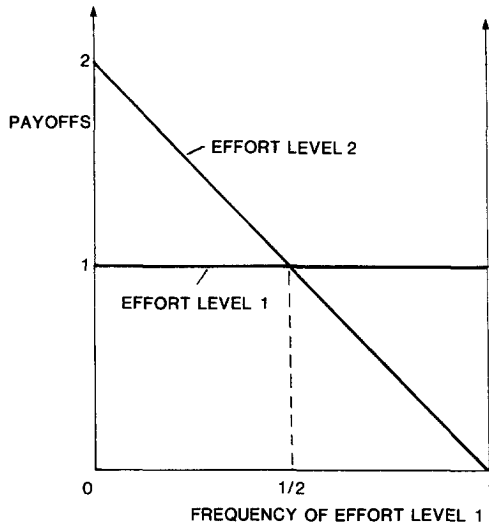
VINCENT P. CRAWFORD

FIG. 1. Stag Hunt with random pairing.

when his partner chooses efforts 1 and 2, and therefore a linear function of the population frequencies. Figure 2 graphs the payoffs for a representative player in the simultaneous-play Stag Hunt. The payoff of effort 1 is again constant, but there is now a discontinuous drop in the payoff of effort 2, which falls from 2 to 0 as the frequency of effort 1 rises above 0. This is due to the fact that in the simultaneous-play Stag Hunt each player's payoff can be substantially altered by a small change in the other players' strategy frequencies, even in a large population.[13]

To avoid confusion with Nash equilibrium, I shall call an equilibrium of the dynamics of the population strategy frequencies a *stationary point*. It is clear from Fig. 1 that there are three stationary points in the random-

[13] This discontinuity is not pathological from the point of view of traditional game theory. Although the payoffs in this game can be represented, for any number of players, only by a discontinuous function of the population strategy frequencies, this kind of continuity is not a substantive restriction if the size of the population is bounded. For any given finite number of players, the payoff function yields a well-defined game that satisfies all of the assumptions of traditional game theory. (Allowing mixed strategies in this game would not alter my conclusions, and would make players' payoffs continuous in the traditional sense in the larger strategy space.) The difference between the simultaneous-play and random-pairing Stag Hunts is that in the former model, small individual probabilities of choosing effort 1 imply a much larger probability that the group minimum effort is 1, while this effect is attenuated by sampling in the latter model. Thus, in terms of the most general model discussed in footnote 12, the essential difference between these two models is whether the subsets of the population that play the Stag Hunt game are large or small.
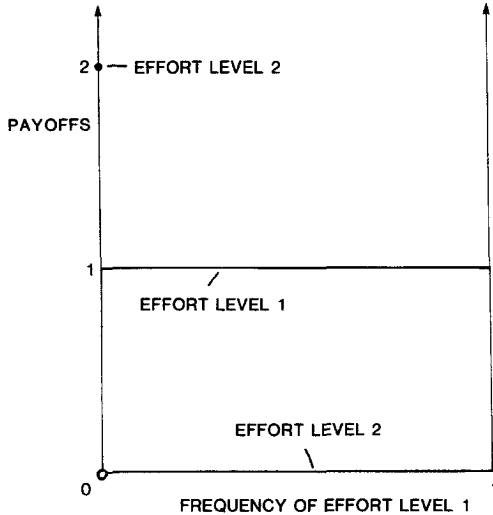
FIG. 2. Stag Hunt with simultaneous play.

pairing Stag Hunt, at population frequencies 0, $\frac{1}{2}$ (the frequency that equalizes the expected payoffs of the two efforts), and 1. Each of these corresponds to a Nash equilibrium in the two-person Stag Hunt game. The stationary points at frequencies 0 and 1 are stable, and the stationary point at frequency $\frac{1}{2}$ is unstable. The outcome is determined by the initial population strategy frequencies. If (as informal experiments suggest) most players initially choose effort 2, the population converges to the efficient Nash equilibrium. It is also possible, however, for the population to get stuck forever at the inefficient equilibrium.

In the simultaneous-play Stag Hunt there are two stationary points, at population frequencies 0 and 1. Each corresponds to a Nash equilibrium in the large-group Stag Hunt game. The stationary point at frequency 1, which corresponds to the inefficient Nash equilibrium, is stable. However, the stationary point at frequency 0, which corresponds to the efficient Nash equilibrium, is unstable, because if even a single player switches to effort 1, it then has the higher payoff. Thus, even though the efficient equilibrium is strict, unless *all* players initially choose effort 2 the population converges to the inefficient equilibrium.

Although these conclusions are encouragingly close to VHBB's results, the informal analysis just given raises as many questions as it answers: Why does selection, in which a "beggar-thy-neighbor" mutation that lowers a player's payoff but lowers other players' payoffs even more can be evolutionarily advantageous, lead to Nash equilibrium, in which play-

ers maximize purely individualistic payoffs? Does the tendency of the population dynamics to converge to Nash equilibrium extend to the finite populations, larger strategy spaces, and more complex payoff determination of VHBB's experimental environments? How does evolutionary stability discriminate between equilibria in those environments?

These questions can be answered only by a formal analysis. It is simplest to begin with large populations. The definitions that follow apply to simultaneous play as well as random pairing. Let the probability vectors r and s represent a player's mixed strategy and the population strategy frequencies, respectively, and let $E(r|s)$ denote the payoff of a player playing r when the population frequencies are given by s. (In a large population, it does not matter whether or not s includes the player's own strategy, because it is a negligible part of the population frequencies.) I shall abuse terminology somewhat by calling $E(\cdot|\cdot)$ the *payoff function*. In the random-pairing model, $E(r|s)$ is linear in s for any given value of r, as noted above; and the linearity in probabilities of expected payoffs makes $E(r|s)$ linear in r for any given value of s. $E(r|s)$ is also linear in r in the simultaneous-play model, for the same reason; but it is in general nonlinear in s.

A large-population ESS can be formally defined as a mixed strategy, p, such that, for each mixed strategy $q \neq p$, there exists an $\bar{\varepsilon}(q) > 0$ for which

$$E(p|(1 - \varepsilon)p + \varepsilon q) > E(q|(1 - \varepsilon)p + \varepsilon q) \qquad (1)$$

whenever $0 < \varepsilon < \bar{\varepsilon}(q)$.[14] A large-population ESS is thus a strategy that, if initially played by all members of a monomorphic population, has strictly higher payoff than any mutant strategy that enters the population with sufficiently low frequency, after the mutants enter.

Although the large-population idealization is often helpful, it is not fully adequate for an analysis of VHBB's games. For a player's payoff in a simultaneous-play minimum game to be determined by his own strategy and the population strategy frequencies when each player's strategy is a negligible part of the population frequencies, it must depend on the minimum effort with positive frequency in the population, not on the minimum effort. (An analogous observation applies in defining the payoff function for large-population median games.) But then lowering a player's effort lowers his effort cost without reducing his output share, so the efficient

---

[14] It is sometimes useful to strengthen this definition by requiring that there exists an $\bar{\varepsilon} > 0$, independent of $q \neq p$, such that (1) holds whenever $0 < \varepsilon < \bar{\varepsilon}$. Crawford (1990a) shows that this stronger definition is equivalent to the one in the text for any p for which there exists a neighborhood of $s = p$ throughout which $E(r|s)$ is continuous in s for all r. Thus, this section's results are readily extended to the stronger definition.

strategy combination in which all players choose the highest effort is not a Nash equilibrium, even though it is an equilibrium in any finite population, however large. This discontinuity, in conjunction with the finiteness of the populations in VHBB's (or anyone else's!) experiments, suggests that the large-population assumption should be used with care, checking its conclusions against those of a parallel finite-population analysis.

Schaffer (1988) and Maynard Smith (1988) argued that the large-population ESS definition remains adequate for finite populations (with one qualification, discussed below), provided that individual players' effects on the population strategy frequencies are taken into account in comparing the payoffs of mutants and ESS-players.[15] Let $E(r|s)$ again denote the payoff function, with r representing the player's mixed strategy and s representing the frequencies of the other players' strategies. (This definition is consistent with my notation for large populations.) Assuming that any mutants switch strategies from the ESS, so that the size of the population remains fixed at N players, a finite-population ESS can be defined as a mixed strategy, p, such that for each mixed strategy $q \neq p$,

$$E\left(p \,\middle|\, \frac{N-2}{N-1}p + \frac{1}{N-1}q\right) > E(q|p). \qquad (2)$$

Condition (2) plays the role of condition (1) in the large-population ESS definition, requiring that ESS-players must have greater payoffs than any mutant that enters the population with sufficiently low frequency—in this case, $1/N$.[16]

I now turn to an evolutionary analysis of the games that describe VHBB's experimental treatments. In these games, as in the Stag Hunt models discussed above, the assumption that players play only pure strategies makes it possible to express each player's payoff as a function of the population strategy frequencies and his own strategy, with the form of this function the same for all players and independent of population size.

In the large-group minimum treatments A and A', lower efforts always yield the same output shares as higher efforts, but at lower cost. Thus, in

---

[15] Here, as in Schaffer's and Maynard Smith's analyses, players' realized payoffs are identified with their mathematical expectations, even though the law of large numbers does not fully justify this substitution in a finite population.

[16] For reasons explained in Schaffer (1988) and Crawford (1990b), no nondegenerate mixed strategy p can satisfy the strict inequality in (2) in a random-pairing model. Schaffer (1988) and Maynard Smith (1988) therefore weakened the finite-population ESS definition by allowing equality in (2), as long as p has strictly higher payoff than any mutant that enters the population with frequency higher than $1/N$. The ESS's in VHBB's games are all monomorphic, and all satisfy the strict inequality in (2). Thus, even though I allow mixed-strategy ESS's to avoid ruling out polymorphic equilibria a priori, the simpler definition in the text is adequate for my purposes.

any population, large or finite, a mutation to an effort that is the minimum in the population, after the mutation, then has higher payoff than all other efforts. The strategy configuration in which all players choose the lowest effort is therefore the unique ESS in these treatments.

Because effort is costless in minimum treatment B, it is impossible for a strategy to satisfy the strict inequality in (1) or (2) (or the weaker finite-population ESS definition discussed in footnote 16). There is, therefore, no ESS in this treatment. It is often possible, in cases like this, to restore existence by generalizing the ESS definition to allow sets of strategies that collectively resist invasion by mutants, but not by each other. However, it seems more informative in this case to increase the discriminatory power of evolutionary stability by assuming, following Selten (1983), that players implement their strategies with small probabilities of error. Weakly dominated strategies then have uniformly lower expected payoffs, and the weakly dominant highest effort is therefore the unique ESS.[17]

To see what evolutionary stability entails in treatment C_d, consider a pure-strategy mutant x invading a population playing the pure strategy y. A mutation to $x > y$ plainly yields lower payoff than y, so it can be assumed that $x < y$ without loss of generality. It is easily verified that if $E(r|s)$ denotes the payoff function in treatment C_d and $x < y$, then $E(x|y) \equiv E(x|x) \equiv (a - b)x + c$ and $E(y|x) \equiv ax - by + c$. The ESS condition (2) becomes

$$E\left(y \middle| \frac{N-2}{N-1} y + \frac{1}{N-1} x\right) \equiv \frac{N-2}{N-1} E(y|y) + \frac{1}{N-1} E(y|x)$$

$$\equiv \frac{N-2}{N-1} [(a-b)y + c] \qquad (3)$$

$$+ \frac{1}{N-1} [ax - by + c] > (a-b)x + c \equiv E(x|y),$$

which reduces, when $a > b > 0$, to $N > (2a - b)/(a - b)$. Because $a = \$0.20$ and $b = \$0.10$ in VHBB's minimum experiments, any pure-strategy effort is a monomorphic ESS whenever $N \geq 4$, hence in their experiments with this treatment, in which N ranged from 14 to 16.[18]

---

[17] The error probabilities assumed in this case are clearly a step in the direction of realism, but they may appear ad hoc because they are not used in my evolutionary analyses of VHBB's other treatments. Those analyses, however, rest on strict fitness inequalities, so that small error probabilities would not alter their conclusions.

[18] For completeness, I record the conclusions for $N = 2$ and 3. When $N = 2$, any $x < y$ can invade, because $(2a-b)/(a-b) > 2$; thus, only the lowest effort is an ESS in this case. When $N = 3$, (3) holds with equality when $a = \$0.20$ and $b = \$0.10$, but the more general version of the ESS definition discussed in footnote 16 can be used to show that only the lowest effort is an ESS.

In treatment C_f, unlike in the other treatments, repeated-game strategies were clearly essential, and were recognized as such by VHBB's subjects. Because the experiments were too short to allow significant learning about repeated-game strategies, an evolutionary analysis of this treatment may not be very helpful in explaining VHBB's results. I therefore confine myself to the remark that intuition, aided by the parallel traditional analysis in Crawford and Haller (1990) and the evolutionary analysis in Kim (1990), suggests that the random-pairing model that describes this treatment has evolutionarily stable repeated-game strategies that correspond closely to VHBB's results (and, possibly, some that do not).

In median treatments Γ and Γdm, a player's payoff is always higher, other things equal, when his effort is closer to the group median. (Finite populations are assumed for simplicity to contain odd numbers of players in median games.) It follows that a mutation to an effort that is the median in the population, after the mutation, then has higher payoff than any other effort. Although such a mutation might change the median, in a polymorphic population there must always be at least one player (or set of players with positive frequency in a large population) who can switch strategies to the current median without changing it. It follows that polymorphic ESS's are impossible. Further, because no low-frequency mutation can change the median in a monomorphic population, any strategy configuration in which all players choose the same effort corresponds to an ESS in this game. Similar arguments extend this conclusion to median treatments Ω and Φ.

Thus, the conclusions of the evolutionary analysis correspond closely to subjects' limiting behavior in the minimum experiments. And although evolutionary stability *per se* does not help to explain subjects' behavior in the median experiments, because any strict Nash equilibrium is an ESS in those games, the evolutionary dynamics at least suggest an explanation for the very strong history-dependence observed in the median treatments. When combined with the analysis of subjects' initial choices outlined in Section 4, this yields a plausible interpretation of subjects' behavior.

The fact that evolutionary stability discriminates between strict Nash equilibria in the large-group minimum game of treatments A and A' is crucial to explaining the systematic patterns of discrimination VHBB observed. It is also somewhat puzzling, because this kind of discrimination cannot occur in the large-population random-pairing model most familiar to economists (and biologists). From a more general point of view, it may also seem puzzling that discrimination between widely separated equilibria follows from an ESS definition designed to ensure only local stability of the dynamics. In any case, because any adaptive explanation

of VHBB's results must imply a similar pattern of discrimination between equilibria, it is important to understand why it occurs.

Because the pattern of discrimination takes precisely the same form in large and finite populations, population size seems unlikely to be the key to understanding it. The graphical analysis of the simultaneous-play Stag Hunt model suggests that it is due to discontinuities in the payoff function. I now confirm this by generalizing the correspondence between Nash equilibrium and evolutionary stability from the large-population random-pairing model to large- and finite-population simultaneous-play models in a way that makes clear the role of continuity.

The main result for large populations can be stated as follows:

THEOREM 1. *If p is a large-population ESS and the payoff function, $E(r|s)$, is continuous in s at s = p for all r, then p corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the players in the population. Conversely, if p corresponds to a strict, symmetric Nash equilibrium in that game, and $E(r|s)$ is continuous in s at s = p for all r, then p is a large-population ESS.*[19]

Proof. If p is a large-population ESS and the payoff function $E(r|s)$ is continuous in s at $s = p$ for all r, then letting $\varepsilon \to 0$ in (1) yields

$$E(p|p) \geq E(q|p) \qquad (4)$$

for all $q \neq p$. Because excluding a player's strategy has a negligible effect on the strategy frequencies in a large population, (4) implies that p maximizes any player's payoff, given that the other players are playing p. Thus, p corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the players in the population.

Conversely, if p corresponds to a strict, symmetric Nash equilibrium in that game, so that (4) holds with strict inequality for all $q \neq p$, then the continuity of $E(r|s)$ in s at $s = p$ for all r implies that, for each $q \neq p$, there exists an $\bar{\varepsilon}(q) > 0$ such that (1) holds whenever $0 < \varepsilon < \bar{\varepsilon}(q)$. Thus, p is a large-population ESS. ∎

Theorem 1's continuity condition is implied by the linearity of $E(r|s)$ in s in the random-pairing model, which explains why discrimination between strict Nash equilibria cannot occur in that model. It is clear from

[19] See, for example, Maynard Smith (1982, p. 24) or Schaffer (1988, pp. 470–471), where this correspondence is discussed without examining the role of continuity. Theorem 1 is a true generalization of the familiar correspondence for random pairing: It is easy to show that strategies that are in symmetric Nash equilibrium in the game played simultaneously by the entire population, with the expected payoffs of players' strategies evaluated before they are paired, are also in symmetric Nash equilibrium in the two-person game they play once paired, and vice versa. A somewhat more general version of Theorem 1 is stated without proof in Crawford (1990b).

the Stag Hunt example that this continuity condition need not hold in nonpathologic simultaneous-play models. As the theorem shows, the ESS condition implies the global best-response property of a Nash equilibrium, even though it is designed to ensure only local stability of the population frequencies, because it includes the ability to resist *all* low-frequency mutations. (Low-frequency mutations, even to distant strategies, induce small changes in the population frequencies.)

It is easy to construct examples to show that, without Theorem 1's continuity condition, a large-population ESS need not correspond to a Nash equilibrium. This can be done, for instance, by modifying the simultaneous-play Stag Hunt example, raising the payoff of effort 2 above that of effort 1 when (and only when) the population frequency of effort 1 is 1. Then, as required for effort 1 to be an ESS, mutations at low but positive frequencies have payoffs lower than effort 1's payoff. But any individual player can increase his payoff by switching from effort 1 to effort 2, so that effort 1 does not correspond to a Nash equilibrium. A similar conclusion is reached below for finite populations.

I now turn to finite populations. Schaffer (1988) and Maynard Smith (1988) presented a simple random-pairing example in which the ESS differs from the Nash equilibrium in any finite population, however large, but converges to it as the population grows. Thus, in general, the correspondence between Nash equilibrium and evolutionary stability is at best approximate in finite populations. Theorem 2 shows that, as their example suggests, a correspondence between Nash equilibrium and evolutionary stability like that found in large populations holds approximately for sufficiently large finite populations in any random-pairing or simultaneous-play model with finite numbers of pure strategies, provided that the payoff function satisfies a continuity condition like Theorem 1's. Moreover, this correspondence is exact for strict, pure-strategy equilibria in any sufficiently large finite population. (The correspondence is inexact for Schaffer's and Maynard Smith's example because it had a mixed-strategy equilibrium, which was not strict.)

THEOREM 2. *If a sequence of finite-population ESS's converges, as the population grows, to a limit, p, such that the payoff function, $E(r|s)$, is continuous in s at $s = p$ for all r, then p corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the players in the limiting large population. Conversely, if p corresponds to a strict, symmetric Nash equilibrium in that game and there exists a neighborhood of $s = p$ throughout which $E(r|s)$ is continuous in s for all r, then p is a finite-population ESS in any sufficiently large finite population.*[20]

---

[20] A more general version of Theorem 2 is stated without proof in Crawford (1990b).

PROOF. Recall that, under my assumptions, the same payoff function can be used to describe the relationship between a player's payoff and players' strategies in large and finite populations, and that in either case, $E(r|s)$ gives the payoff of r when the frequencies of the other players' strategies are given by s. If p satisfies the conditions in the first part of the theorem, then letting $N \to \infty$ in (2) and using the continuity of $E(r|s)$ at $s = p$ for all r yields (4) for all $q \neq p$. It follows that p corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the players in the population.

Conversely, if p corresponds to a strict, symmetric Nash equilibrium, so that (4) holds with strict inequality for all $q \neq p$, then the continuity of $E(r|s)$ in s at $s = p$ for all r implies that, for each $q \neq p$, there exists an $\underline{N}(q)$ such that (2) holds whenever $N > \underline{N}(q)$. If $E(r|s)$ is also continuous in s throughout a neighborhood of $s = p$, then, by the argument of Crawford (1990a) (see footnote 14), there also exists an $\underline{N}$, independent of q, such that (2) holds for all $q \neq p$ whenever $N > \underline{N}$. Thus, p is an ESS in any sufficiently large finite population. ∎

The proof of Theorem 2 makes clear why the beggar-thy-neighbor aspect of selection need not "bias" a large-population ESS away from Nash equilibrium: As Schaffer (1988) showed, because $q/(N - 1) \to 0$ as $N \to \infty$, the effect of low-frequency mutations on the payoffs of ESS-players becomes negligible in large populations, at least when the payoff function is continuous in the population strategy frequencies. As my evolutionary analysis of the simultaneous-play minimum game makes clear, however, in games with discontinuous payoff functions beggar-thy-neighbor selection can occur even in large populations.

The discrimination between strict Nash equilibria that results in that game appears to contradict the common intuition (see, for example, Fudenberg and Kreps, 1988) that any strict equilibrium must be locally stable for sensible adaptive dynamics. However, small perturbations in the population pure-strategy frequencies are quite different from the small perturbations in the population mixed-strategy profile contemplated in the common intuition.[21] As Theorems 1 and 2 show (recalling that their conclusions would not be altered if individual players were allowed to play mixed strategies), this difference does not destroy the correspondence between Nash equilibrium and evolutionary stability for games with continuous payoff functions. But, although the strong payoff interactions that show up as discontinuities in the population strategy frequencies are not pathological, they can make the implications of evolutionary

---

[21] The payoff implications of each player in the simultaneous-play Stag Hunt deciding, by independent randomization, whether to reduce his effort from 2 to 1 are very different from the implications of a subset of the players with the same expected frequency, however small, doing so with probability one.

stability very different from those of the more familiar notion of local stability.

Theorems 1 and 2 clarify the correspondence between Nash equilibrium and evolutionary stability in a wide class of economically interesting environments. They also help to identify environments in which evolutionary stability discriminates between strict Nash equilibria, and to predict the pattern of discrimination. In the simultaneous-play minimum game of treatments A and A', for instance, the payoff function is clearly continuous whenever the population frequency of the lowest effort is positive, hence throughout a neighborhood of the strategy configuration at which all players choose that effort. But raising the frequency of the lowest effort above zero causes a discontinuous drop in output, and the payoff function violates the continuity condition of Theorems 1 and 2 whenever the frequency of the lowest effort is zero, hence at each of the six symmetric strategy configurations with higher efforts. Although these configurations correspond to strict, symmetric Nash equilibria in any finite population, however large, none of the higher efforts is an ESS. In VHBB's median games, by contrast, there are similar discontinuities, but they do not occur at the configurations that correspond to those games' strict, symmetric Nash equilibria, and each of those equilibria corresponds to an ESS.

4. AN "EVOLUTIONARY" INTERPRETATION

Although Section 3's analysis suggests that evolutionary game theory can be of considerable help in understanding how strategic uncertainty affects behavior, it is important not to rest content with a mechanical application of the ESS definition. Although strategic behavior is usually less sophisticated than traditional game theory assumes, it is plainly more complex than the simple characterization that underlies the evolutionary dynamics. Human players often draw subtle inferences from their observations and anticipate the effects of their strategy choices, taking the structure of the environment into account. Further, there is no guarantee that evolutionary stability's success in capturing the effects of strategic uncertainty in VHBB's experimental environments will extend to other settings in which coordination is difficult. (Modifying the rules of the large-group Stag Hunt to allow a small fraction of the population to choose effort 1 without destroying the payoff advantage of effort 2, for instance, makes the efficient strategy configuration in which all players choose effort 2 evolutionarily stable; but few analysts would expect subjects' behavior to be significantly altered by this change.[22]) Finally, there

---

[22] The fragility of the efficient strategy configuration is reflected in the smallness of its "basin of attraction," but this is unlikely to fully capture the effects of strategic uncertainty. Smoothing this example makes it clear that payoff discontinuities are not necessary for discrimination between strict Nash equilibria to occur in practice.

are strong regularities in the dynamics of players' strategy choices that are not explained by evolutionary stability; and the final outcome often depends on players' initial choices, about which the evolutionary analysis is completely silent.

These considerations suggest that an adaptive model, combining the advantages of the evolutionary framework with a more sophisticated view of players' behavior and an explicit account of their uncertainty about each others' choices, would yield a better explanation of VHBB's experimental results. This section outlines such a model. The interpretation of VHBB's results it suggests is close to theirs in most respects. My goals here are to highlight the issues that must be confronted in constructing an explanation and to show that an "evolutionary" analysis, based on simple, plausible hypotheses about subjects' behavior, can provide one.

The major hurdles the explanation must surmount are explaining the differences between VHBB's results for large- and small-group minimum games (sharp discrimination between Nash equilibria in large groups but not in small groups) and between their results for large-group minimum games and median games (no history-dependence in the former but very strong history-dependence in the latter). It is the absence of history-dependence in the large-group minimum game and the concomitant discrimination between equilibria that are most difficult to explain. Why would an intelligent, well-informed player ever move away from the efficient Nash equilibrium (or any strict equilibrium) in that game? The evolutionary analysis suggests why plausible adaptive dynamics might keep such a process going once it began: Although a mutation to a lower effort reduces the payoff of the mutant, it reduces the payoffs of players who choose higher efforts even more. Thus, once such mutants enter the population, they take it over in beggar-thy-neighbor fashion. But VHBB's subjects' strategy choices were voluntary, and the first step in this evolutionary process appears to depend crucially on the involuntary character of mutations.

The key to this puzzle lies in recognizing that, although players' strategy choices are voluntary, the beliefs that underlie them are involuntary. It is shown below that in environments with significant strategic uncertainty, the differences between players' beliefs can drive the adaptive dynamics like the mutations that drive the evolutionary dynamics.

The adaptive analysis follows Section 3's evolutionary analysis in most respects. Players are again restricted to pure stage-game strategies. Each player's strategy choices are governed by his beliefs about the group minimum or median (or, more precisely, how it is likely to depend on his choice), the only payoff-relevant aspect of other players' choices. In median games and large-group (but not small-group) minimum games, it is assumed that each player treats the probability that his effort influences

the group median or minimum as negligible.[23] In each stage, each player adjusts his strategy in the direction of increased payoffs, given his beliefs.[24] These adjustments determine a new group minimum or median. Players then observe it, update their beliefs, and the process continues.

Strategic uncertainty can affect the outcome of this process in two ways: through its effect on players' beliefs in initial treatment stages, and through its effect on how they interpret their shared experience in subsequent stages. The latter effect can be thought of as due to differences in how players model the stochastic process that generates the group minimum or median. Although any reasonable model must predict the minimum or median correctly if it remains constant long enough, strategic uncertainty may lead to differences between players' predictions of the minimum or median before it converges, even if their initial beliefs are the same and they always observe the same history.

I defer, for the moment, discussion of how players' initial beliefs are determined, and focus on the dynamics. A simple observation clarifies how strategic uncertainty can affect the outcome. Suppose that there are no differences in the models players use to interpret their experience, so that differences in their initial beliefs are the only possible source of differences in their beliefs in subsequent stages.[25] Then, under plausible assumptions, players never adjust their strategy choices beyond the current group minimum or median. The minimum or median therefore remains constant as they adjust, because it is an order statistic and there is no overshooting. It follows that as players' experience accumulates, their beliefs and efforts converge to the initial minimum or median, so that the final outcome necessarily coincides with it. The dynamics are perfectly history-dependent, and discrimination between equilibria occurs only through players' initial choices.

This perfect history-dependence corresponds closely to the results of the median experiments. However, it is inconsistent with the results of the minimum experiments, in which history influenced the outcome

---

[23] In the large-group minimum experiments of treatments A and A', subjects were alone at the group minimum in only 10 out of 70 stages and 0 out of 32 stages, respectively. An individual subject's effort therefore determined the group minimum only rarely. Similarly, an individual subject's effort rarely influenced the group median in the median experiments. A player would, of course, expect to have a nonnegligible influence on the group minimum if he chose an effort below what he expected the minimum of the others' efforts to be, but he has no incentive to do this.

[24] VHBB's subjects' adjustments were on average approximately proportional to the distance between their current efforts and those that would have maximized their payoffs, and usually partial, in that they stopped short of the maximizing efforts.

[25] Recall that each player receives the same information, except that only he observes his own effort; except in two-person groups, this difference is negligible, because the probability that his effort influences the minimum or median is negligible.

weakly in small groups and not at all in large groups, and overshooting was not uncommon. To capture this difference, I assume that players' beliefs are subject to idiosyncratic random shocks, which reflect the differences in how they interpret their experience. Because these shocks represent an unpredictable component of players' beliefs, it is natural to assume that they are distributed independently across players. Their influence on players' adjustments can be expected to die out over time. because as players' shared experience accumulates, their beliefs become more precise and they are more likely to choose the common value of effort their experience suggests is optimal. Before their influence dies out, however, these shocks can significantly alter the course of players' adjustments, and thereby the final outcome. This influence varies across treatments in a way that allows a unified explanation of VHBB's results.

For simplicity, I describe each player's beliefs in terms of its implications for his effort choice, following Woodford (1990); provided that the adjustment process reflects the increasing precision of beliefs, this involves no significant loss of generality. In the large-group minimum games, I assume that a player's next-stage effort choice is most likely to equal the current minimum, but lies above or below it with small probabilities (determined by the distribution of his idiosyncratic shock), which decline over time as experience accumulates and his beliefs become more precise. Behavior is similar in small-group minimum games, except that a player whose effort is at the minimum then has a higher probability of increasing his effort, because he knows that there is a good chance that it alone determined the minimum. Thus, the average player whose effort is above the minimum reduces it; and the average player whose effort is at the minimum increases it in small groups, and leaves it approximately unchanged in large groups.

In the median games, I assume, similarly, that a player's next-stage effort choice is most likely to equal the current median, but lies above or below it with small probabilities, which again decline over time as his beliefs become more precise.

The model could now be closed by specifying how the probability distributions of players' idiosyncratic shocks depend on the history of the system. However, its implications are largely independent of the details of this specification. I therefore proceed by showing in general terms why it implies behavior like that VHBB observed.

For the group minimum to go up in the large-group minimum game of treatments A and A', all players at the minimum must increase their efforts, and all players above it must reduce theirs by less than the distance to the minimum. The former possibility is highly unlikely without significant idiosyncratic shocks to their beliefs, and in a large group the latter possibility is highly unlikely with them. Because the shocks are

independent across players, the probability that the minimum in a large-group minimum game ever goes up is therefore near zero with or without shocks. In fact the minimum went up (by one unit) in only one out of 102 stages in VHBB's experiments with treatments A and A'.

As noted above, once players reach a stage after which the effects of the shocks to their beliefs are negligible, their efforts converge to the Nash equilibrium determined by the minimum in that stage. Until such a stage is reached, however, the group minimum has a strong tendency to fall. Although the average player whose effort is above the current minimum does not reduce his effort below it, there is a positive probability that any such player's idiosyncratic shock will lead him to do so. Suppose for the sake of argument that there are L players with efforts above the minimum, that this probability is z for each of them, and that players at the minimum never reduce their efforts below it. (This view of players' adjustments is reinforced by VHBB's observation that a given subject rarely remained at his group minimum in successive stages.) Because players' shocks are independent, the probability that the minimum is lower in the next stage is then $1 - (1 - z)^L$. This probability converges to zero with z, as it must for the common intuition about local stability to be valid in the absence of significant strategic uncertainty. The rate of convergence is very slow, however, unless L is very small. If, for instance, $L = 14$ (a typical value in the early stages of treatment A), then $1 - (1 - z)^L = 0.13$ if $z = 0.01$, 0.25 if $z = 0.02$, 0.51 if $z = 0.05$, 0.69 if $z = 0.08$, and 0.77 if $z = 0.10$. In fact, in the experiments with treatment A the minimum effort went down in 9 out of the 13 stages (69%) in which it was not already at the lowest possible level.

The outcome in a large-group minimum game is therefore determined by how quickly players' learning from their experience reduces their strategic uncertainty, relative to the rate at which that uncertainty makes the group minimum fall. It is theoretically possible for players to learn so quickly that they converge to a Nash equilibrium with efforts above the lowest level; this possibility can even have positive prior probability under my assumptions. However, VHBB's subjects' predictions for the first stage of treatment A and their subsequent behavior indicate that the large-group minimum game initially generated a great deal of strategic uncertainty, much of which persisted until the minimum reached the lowest effort.[26] The local stability of this game's strict Nash equilibria therefore had little relevance for the probability distribution of final outcomes.

The analysis is essentially the same for treatment A', which restored the payoffs of treatment A following those of treatment B. Subjects' pre-

---

[26] It is clear that this result would extend to some games without the extreme form of payoff interactions that lead to discontinuities in the large-group minimum game.

vious experience with this game appeared to reduce their initial level of strategic uncertainty below that observed for treatment A. However, this reduction consisted mainly of making their beliefs closer to the final outcome of treatment A. This kind of reduction tends to make the dynamics converge more quickly to that outcome, as occurred in VHBB's results for this treatment.

The analysis of treatment B (in which the highest effort was a weakly dominant strategy) is even simpler. Because players' shocks make them choose different efforts with positive probability, higher efforts generally yield higher expected payoffs in treatment B; players therefore tend to adjust their efforts upward until they reach the highest effort, as in VHBB's experiments.

The adaptive dynamics have quite different implications in treatment C_d. Recall that players in this treatment learned only their current pair's minimum after each stage. As in treatments A and A', the average player whose effort is above his current pair's minimum reduces his effort and the average player whose effort is at his current pair's minimum increases his effort. These upward and downward adjustments should be roughly equal in magnitude. (VHBB found, in fact, that the average subject at the minimum in treatment C_d increased his effort by 1.1, and the average subject above the minimum reduced his effort by 1.0.) Because there are roughly equal numbers of players in the population who are above and below their pairs' minima, players' efforts can be expected to drift over time, without strong trends, until experience eliminates their uncertainty.

This explanation of the difference between VHBB's results for small- and large-group minimum games requires that each player make sophisticated use of the information he receives, taking into account the fact that he is a nonnegligible fraction of the pairs he forms. Otherwise, a player whose effort is at his current pair's minimum would have no strong tendency to increase his effort. Players' idiosyncratic shocks would then tend to make their efforts drift downward over time, on average, perhaps converging to the lowest effort. Subjects' efforts in treatment C_d had, if anything, an upward trend, and there is certainly no evidence that they were converging to the lowest effort.

In median games, the adaptive dynamics tend to make players' efforts move toward the current group median. The idiosyncratic shocks to players' beliefs do not induce a significant trend in the median, because of its symmetry. The dynamics are therefore most likely to converge to the Nash equilibrium at which players' efforts equal the initial group median, although they may converge, with positive probability, to equilibria with other common efforts. These conclusions correspond closely to VHBB's results: In the 27 treatments of the median experiments (12 Γ, 3 Γdm, 9 Ω, and 3 Φ), totaling 195 stages, the median never changed during a treat-

TABLE I

		Minimum treatment				
		A (%)	B (%)	A' (%)	C_d (%)	C_f (%)
Subject's	7	33 (31)	76 (84)	23 (25)	11 (37)	13 (42)
initial	6	10 (9)	1 (1)	1 (1)	1 (3)	0 (0)
effort	5	34 (32)	2 (2)	2 (2)	2 (7)	6 (19)
	4	18 (17)	5 (5)	7 (8)	5 (17)	2 (6)
	3	5 (5)	1 (1)	7 (8)	3 (10)	1 (3)
	2	5 (5)	1 (1)	17 (19)	1 (3)	1 (3)
	1	2 (2)	5 (5)	34 (37)	7 (23)	8 (26)
Totals		107 (101)	91 (99)	91 (100)	30 (100)	31 (99)

		Median treatment		
		Γ, Γdm (%)	Ω (%)	Φ (%)
Subject's	7	8 (15)	14 (52)	2 (7)
initial	6	4 (7)	1 (4)	3 (11)
effort	5	15 (28)	9 (33)	9 (33)
	4	19 (35)	3 (11)	11 (41)
	3	8 (15)	0 (0)	2 (7)
	2	0 (0)	0 (0)	0 (0)
	1	0 (0)	0 (0)	0 (0)
Totals		54 (100)	27 (100)	27 (99)

ment, and, with three minor exceptions, each subject's effort converged to the Nash equilibrium determined by the initial treatment median.

The history-dependence in the results for the median treatments (and, to a lesser extent, in minimum treatment C_d) makes a full explanation of VHBB's results depend on understanding their subjects' choices in initial treatment stages. There were strong regularities in subjects' initial choices throughout the experiments. In the median treatments, for instance, no subject ever began with an effort below $e_i = 3$. These regularities are not explained by evolutionary stability or by traditional equilibrium refinements. I now consider whether they can be understood as sensible responses to strategic uncertainty.

VHBB's subjects' initial effort choices can be summarized in Table I (the median table, which is adapted from Table II in VHBB (1991), includes only the first treatments in each sequence).

As far as I am aware, the only systematic attempt to make precise predictions of behavior in coordination games is Harsanyi and Selten's (1988) (henceforth "HS") general theory of equilibrium selection. HS's theory follows the traditional approach of assuming that players' beliefs about each other's strategy choices always converge to a particular Nash

equilibrium before play begins; but the mental tâtonnements they use to model the convergence process are sensitive to strategic uncertainty. Their predictions are therefore a useful starting point for examining VHBB's results for initial treatment stages.

I again assume that players focus on their choices of stage-game strategies, ignoring any possible influence of their choices on future developments. As noted in footnote 6, HS's theory gives priority to payoff-dominance, and therefore predicts the $e_i = 7$ equilibrium in the initial stages of all treatments except treatment Φ. In treatment Φ, all pure-strategy equilibria yield the same payoffs and HS's theory therefore predicts the $e_i = 4$ equilibrium in response to the symmetries of the game. These predictions correspond to the modal first-stage choices in treatments B, C_f, C_d, Ω, and Φ; run a close second in treatment A and a more distant second in treatment A'; and tie for a distant third in treatments Γ and Γdm. (The correspondence for treatment C_f should be discounted because it may be due in part to players' attempts to influence each other's future choices.) Thus, HS's solution theory does fairly well overall in predicting subjects' modal initial responses. However, their assumption that players can perfectly coordinate their beliefs before play begins leads to severe underestimates of the amount of strategic uncertainty, and their theory predicts the first-stage group minimum or median effort incorrectly in most treatments.[27]

Subjects' uncertainty often seemed to take the form of doubts about which of two (or more) equilibrium selection principles would govern other subjects' choices. (See Schelling (1960) and Roth (1985) for evidence of similar doubts in other settings.) In treatment A, for instance, subjects' initial choices were bimodal, with peaks at efforts $e_i = 7$ and $e_i = 5$: the "obvious" efficient equilibrium and a conservative, but still optimistic, response to uncertainty about whether it would be obvious to

----

[27] The resulting first-stage inefficiency suggests that HS's solution theory might predict better if it did not give priority to payoff-dominance. (HS's arguments in support of payoff-dominance are almost completely separable from their arguments in support of the rest of their theory; and it is the latter arguments that embody most of their ideas about modeling players' responses to strategic uncertainty.) HS's theory without payoff-dominance would make the same predictions in treatments B (because HS also give priority to the elimination of weakly dominated strategies by working with the limits of "uniformly perturbed" games); in treatments Γ, Γdm, and Ω (because risk-dominance also favors higher efforts in these games); and in treatment Φ (because of the symmetries in this game). HS's theory would predict the equilibrium with common effort $e_i = 1$ in treatments A and A' (because risk-dominance favors lower efforts in large-group Stag Hunt games) and the equilibrium with common effort $e_i = 4$ in treatment C_d with two players (because risk-dominance is neutral in two-person Stag Hunt games, and HS's theory therefore applies the tracing procedure to a uniform prior over the undominated strategies). Thus, without payoff-dominance, HS's theory predicts subjects' first-stage behavior worse in treatments A and C_d, and better in treatment A'.

everyone. In treatments A' and C, subjects' initial choices also seemed to be influenced by doubts about the extent to which other subjects would treat the outcomes of the preceding A and B treatments as relevant precedents. Interestingly, subjects seemed more optimistic at the start of treatment C than at the start of the immediately preceding A' treatment, although their negative experience in the A' treatment might have been expected to make them less optimistic.

The distributions of initial choices in median treatments Γ, Γdm, and Ω were similar, in this respect, to those of minimum treatment A. The simple payoff table of treatment Ω (which had zeros everywhere off the main diagonal) seemed to give subjects somewhat more confidence in the efficient equilibrium. The symmetries of the game used in treatment Φ and the fact that it does not have a uniquely efficient equilibrium led to qualitatively different responses. Some subjects responded to the symmetries as HS's theory predicts, by choosing $e_i = 4$, but most "shaded" their choices upward, perhaps due to a "top-left" focal point effect or a fear that other players would be influenced by such an effect. (This effect makes the inefficiencies observed in the other treatments more striking, because the efficient Nash equilibrium occupied the top-left position in their payoff tables.)

The sequencing of treatments in the median experiments revealed another interesting regularity, which may prove important in other environments. In the median experiments, the group median never fell from one treatment to the next, even when the sequencing of treatments was reversed. With treatment sequence $\{\Gamma, \Omega, \Gamma\}$, the observed sequences of (initial and final) treatment medians were $\{4, 7, 7\}$, $\{5, 7, 7\}$, and $\{5, 5, 5\}$. Treatment sequence $\{\Gamma$dm$, \Omega\}$ yielded median sequences $\{4, 7\}$, $\{4, 7\}$, and $\{5, 5\}$; and treatment sequence $\{\Omega, \Gamma\}$ yielded median sequences $\{7, 7\}$, $\{5, 7\}$, and $\{7, 7\}$. Finally, treatment sequence $\{\Phi, \Gamma\}$ yielded median sequences $\{4, 6\}$, $\{5, 7\}$, and $\{5, 7\}$. (The increases in this last sequence should be discounted, because the $e_i = 7$ equilibrium was not uniquely efficient in treatment Φ.) Thus, when one treatment was replaced by another, subjects did not just continue playing the Nash equilibrium to which they had converged in the "old" game (which was always an equilibrium in the "new" game). Instead they experimented, treating the old equilibrium as a benchmark, from which improvements could safely be sought in the new game, in the "sensible" direction. Changing treatments seemed to enable subjects to coordinate the timing of their experiments, increasing the chance of success.

REFERENCES

BANKS, J., PLOTT, C., AND PORTER, D. (1988). "An Experimental Analysis of Unanimity in Public Goods Provision Mechanisms," *Rev. Econ. Stud.* **55**, 301–322.

BRYANT, J. (1983). "A Simple Rational Expectations Keynes-Type Model," *Quart. J. Econ.* **98**, 525–528.

COOPER, R., AND JOHN, A. (1988). "Coordinating Coordination Failures in Keynesian Models," *Quart. J. Econ.* **103**, 441–463.

CRAWFORD, V. P. (1989). "Learning and Mixed-Strategy Equilibria in Evolutionary Games," *J. Theoret. Biol.* **140**, 537–550.

CRAWFORD, V. P. (1990a). "On the Definition of an Evolutionarily Stable Strategy in the 'Playing the Field' Model," *J. Theoret. Biol.* **143**, 269–273.

CRAWFORD, V. P. (1990b). "Nash Equilibrium and Evolutionary Stability in Large- and Finite-Population 'Playing the Field' Models," *J. Theoret. Biol.* **145**, 83–94.

CRAWFORD, V. P., AND HALLER, H. (1990). "Learning How to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica* **58**, 571–595.

FUDENBERG, D., AND KREPS, D. (1988). *A Theory of Learning, Experimentation, and Equilibrium in Games*, manuscript, Stanford University.

HARSANYI, J., AND SELTEN, R. (1988). *A General Theory of Equilibrium Selection in Games.* Cambridge, MA: MIT Press.

HINES, W. G. S. (1980a). "Three Characterizations of Population Strategy Stability," *J. Appl. Probab.* **17**, 333–340.

HINES, W. G. S. (1980b). "Strategy Stability in Complex Populations," *J. Appl. Probab.* **17**, 600–610.

ISAAC, R. M., SCHMIDTZ, D., AND WALKER, J. (1989). "The Assurance Problem in a Laboratory Market," *Public Choice* **62**, 217–236.

KEYNES, J. M. (1936). *The General Theory of Employment, Interest, and Money.* London: Macmillan.

KIM, Y.-G. (1990). "Evolutionary Analysis of Two-Person Finitely Repeated Coordination Games," manuscript, University of California, San Diego.

KOHLBERG, E., AND MERTENS, J.-F. (1986). "On the Strategic Stability of Equilibria," *Econometrica* **54**, 1003–1037.

MAYNARD SMITH, J. (1974). "The Theory of Games and the Evolution of Animal Conflicts," *J. Theoret. Biol.* **47**, 209–221.

MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games.* Cambridge, UK: Cambridge Univ. Press.

MAYNARD SMITH, J. (1988). "Can a Mixed Strategy Be Stable in a Finite Population?" *J. Theoret. Biol.* **130**, 247–251.

ROTH, A. (1985). "Toward a Focal-Point Theory of Bargaining," in *Game-Theoretic Models of Bargaining* (A. Roth, Ed.), pp. 259–268. New York: Cambridge Univ. Press.

ROUSSEAU, J.-J. (1973). "A Discourse on the Origin of Inequality," in *The Social Contract and Discourses* (translated by G. D. H. Cole), pp. 27–113. London: Dent.

SCHAFFER, M. (1988). "Evolutionarily Stable Strategies for a Finite Population and a Variable Contest Size," *J. Theoret. Biol.* **132**, 469–478.

SCHELLING, T. (1960). *The Strategy of Conflict.* Cambridge, MA: Harvard Univ. Press.

SCHELLING, T. (1973). "Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities," *J. Conflict Resolution* **17**, 381–428.

SCHELLING, T. (1978). *Micromotives and Macrobehavior.* New York: Norton.

SELTEN, R. (1983). "Evolutionary Stability in Extensive Two-Person Games," *Math. Soc. Sci.* **5**, 269–363.

SELTEN, R., AND STOECKER, R. (1986). "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach," *J. Econ. Behav. Organ.* **7**, 47–70.

TAYLOR, P., AND JONKER, L. (1978). "Evolutionarily Stable Strategies and Game Dynamics," *Math. Biosci.* **40,** 145–156.

VAN HUYCK, J., BATTALIO, R., AND BEIL, R. (1989). "Asset Markets as an Equilibrium Selection Mechanism: Coordination, Game Form Auctions, and Forward Induction," Working Paper 88–40, Texas A&M University, revised 1989.

VAN HUYCK, J., BATTALIO, R., AND BEIL, R. (1990). "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *Amer. Econ. Rev.* **80,** 234–248.

VAN HUYCK, J., BATTALIO, R., AND BEIL, R. (1991). "Strategic Uncertainty, Equilibrium Selection Principles, and Coordination Failure in Average Opinion Games," *Quart. J. Econ.* **106,** in press.

WOODFORD, M. (1990). "Learning to Believe in Sunspots," *Econometrica* **58,** 277–307.

ZEEMAN, E. C. (1979). "Population Dynamics from Game Theory," in *Global Theory of Dynamical Systems* (Z. Nitecki and C. Robinson, Eds.), Lecture Notes in Mathematics, Vol. 819. Berlin/New York: Springer-Verlag.