Coordination and Shared Mental Models

Diana Richards University of Minnesota

Preferences may be structured by social constraints, by institutional procedures, or, as in the focus of this article, by knowledge representations. This article explores the prospects for successful coordination when players have conflicting preferences but have similar cognitive representations of the decision context. A "knowledge-induced equilibrium" is a stable outcome reached under players' mutual understandings of the empirical context. The purpose of this article is to develop a formal framework that combines strategic rationality with social or cognitive components of knowledge.

Theories of coordination are concerned with how individuals can coordinate their conflicting preferences in the absence of credible communication. Coordination problems abound in politics. For example, voting for a third-party candidate is a coordination problem with other strategic voters (Cox 1997; also Myerson and Weber 1993). The decision to join a risky mass protest (Chong 1991; Lohmann 1994) or to contribute to a public good (Taylor 1987; Ostrom 1990) or to protest a government by withholding tax payments or resisting a military draft (Levi 1988, 1997) are all coordination problems. The establishment of stable institutional arrangements such as norms, conventions, contracts, or principal-agent relationships are also coordination problems (e.g., Axelrod 1986; Spruyt 1994; Young 1998), as are tacit agreements such as restraint in warfare (e.g., Legro 1995) and formal negotiated settlements in conflicts (e.g., Schelling 1960).

However, despite the empirical prevalence of coordination problems, we have failed to achieve a full theoretical account of coordination. The theoretical cul-de-sac arises because coordination problems by definition have multiple equilibria, resulting in indeterminate predictions. The most prevalent solution is Schelling's (1960) idea of a *focal point.* Schelling surmised that coordination could occur if there was some shared interpretation of the salient features of a decision context. Forty years later, the presence of a focal point remains the most frequently invoked concept to explain coordination in politics. However, the concept has largely remained extra-theoretical in that it is seldom formally defined and is typically invoked as a post-hoc explanation of an observed empirical outcome.[1]

There are several different angles from which to develop a more rigorous theory of coordination, most of which include some form of intersubjective understanding among players. One promising solution focuses on the emergence of *precedents and conventions* over time (e.g., Crawford and Haller 1990; Young 1998). However, this solution is most

---

Diana Richards is Associate Professor of Political Science, University of Minnesota, 267 19th Avenue South, Minneapolis, MN 55455 (richards@polisci.umn.edu).

Supported by NSF grant SBR-9729847. The author is grateful to workshop participants at the University of Minnesota, Hoover Institution, Caltech, the Santa Fe Institute workshop on Institutions: Complexity and Difficulty, and the Complex Systems Modeling Team at Los Alamos National Laboratory. The Mathematica program was provided by Whitman Richards. Computing time provided by the Minnesota Supercomputing Institute.

[1]Experimental studies are an exception to the post-hoc use of a focal point (e.g., Mehta, Starmer, and Sugden 1994; Bacharach and Bernasconi 1997).

American Journal of Political Science, Vol. 45, No. 2, April 2001, Pp. 259–276

applicable to coordination settings where repeated interaction can institutionalize coordination solutions over time, as in long-term economic or social institutions. However, although the folk theorem points out that repeated play can induce cooperation, repeated play can also exacerbate the number of equilibria. Furthermore, coordination in politics often lacks the long-term dynamics that are necessary for evolutionary analysis, as in one-time negotiated settlements to resolve an international disagreement, or voters' coordination on third-party support in a single election, or the decision to walk the town on the Monday evening that became the Leipzig demonstrations in East Germany.

Another approach to coordination looks to the *salience* of particular outcomes resulting from the players' choice of a frame (e.g., Schelling 1960; Sugden 1995; Bacharach and Bernasconi 1997). In this approach, players' strategies are broadened to include not only the coordination act, but also the choice of a payoff-independent labelling scheme that partitions the alternatives into subsets such that one subset is smaller (hence "rarer" and more "salient") by virtue of its feature classification. In these models, unlike traditional game-theoretic models, the labels of the choices matter. For example, Sugden (1995, 549) suggests a labelling scheme based on the extent to which alternatives have been empirically mentioned in the past. Applied to Cox's (1997) multiparty coordination problem, Sugden's model suggests that voters should independently label the small parties in terms of the frequency with which they have been mentioned (such as in the media) and coordinate on the most salient third party under this labelling scheme. However, by treating the labelling of alternatives as independent of players' payoffs, this solution neglects the voters' preferences over parties.

Another approach to coordination focuses on the role of *culture or ideas* as resolving the indeterminacy from multiple equilibria (e.g., Kreps 1991; Ferejohn 1991; Weingast 1995; Schiemann 2000). For example, Kreps (1991) proposes that the presence of a "corporate culture" mediates between actions and outcomes in economic games. Weingast (1995) uses Converse's (1964) idea of a shared system of beliefs to model how a shared understanding of sovereignty maintains international cooperation by removing ambiguity due to differing interpretations about others' actions. Schiemann (2000) discusses the promising gains to be had from merging strategic rationality with intersubjective knowledge, but stops short of providing a formal framework of such a union. Thus, the role of culture or ideas in coordination remains largely at the conceptual stage rather than undertaking the task of developing a formal model of how shared beliefs intersect with players' coordination decisions.

This article develops a formal model of coordination which focuses on the information provided by the participants' mental models. When players have similar mental models of a choice setting, the choices and the relationships between choices have a common underlying structure. This article adapts an equilibrium concept originally developed for social choice (Richards, McKay, and Richards 1998). As in the approaches outlined above, the empirical properties of the alternatives and the intersubjective understandings among players remain important. However, rather than an external labelling scheme that is independent of payoffs, my focus is on the players' internal cognitive representations of the strategy set, which are assumed to be closely coupled with players' preferences. The contribution of this article is to bring together several existing ideas: mental representations from cognitive science, maximum-likelihood from statistics, and coordination games from noncooperative game theory, to develop the concept of a *knowledge-induced equilibrium.*

Shared Mental Models

When the assumption of rationality is relaxed, it is typically to emphasize the limits of humans' abilities to comprehend a complex environment and their need to rely on the use of lower-level algorithmic routines such as myopic searches, satisficing, or mimicking. Clearly there is evidence that these shortcut routines are used by decision makers. However, another way that humans cope with a complex empirical environment is to rely on their powerful mental modeling abilities. These two approaches are both consistent with a rational framework (e.g., Kollman, Miller, and Page 1992; Denzau and North 1994) and can coexist: the former focuses on the implementation of decisions and the latter emphasizes the representation of a decision context. In this article, the emphasis is on the effect of cognitive structures. I refer to the cognitive organization of an empirical domain as a *mental model* and a *knowledge structure* as the representation of a mental model. An organization of knowledge is a *structure* in that it mediates between individuals and their world—much as social constraints or political institutions are also structures (e.g., Converse 1964; Shepsle 1979).

Mental models have diverse organizational form and content. The mental landscape of political parties is a form of a mental model (Poole and Rosenthal 1991;

Hinich and Munger 1994). A mental model may focus on categories and features, as in Clausen's (1973) account of congressional politics. Mental models may also take the form of cause-and-effect models of how the world works and relate to beliefs of what action is appropriate, as in Chamberlin's and Churchill's different causal models and interpretations during World War II. Narratives and stories and plots, like other forms of linguistic communication, are also mental models in that they form a known intrinsic structure in order for the meaning to be understood by the audience. Schemas and analogies are mental models in that they are heuristic narratives that structure understandings of a class of events (e.g., Axelrod 1973; Khong 1992, 25). Mental models organize the empirical world and thus organize interpretations, communication, and behavior.

In this article, a mental model is modeled simplistically with two components: a set of categories and similarity relations among the categories. Specifically, the mental model is represented as a graph, where each node is a category and a link between two nodes indicates that the two categories are closely related in a player's mental organization. Categories that are not adjacent are more cognitively distinct in a player's mental organization. This graph is referred to as the knowledge structure, as it is a representation of players' mental models. It is important to emphasize that this is a "feature-based" rather than exclusively metric representation (Tversky 1977).

Example 1 *Organization of Political Parties*

One way to organize political parties is to place them on a left-right continuum. Traditionally this organization is modeled spatially (e.g., Downs 1957; Black 1958; Hinich and Munger 1994). A more general nonmetric cognitive organization would allow voters to have a mental map of the set of political parties, such as that the Green Party is similar to the Citizen Party and more similar to the Democratic Party than to the Republican Party. ■

Example 2 *Organization of Proposals under Negotiation*

Negotiation can take place over relatively trivial categories, such as how to spend the evening in the battle-of-the-sexes game, or over categories with profound historical impact, such as which set of institutional rules to implement as a framework for governing the United States, or which armistice agreement to abide by during World War I. Whatever the content of the negotiation, participants organize the set of proposals comparatively in a mental framework in order to understand their relative merits. For example, in the 1787 Convention, where collections of institutional rules such as the Nationalist Plan or the Federalist Plan were debated, much of the debate among the delegates centered on organizing the similarity and differences between the various plans. ■

Example 3 *A Knowledge Representation of Political Organizations*

Ideology itself is an organization of knowledge in that it summarizes relationships between political ideas. Using data from a larger experimental study, Figure 1 summarizes how a group of fifteen students collectively organized fourteen American political organizations.[2] Each student completed a survey on how he or she organized the categories both in terms of similarity and in terms of adjacency triples.[3] The data were analyzed to identify statistically significant pairs and triples of categories across all the subjects' responses. Figure 1 shows a simplified version of the results of these multidimensional scaling techniques. Each node is a category and each edge is a statistically significant similarity relation between those two categories based on the students' pooled data. For example, the radical environmental group Earthfirst! was placed adjacent to Greenpeace and PETA, but not directly adjacent to the moderate Sierra Club. The American Civil Liberties Union was placed as similar to organizations both on the traditional right, such as the National Rifle Association (advocating the right to bear arms) and the traditional left, such as the National Organization for the Reform of Marijuana Laws (advocating the legalization of marijuana).[4] ■

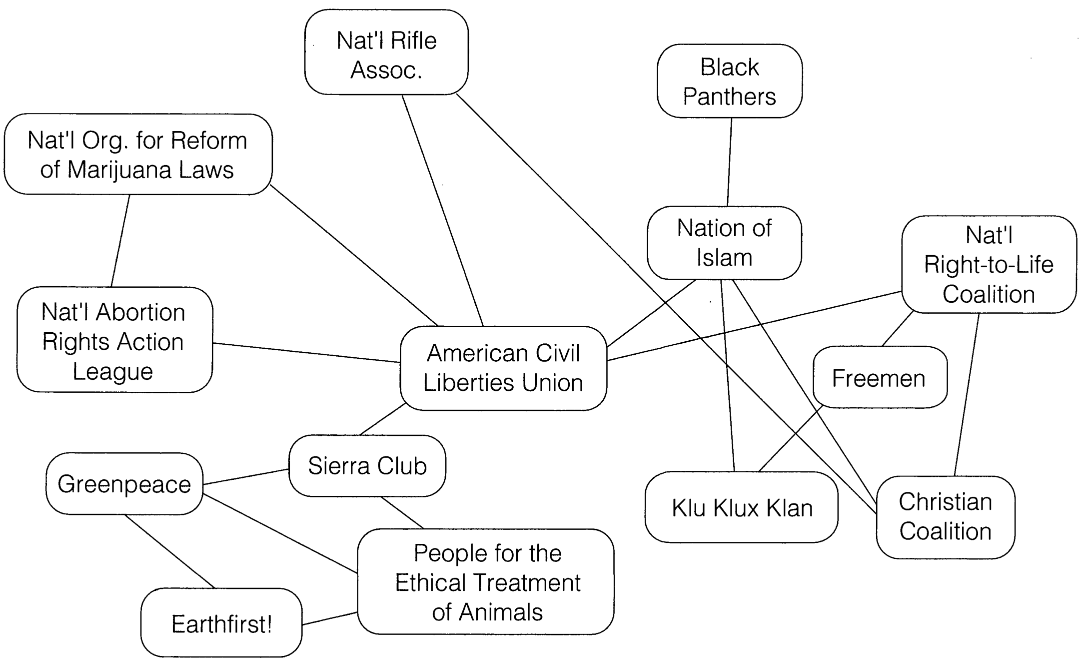
In the formal model presented below, I assume that mental models are shared across decision makers. The idea of shared knowledge is not new in political science. Philip Converse's work (1964) emphasized the role of a shared system of beliefs in politics. Thomas Schelling's (1960) focal point solution is basically an appeal to intersubjective or cultural understandings. Both recent constructivist approaches and studies of epistemic communities in international relations emphasize the importance of collective knowledge (e.g., Wendt 1999; Haas

[2] Since the data for this example is part of a larger experimental project, details on the experimental protocol are not given. However, standard experimental techniques were used, including the independent creation of the list of categories (based on a separate survey asking for salient contemporary political organizations), randomization of survey questions, anonymity of subjects' responses, and compensation for subjects' time.

[3] These triples are called "trajectories" in a mental representation. The procedure is described in Richards and Koenderink (1995).

[4] This example illustrates a knowledge structure as a graph of categories and similarity relationships but should be not be interpreted as a coordination example.

FIGURE 1 Graph Representation of Knowledge: Experimental Data on Political Organizations



All edges shown are significant at .01 level based on multinomial distribution of pooled data on adjacency triples from 15 subjects (n = 210). Lengths of edges are irrelevant (although the spatial layout of the categories is guided by multidimensional scaling on subjects' pooled responses on the pairwise similarity of the categories.

1990). Within the formal literature, the Condorcet Jury Theorem implicitly includes shared knowledge in the common value assumption of a shared probability of choosing correctly (e.g., Miller 1986; Austen-Smith and Banks 1996). Some evidence of shared knowledge from other fields includes the shared linguistic structure of grammar and phonetics, the common semantic structure of kinship terms (Romney et al. 1996), perceptual saliency in cognitive science where humans all pick the same key features when shown an empirical context (Ullman 1996), and shared reciprocity relations (Cosmides and Tooby 1992; Richards 2001).

However, the extent to which knowledge is shared is an important empirical question. Certainly many beliefs and ideas vary greatly across individuals due to differences in socioeconomic position, information access, culture, or experience (e.g., Wittkopf and Maggioto 1983; Conover and Feldman 1984). The extent to which mental models are shared, in terms of agreement over categories and relationships, is a potentially important independent variable. Political disagreement may stem from differences in preferences or from different conceptions of basic category relations. The variation in the extent to which models are shared may be manifested across issues (as in the disagreement over basic category relations in the debate over affirmative action) or across subgroups (such as between elites and masses [Converse 1964]). The purpose of this article, like that of Shepsle's (1979) insights regarding institutional structures, is to demonstrate that knowledge structures, when shared, are a source of stability in collective decision making.

Organizing Outcomes in Coordination Games: Two Examples

In *coordination* or *bargaining games,* players have a common interest in reaching some agreement but have different preferences over the terms of agreement and are often uncertain of other players' preferences. Two classic representations of coordination problems are the battle-of-the-sexes game (Luce and Raiffa 1985, 91; Banks and Calvert 1992) and Schelling's parachutist game (Schelling 1960, 58–59; Gauthier 1975; Sugden 1995). In this section I explore these two classic examples through the perspective of players placing an organizational structure on the choice context with shared mental models.

Battle-of-the-Sexes

In the traditional narrative of the battle-of-the-sexes game, two players must coordinate on one activity for the evening, such as between a prize fight and a ballet, where the players disagree over the ranking of the activities, but prefer to go to any activity with the other than to spend the evening alone. Thus, players face multiple equilibria and the danger that if they fail to coordinate at one of the equilibrium outcomes the result will be inferior. This game has been used to model bargaining and repeated Prisoners' Dilemma games (Schelling 1960; Hardin 1982; Taylor 1987). The traditional solution to the game is the mixed-strategy equilibrium, the only symmetric equilibrium of the game, where players randomize over their choices of the evening's activities. Recent extensions also consider the role of communication (e.g., Banks and Calvert 1992).

In this example I consider the contribution of players' mental models. Communication and repeated play is removed in this example to illustrate the role of knowledge structures. Elaborating on Luce and Raiffa's original story, suppose that a couple arranged to meet at "Cinema One-2-Many" to watch a film but they did not decide which film to watch. Upon arriving a few minutes late to the cinema, each player quickly scans the list of current showings: an adventure film, a comedy, a drama, a mystery, a suspense thriller, and a war movie. Each person has their own private preferences for the evening's film but prefers watching any movie with the other person to watching their top-ranked choice alone. Assume that neither person knows the other person's preferences for that evening's film. Which film should they choose?

Film genres, like other forms of narrative, are organized using mental models derived from an understanding of the attributes of the film categories. For example, categories of films differ in their mood, level of violence, and tension in the plot. Assume that prior to any coordination choice, players cognitively organize these outcomes in a mental map. This mental map allows each player to understand what it means to say that a film is a "suspense thriller" and informs that player's preference formation. I begin by assuming that the basic organization of the outcomes is shared, namely that although players may disagree over their rankings of the film genres, they both organize them in the same abstract cognitive arrangements. Figure 2 shows a shared organization of the film genres from a larger experimental study.[5]

[5] This graph is an excerpt from more detailed experimental tests conducted with Whitman Richards using multidimensional scaling techniques.

The formal model described in detail in the following section assumes that players' preferences over the coordination equilibria follow from their organization of the empirical context. In other words, the knowledge structure is assumed to contain information about feasible and consistent preferences. Let $\succ$ denote preference between two actions and $\sim$ denote indifference. If player 1 most prefers watching the drama, then her preferences over the remaining outcomes are assumed to follow from her organization of the activities: drama $\succ$ comedy $\sim$ adventure $\sim$ mystery $\succ$ suspense $\sim$ war. The assumptions of the model imply six preference orderings over outcomes consistent with the knowledge structure of Figure 2:

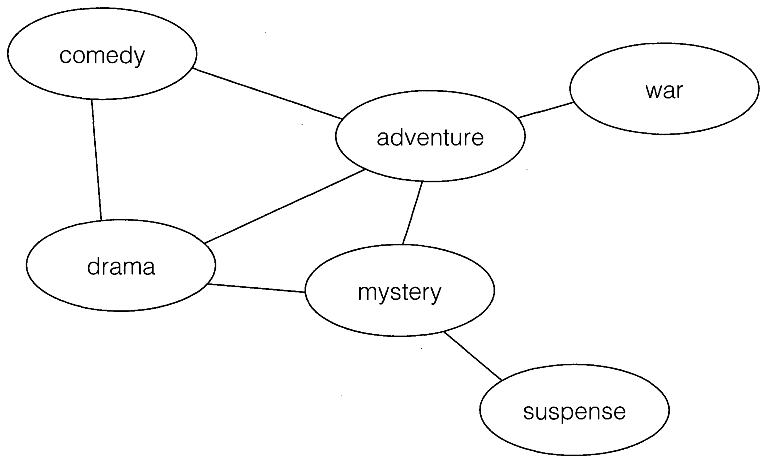
$$\begin{aligned}
\tau_a &\equiv \text{adventure} \succ \text{war} \sim \text{mystery} \sim \text{drama} \sim \text{comedy} \\
&\quad \succ \text{suspense}, \\
\tau_c &\equiv \text{comedy} \succ \text{adventure} \sim \text{drama} \succ \text{mystery} \sim \text{war} \\
&\quad \succ \text{suspense}, \\
\tau_d &\equiv \text{drama} \succ \text{comedy} \sim \text{adventure} \sim \text{mystery} \\
&\quad \succ \text{suspense} \sim \text{war}, \\
\tau_m &\equiv \text{mystery} \succ \text{suspense} \sim \text{adventure} \sim \text{drama} \\
&\quad \succ \text{comedy} \sim \text{war}, \\
\tau_s &\equiv \text{suspense} \succ \text{mystery} \succ \text{adventure} \sim \text{drama} \\
&\quad \succ \text{comedy} \sim \text{war}, \\
\tau_w &\equiv \text{war} \succ \text{adventure} \succ \text{comedy} \sim \text{drama} \sim \text{mystery} \\
&\quad \succ \text{suspense}. \qquad (1)
\end{aligned}$$

The important theoretical point is that each player's mental model is a structure that organizes preferences. Players can use the information embedded in this structure to collectively maximize the probability of coordination. Specifically, given shared mental models, agents can use a maximum-likelihood rule for determining which alternative beats all other alternatives in a particular choice context. Later it will be shown formally that this alternative is the action that is highest ranked over the distribution of preference types; furthermore, this alternative very often can be identified using a simple heuristic from the knowledge structure.

Assume for simplicity that players' preferences are distributed uniformly over the m ideal points and let the cost function simply be the path length in Figure 2 from a player's ideal point to that action. Then the sum of the rankings of each action from the orderings in (1) are:

$$\begin{aligned}
\text{adventure}: &\ 0+1+1+1+2+1=6, \\
\text{comedy}: &\ 1+0+1+2+3+2=9, \\
\text{drama}: &\ 1+1+0+1+2+2=7, \\
\text{mystery}: &\ 1+2+1+0+1+2=7, \\
\text{suspense}: &\ 2+3+2+1+0+3=11, \\
\text{war}: &\ 1+2+2+2+3+0=10. \qquad (2)
\end{aligned}$$

FIGURE 2 Shared Mental Map of the Film Genre



Adventure-mystery edge significant at .10 level; all other edges significant at .01 level.

From the expressions in (2), the activity with the lowest sum is the activity that is highest ranked over all preferences induced from organization of outcomes, which in this case is the adventure film. The outcome "meet in the adventure film theatre" is defined here as a knowledge-induced equilibrium. This outcome is the alternative that beats all other alternatives in this choice context using a maximum-likelihood rule.

Schelling's Parachutist Game

The logic outlined above also applies to games with incomplete information and to games with more than two players. To illustrate, consider Schelling's Parachutist Game extended to three players. Three parachutists each have a choice of m strategies, namely where to walk to meet the other parachutists (Figure 3a). Each player wants to reduce his costs of walking (by meeting at the location closest to his landing site) but will not receive the positive benefit of meeting unless all players coordinate on the same location. (It is assumed that players are unable to credibly communicate.) Each player knows the set of meeting places and is only informed about his own location and preferences over meeting places (referred to as a player's "type"). These informational conditions, as well as the distribution over players types, are common knowledge. Unfortunately, there are multiple Bayesian Nash equilibria in this game and players' preferences over these equilibria are in conflict.

If players form a mental map of the decision context, then it might look like that in Figure 3b, where players identify salient features of the landscape and connect these features based on empirical knowledge (such as that water runs downhill, bridges cross rivers, and farms are accessed by roads).[6] Players' preferences are assumed to be consistent with the empirical organization of the decision context. Assume in this case that the probability distribution over players' types (tributary, bridge, pond, road, driveway, farmhouse) is (.1,.3,.1,.3,.1,.1). Players know this aggregate information, although they do not know where particular other players land, because they have shared information about factors that affect where each might land, such as the topography or wind conditions. Then the outcome "players of all types meet at the road junction," with minimum weighted sum of rankings $(.1) \cdot 2 + (.3) \cdot 1 + (.1) \cdot 1 + (.3) \cdot 0 + (.1) \cdot 1 + (.1) \cdot 2 = .9$, is the outcome that beats all other outcomes in this choice context using a maximum-likelihood rule. This outcome is a Bayesian Nash equilibrium and is the unique prescription of the shared knowledge structure.

The Formal Model

A pure coordination game is a one-shot game with symmetric payoffs and no credible communication between players. The advantage of such stylized constructions of coordination problems (as in Schelling 1960; Gauthier 1975; Sugden 1995) is that they illustrate at its barest the problems of indeterminacy and belief convergence in

[6]Ullman (1996) presents empirical evidence that humans do identify the same key features (referred to as *perceptual saliency*). Evidence of the use of shared knowledge of empirical relations is in Knill and Richards (1996).

FIGURE 3a Map of Parachutists' Bargaining Game

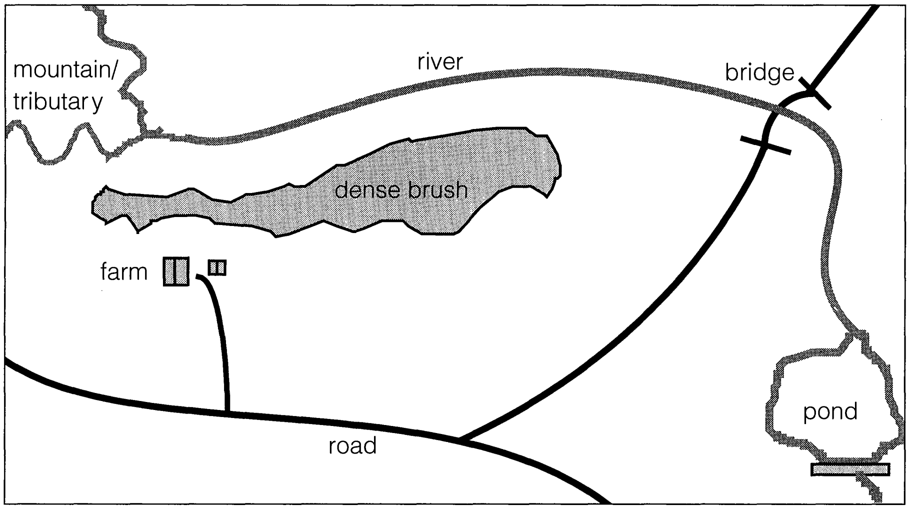
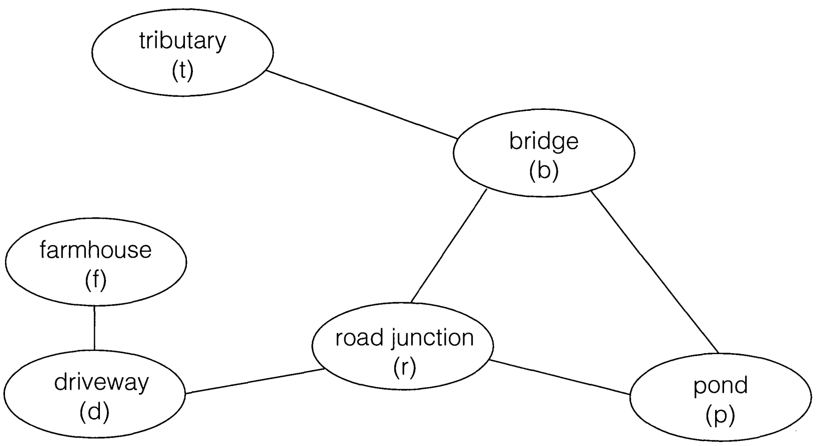


FIGURE 3b Hypothetical Mental Map of Figure 3a



games with multiple equilibria. This construction is empirically artificial, but the effect of shared knowledge structures in this game generalizes to more complicated settings as well.

A *knowledge structure game* Γ consists of a finite set of players, denoted $i = 1,\ldots,n$ and a finite set of actions $A = \{a_1,\ldots, a_m\}$, where each player's action set is symmetric and the set of actions is common knowledge. The set of actions is organized by a *knowledge structure* which is a labeled graph $\mathcal{M}(A,E)$ with vertices A and a set of edges E. A vertex may have one or more edges, but it is assumed that $\mathcal{M}$ is connected.[7] Each edge $e = \{a_j, a_k\}$ of $\mathcal{M}$ linking a_j and a_k corresponds to a similarity link based on the set of features or attributes between actions a_j and a_k. It is assumed that $\mathcal{M}$ is mutual knowledge, i.e., that all players organize the set of actions in the same way (as in Sugden's mutual knowledge about labels [1995, 536]), or more specifically, that a player assigns some probability ρ to the other players organizing the choice set in the same way. This assumption is weaker than common knowledge, which would require that all players know that all other players know that they all organize the choice set in the same way, *ad infinitum*.

In a coordination setting, the organization can also be thought of as an organization over *outcomes* rather than over actions, although the two are related. In a slight abuse of notation, when used to designate an organization over coordination outcomes, the graph will be

[7] The assumption that knowledge structures are connected implies that any feasible alternative in the choice set possesses at least one feature (e.g., Tversky 1977) that allows that alternative to be referentially related to another alternative in the choice set.

denoted as O, where the vertices $a_1,\dots,a_m$ of O represent coordination by all players at one of the m actions. The game is then referred to in shorthand as $\Gamma(O)$.

It is assumed that players' utility functions over A are consistent with their mental organization of A; thus the organization of outcomes constrains the set of feasible utility functions. This assumption has precedent from a variety of sources. For example, Bacharach (1993) refers to a "first phase of decision making where an agent arrives at some way of describing the options to herself" (see also Gauthier 1975; Nozick 1993, 134–135; Wendt 1992). Many of Anthony Downs's (1957) hypotheses stem from assumptions that preferences are connected to the structure of empirical choices (see also Hinich and Munger 1994). Most similarly, Black's theorem (1958) imposes a requirement that preferences are consistent with a linear ordering of the choice set. Black's assumption of a linear left-right continuum can be extended to the notion of ideological constraint in general, where preferences are constrained by a conceptual organization of the alternatives (e.g., Sullivan, Piereson, and Marcus 1978). Black's theorem can be thought of as the special case where $\mathcal{M}$ is a linear ordering; or conversely, the model presented here can be thought of as a graph-theoretic extension of Black's linear ordering of alternatives with its induced single-peaked preferences.

Specifically, each vertex $a_j \in O$ defines a *type* of player, in the Bayesian sense, who most prefers coordination at outcome a_j and whose preferences over the remaining actions follow from the knowledge structure O. Since utility functions are constrained by O, a player's type is defined by identifying that player's top-ranked outcome or *ideal* point. Players' types (or equivalently, ideal points) occur with a probability distribution $\varnothing$ over A. Players are informed about the probability distribution $\varnothing$ as well as their own type, but are unaware of other players' types.[8] Players make simultaneous choices of an action in A.

The graphs $\mathcal{M}$ and O are assumed to be unweighted, allowing players' utility functions for all m types to be defined by the path lengths through O. (Weighted relationships are straightforward but complicate the description.) Let B denote the positive payoff from successful coordination. Let $\beta(a_j;a_\tau)$ denote the cost to a type τ player (a player with ideal point a_τ) of choosing action a_j. For example, assuming an unweighted graph, $\beta(a_j;a_\tau)$ is equivalent to the number of edges on the shortest path length from ideal point a_τ to action a_j, namely $d(a_j, a_\tau)$. Let a^*_{-i} denote the equilibrium action of all other players except player i. The utility function of player i with ideal point a_τ is

$$u_{i,\tau}(a_j) = \begin{cases} B-\beta(a_j;a_\tau) & \text{if } a_j = a^*_{-i} \\ -\beta(a_j;a_\tau) & \text{if } a_j \neq a^*_{-i} \end{cases} \tag{3}$$

for $\tau = 1,\dots,m$. As in the examples above, players receive the benefit B only if they coordinate at the same action yet incur costs $\beta(a_j;a_\tau)$ whether or not they successfully coordinate due to the costs of choosing an action (e.g., Sugden 1995).

The presence of a shared interpretation of the choice environment provides information that allows players to maximize the probability of choosing the action that is top-ranked for all possible preferences. As Young (1986, 1995) shows, if one wants to maximize the probability of getting a correct social ranking over a set of choices, then the Condorcet rule is the "optimal rule."[9] However, in coordination games the full social ranking is unnecessary and one only needs to identify the most-likely top-ranked action. In this case Young shows that the Borda winner, the action that beats the other actions most often in a series, is the optimal rule that yields the maximum-likelihood estimate of the top-ranked alternative.[10] The following example illustrates this logic:

Example 4 *Maximum-likelihood Estimation and Schelling's Bargaining Game*

Consider the Schelling bargaining game of the previous section. In order to simplify the example I will assume that there are two players and that the probability of each type is $1/m$.[11] Since players are unaware of their own type

[8] This assumption is not new to a shared knowledge structure approach. Nearly all models of games of incomplete information assume that the probability distribution is known by the players. The justification is that aggregate information (such as through polls or the media) is more readily available than private, individual preference information.

[9] Young's maximum-likelihood approach applies both to cases where there is an objective true answer (as in the Condorcet Jury Theorem) and where the answer is endogenously derived from voters' preference rankings (e.g., Young 1986).

[10] Note that other maximum-likelihood rules might be appropriate for coordination settings, depending on what is to be maximized. Young's rule maximizes the probability of identifying the single alternative that is most likely to be top-ranked over voters' preferences, which is the fundamental problem of coordination. However, if, for example, robustness is the most important criterion (namely identifying the alternative that is most insensitive to changes in the choice set), then a different maximum-likelihood rule might be derived.

[11] More players requires sampling on a combinatorially larger set, such as *ttt*, *ttb*, *ttp*, ..., *fff* for three players. In general, the distribution of the sampling is proportional to the distribution of players' types $\varnothing$, and the Borda scores are correspondingly weighted by $\varnothing$. For example, with two players and $\varnothing = \left(\frac{1}{2},\frac{1}{2},0,\dots,0\right)$ the sampling is on *tt*, *tb*, *bt*, and *bb*, and the Borda score for r is $\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1+0\cdot(1+0+1+2)$.

ex ante and players remain unaware of each others' types in the interim, each player must maximize over all possible positions of both players. With two players there are m^2 possible combinations of players' types and $\binom{m}{2}$ pairwise comparisons of the actions. Since players do not have information about each other's type, in effect they are drawing a sample of each pairwise comparison of actions and asking which action has the highest probability of being the winner for all possible combinations of players' types. The action a_j that has the maximum-likelihood of being "best" for all possible player types is the action with the greatest probability of winning against all other $a_k \in A$, k j (Young 1986). In this example, outcome r is the maximum-likelihood winner since the probability that r is best is $\left(24+\frac{1}{2}\cdot 9\right)\Big/36$ against outcome t, $\left(15+\frac{1}{2}\cdot 13\right)\Big/36$ against outcome b, $\left(21+\frac{1}{2}\cdot 10\right)\Big/36$ against outcome p, $\left(16+\frac{1}{2}\cdot 16\right)\Big/36$ against outcome d, and $\left(20+\frac{1}{2}\cdot 13\right)\Big/36$ against outcome f, giving an average probability of $\frac{27.3}{36}$ of being the best coordination outcome (see also Young 1986, 117). Thus, action r has the greatest average probability of being best; the next best action is b with an average probability of $\frac{22.5}{36}$ and the action with the lowest average probability of being best is f with an average probability of $\frac{10.4}{36}$.[12] ■

This tedious sampling approach illustrates the connection to maximum-likelihood estimation but is unnecessary since Young (1986, 1988, 1995) shows the equivalency between this procedure and the much simpler procedure of choosing the action with the lowest Borda score (see appendix). Using this shortcut (which is independent of the number of players but not independent of the distribution of types), the outcome r is quickly identified as the maximum-likelihood winner since it has the minimum Borda score (or equivalently, is highest ranked over all types). However, to implement a decision based on the knowledge structure, agents need not engage in any calculations at all but can rely on a simple heuristic: in most cases, the maximum-likelihood winner (or equivalently the Borda winner) is the alternative in the knowledge structure $\mathcal{M}$ with maximum degree, namely the alternative that is adjacent to the greatest number of other alternatives in the choice set. Specifically, for random graphs and random distributions of players' types, the probability that an alternative that is the Borda winner is also an alternative with maximum degree is approximately .75 for choice sets up to ten alternatives.[13]

We can now precisely define a knowledge-induced equilibrium in a coordination game Γ with shared knowledge structure O:

Lemma 1 An action $a_j \in A$ is the Borda winner *iff* a_j is $\arg\min \sum_{a_k \in A} d(a_j, a_k)$

Proof. By the utility functions of Equation (3), the players' rankings over the outcomes $a_j \in O$ correspond to the distance from a_j to the ideal point for that player's type. The Borda winner is the action that is highest ranked over all players' types, which corresponds to the action with the lowest sum of distances to all ideal points, namely the a_j that minimizes $\sum_{a_\tau \in A} \beta(a_j; a_\tau)$. By the definition of $\beta(a_j;a_\tau)$ this is equivalent to the a_j that minimizes $\sum_{a_k \in A} d(a_j, a_k)$.[14] ■

Definition 1 *An outcome* a_j^* *is a* knowledge-induced equilibrium *of* $\Gamma(O)$ *iff* a_j^* *is* $\arg\min \sum_{a_k \in O} d(a_j^*, a_k)$.

Results

A knowledge-induced equilibrium can be thought of as a refinement to the set of (Bayesian) Nash equilibria for games where players share an organization of the action set. This section presents some general properties and results of a knowledge-induced equilibrium. Until the

[12] To recreate these values, make a 36 × 15 table where the rows are all combinations of players' types: *tt*, *tb*, *tp*, ..., *ff*, and the columns are the pairwise comparisons: *t* versus *b*, *t* versus *p*, ..., *d* versus *f*. For each cell, enter the action that is the winner in that pairwise comparison given that distribution of players' types and their utility functions. If two actions a_j and a_k are tied then each is best with probability 1/2 in that cell. The probability that t is best against *b* is the total number of times in the 36 rows that *t* beats *b* plus one-half the number of times *t* ties *b*.

[13] This result is based on simulations with random graphs (probability of an edge = .5) and random distributions of agents (∅) over the feasible preference types for m = 3,...,10,15, 20 with 1000 trials each. Even for choice sets as large as twenty alternatives, the Borda winner is the vertex with maximum degree in nearly two-thirds of the cases (.63) and was nearly always the vertex with either the maximum or one less than the maximum degree.

[14] Note that the maximum-likelihood winner is not equivalent to two common graph-theoretic definitions of centrality: the *center* (the vertex with minimum eccentricity) and the *centroid* of a graph (defined only for trees as the vertex with the minimum-maximum branch weight). Examples are easy to construct where the concepts do not coincide.

relaxation of this assumption in Propositions 6 and 7, it is assumed that the knowledge structure is shared.

Theorem 2 *A coordination game* $\Gamma(O)$ *has at least one knowledge-induced equilibrium.*

Proof. For any graph O, there is always at least one vertex a_j^* for which a_j^* is arg $\min \sum_{a_k \in O} d(a_j^*, a_k)$. ■

Corollary 3 *If* a_j^* *is a knowledge-induced equilibrium, then* a_j^* *is a (Bayesian) Nash equilibrium.*

Proof. In the construction of the coordination game $\Gamma(O)$, the graph O consists of the m coordination outcomes, each of which is a Nash equilibrium of the coordination game. A knowledge-induced equilibrium is a subset of the coordination outcomes of O. Given that players coordinate at a_j^*, they have no incentive to unilaterally deviate from their equilibrium action. ■

A knowledge-induced equilibrium is also a coalition-proof Nash equilibrium, as shown by the following corollary, implying that it is an efficient, self-enforcing agreement under nonbinding preplay communication (Bernheim et al. 1987).

Corollary 4 *A unique knowledge-induced equilibrium is a strong Nash equilibrium.*

Proof. Since coordination requires the choice of the same action by all players, unilateral changes of strategies by any coalition of players results in a lower payoff for at least one member of the coalition. ■

The force of the knowledge structure is to organize and restrict the set of players' preferences in particular consistent ways, thereby allowing for a maximum-likelihood winner to emerge (e.g., Saari 1994; Richards, McKay, and Richards 1998). As in other structural restrictions (e.g., Shepsle 1979; Black 1958), the knowledge structure contains implicit information because it constrains the extent of feasible preference types. To illustrate the force of this structure, the following simple (and familiar) example highlights how a collection of preferences not structured by a knowledge representation yields to a breakdown of a maximum-likelihood winner.

Example 5 *Preferences not consistent with a common knowledge structure.*

Assume three preference types: a $\succ$ b $\succ$ c, b $\succ$ c $\succ$ a, and c $\succ$ a $\succ$ b that do not form a consistent shared $\mathcal{M}$. Using the procedure of Young (1986), the probability that a is best against b is 6/9 and the probability that a is best against c is 3/9. Similarly, the probability that b is best against c is 6/9 and the probability that c is best against a is 6/9, yielding the intransitive information that a is likely to be better than b, c is likely to be better than a, and b is likely to be better than c. The average probability that any alternative is best is .5 for a, b, and c, yielding no maximum-likelihood winner. ■

Uniqueness of Knowledge-Induced Equilibria

In any coordination setting, uniqueness is a virtue. However, it is already apparent that a knowledge-induced equilibrium need not be unique since a graph O may have multiple vertices that minimize $\sum_{a_k=1}^{m} d(a_j^*, a_k)$. Thus, like other coordination prescriptions (e.g., Sugden 1995), a knowledge-induced equilibrium will provide guidance in some cases, but not in others. Before presenting the positive results, I illustrate some of the worst-case scenarios with examples.

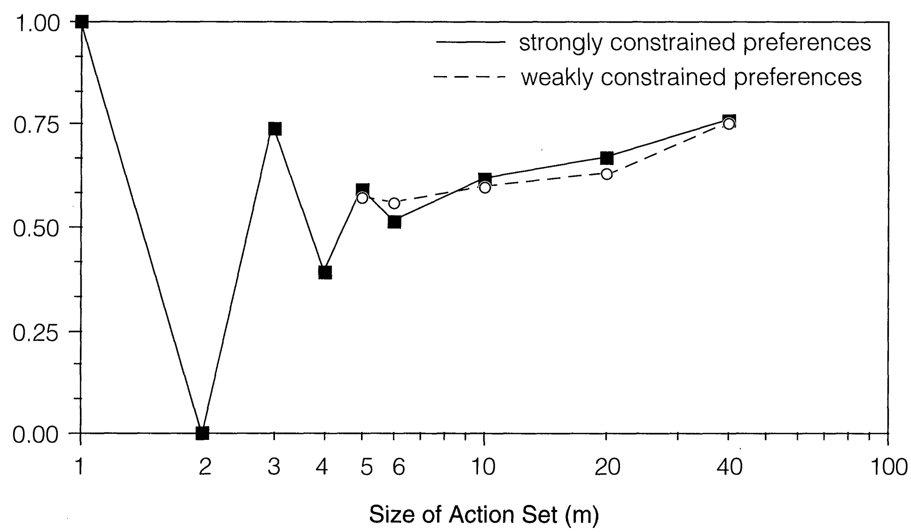
The set of equilibria need not be adjacent in O. For example, if O has edges $E = (a_1a_3, a_1a_5, a_2a_3, a_2a_5, a_3a_4, a_4a_5)$ then the outcomes a_3 and a_5 are both knowledge-induced equilibria but a_3 is not adjacent to a_5. Multiple knowledge-induced equilibria occur when O is symmetric and can potentially include the entire set of actions in nongeneric symmetric cases. Let O be a ring graph R_m. Then there are m knowledge-induced equilibria. It might be conjectured that the cardinality of the set of knowledge-induced equilibria depends solely on the symmetry of the knowledge structure. This is not true. Although the symmetry of the structure plays a role, it is not a perfect predictor, as the following example shows: if O has edges $E = (a_1a_2, a_2a_3, a_2a_6, a_3a_4, a_3a_5, a_4a_5, a_5a_6)$ then O is asymmetric but $\Gamma(O)$ has two knowledge-induced equilibria: outcomes a_2 and a_3.

However, particular forms of knowledge structures do lead to specific properties of the set of knowledge-induced equilibria. An empirically prevalent representational form is a *tree*, as in the taxonomy of animal species, kinship networks, or a city-block ideology metric (e.g., Minsky 1985; Corter 1996; Barberà, Gul, and Stacchetti 1993). The following theorem shows that if the knowledge structure is a tree (a connected acyclic graph), then the set of knowledge-induced equilibria consists of at most two adjacent equilibria. This implies that knowledge representations of these common forms will have special (and nice) coordination properties.

Theorem 5 *If O is a tree then* $\Gamma(O)$ *has at most two knowledge-induced equilibria and they will be adjacent in O.*

Proof. See appendix.

FIGURE 4 Probability of a Unique Knowledge-Induced Equilibrium



Examples are useful to illustrate the possibilities of uniqueness and its failures; however, to explore the question of how often a knowledge-structure game prescribes a unique equilibrium we need to turn to Monte Carlo simulations. The procedure is to generate random graphs with m vertices and estimate the percent of graphs which prescribe a unique equilibrium in a knowledge-structure graph. Each graph is a random (connected) graph with the probability of an edge equal to $\frac{1}{2}$, which corresponds to sampling uniformly from all labelled graphs on m vertices and which generates the greatest variety of nonisomorphic graphs for any fixed-edge probability and number of vertices.[15] Figure 4 summarizes the results. (Note the log scale.) There are several interesting observations. First, although a knowledge-structure game may have multiple equilibria, this is not the norm based on an examination of the set of random graphs. Obviously three-quarters of all knowledge structures with a choice set of three actions prescribe a unique equilibrium. Less intuitively, choice sets with five or six actions also have a probability of .5 or better of a unique equilibrium. However, a second observation is that some choice sets are clearly worse than others: most notable in this category are the cases with $m = 2$ and $m = 4$. The symmetry effects of even-versus-odd choice sets create drastic oscillation of the probability results for small choice sets, but disappears as m increases. Third, and probably most surprising, the prospects for a unique equilibrium *improve* rather than decline as the size of the choice set increases. Even for choice sets that are extremely large by empirical standards, such as $m = 40$, the probability of a unique equilibrium prescription in a knowledge-structure game is still approximately .75. Furthermore, the assumption that preferences are perfectly consistent with the knowledge structure can be relaxed in various ways. Similar results hold if players' preferences are structured by $\mathcal{M}$ only over a subset of vertices near an agent's ideal point (with indifference thereafter) rather than over the entire choice set. The dashed line in Figure 4 shows the case where preferences are constrained only within two edge steps from each agent's ideal point.

Uncertainty Over Others' Mental Models

Up until this point the model assumed that players organize the set of alternatives using similar abstract mental models. However, there are many reasons why players may not have shared models or may be unsure of other players' understanding of the decision context. In this section I relax the assumption of mutual knowledge to allow some probability that a subset of players do not hold the same shared model. Thus, even though a unique coordination equilibrium may exist, the possibility of others "not being on the same page" may undermine the equilibrium. This logic also occurs in the well-known Stag Hunt game, where the presence of uncertainty about

[15] Five hundred graphs were generated for cases with m 20. One hundred graphs were generated for the forty-vertex case. Only 100 trials were run for $m = 40$ because a single trial took thirty minutes of computer time. To verify the accuracy of the Monte Carlo simulations, the program was also run on the complete set of all connected graphs for $m = 3, 4, 5$, and 6 (from the appendix in Harary 1969).

others' understandings of the game may undermine the unique Pareto-optimal equilibrium. Factors such as cultural differences across players, high-risk situations, very large numbers of players, or a lack of institutional or social norms, may all contribute to a higher probability that miscoordination occurs (Jervis 1978).

Let $\mathcal{M}_i$ be the knowledge representation held by player i with knowledge-induced prescription k_i^*. The problem for coordination is that if one or more players has a different mental model then there is some probability that those players have a different knowledge-induced prescription and hence players will fail to coordinate. For simplicity, assume the uncertainty is constant across players and denote the probability of a shared knowledge structure as $\rho = prob(\mathcal{M}_j = \mathcal{M}_i)$ for all $j \quad i$. Then we are interested in the value of the expression $prob(k_i^* = k_j^*)$ for all $j \quad i$ given the probability ρ. Furthermore, we can explore the minimum-threshold values of ρ, denoted $\overline{\rho}$, such as when $prob(k_i^* = k_j^*) \geq .5$ for all $j \quad i$.

Let L_m denote the number of connected labeled graphs with m vertices and $\{L_m\}$ denote the set of such graphs. The following result summarizes the probability of coordination as a function of the number of players n, the number of strategies m, and the probability of a shared structure ρ.

Proposition 6 *Let $0 < \rho < 1$ be the probability that a player j's ($j \quad i$) knowledge representation $\mathcal{M}_j = \mathcal{M}_i$. The probability that all n players coordinate at a unique knowledge-induced equilibrium of $\mathcal{M}_i$ is*

$$\sum_{j=0}^{n-1} \binom{n-1}{j} \rho^{n-1-j} (1-\rho)^j \left(\frac{L_m - m}{L_m - 1} \cdot \frac{1}{m} \right)^j . \qquad (4)$$

Proof. The probability that all $n-1$ players have representation $\mathcal{M}_i$ (and hence the same knowledge-induced equilibrium) is ρ^{n-1}. The probability that $n-2$ players have representation $\mathcal{M}_i$ and exactly one player has a different knowledge representation is $\binom{n-1}{1} \rho^{n-2}(1-\rho)$. However, any player with a knowledge structure $\mathcal{M}_j \quad \mathcal{M}_i$ may still choose k_i^* if k_i^* is the knowledge-induced equilibrium of $\mathcal{M}_j$. By the symmetry of $\{L_m\}$, each vertex is a knowledge-induced equilibrium in $\frac{1}{m}$ of the graphs, hence the original knowledge-induced equilibrium k_i^* occurs with probability $\frac{1}{m}$. However, $\mathcal{M}_j$ is not drawn from the full set $\{L_m\}$ but from $\{L_m\}$ - $\mathcal{M}_i$. The probability of k_i^* occurring in $\{L_m\} - \mathcal{M}_i$ is $\frac{1}{L_m - 1}\left(\frac{L_m}{m} - 1\right)$. (The probability of any of the $m-1$ alternatives not equal to k_i^* occurring in $\{L_m\} - \mathcal{M}_i$ is $\frac{1}{L_m - 1}\left(\frac{L_m}{m}\right)$ which gives $\frac{1}{L_m - 1}\left(\frac{L_m}{m} - 1\right) + (m-1)\frac{1}{L_m - 1}\left(\frac{L_m}{m}\right) = 1$. Generalizing this logic into the weighted binomial sum for 2,3,...,$n-1$ players having different models yields

$$\rho^{n-1} + \binom{n-1}{1} \rho^{n-2}(1-\rho) \left[\frac{1}{L_m - 1}\left(\frac{L_m}{m} - 1\right) \right] + \binom{n-1}{2} \rho^{n-3}(1-\rho)^2 \left[\frac{1}{L_m - 1}\left(\frac{L_m}{m} - 1\right) \right]^2 + \ldots,$$

which simplifies to Equation (4). ■

Table 1 shows how confident a player must be about others' mental models in order to achieve at least a 50 percent probability of coordination if he plays his own knowledge-structure prescription. For example, with five players and twenty strategies, a player needs to place an 83 percent chance on each other player sharing his own representation in order to guarantee a 50 percent probability of coordination at his knowledge structure prescription. Although the probability of coordination overall remains approximated by ρ^n for large numbers of alternatives, this analysis points out that some uncertainty does not undermine the idea of a knowledge-structure prescription—*provided the number of players is not too large.* The following proposition provides the information for the last column of Table 1.

Proposition 7 *As $m \to \infty$, the value of $\overline{\rho}$ such that $prob(k_i^* = k_j^*) \geq .5$ for all $j \quad i$ ap proaches $\overline{\rho} = (.5)^{\frac{1}{n-1}}$, where n is the number of players.*

TABLE 1 Confidence in Shared Models for prob(coordination) 50%. ($\overline{\rho}$ such that $prob(k_i^* = k_j^*) \geq .5$ for all $j \quad i$)

		Number of Alternatives (m)					
		3	4	5	10	20	∞
Number of Players (n)	3	.67	.62	.63	.67	.69	.71
	4	.77	.73	.74	.77	.78	.79
	5	.82	.79	.80	.82	.83	.84
	6	.85	.83	.84	.86	.86	.87
	7	.88	.86	.86	.88	.89	.89
	10	.92	.90	.91	.92	.92	.93
	20	.959	.953	.955	.960	.962	.964

Proof. Using the binomial expansion
$\sum_{x=0}^{n-1}\binom{n-1}{x}a^x b^{n-1-x} = (a+b)^{n-1}$ with
$a = (1-\bar{\rho})\left(\frac{L_m - m}{L_m - 1}\right)\left(\frac{1}{m}\right)$ and $b = \bar{\rho}$ from Equation (4)
gives the inequality

$$\left((1-\bar{\rho})\left(\frac{L_m - m}{L_m - 1}\right)\left(\frac{1}{m}\right)+\bar{\rho}\right)^{n-1} \geq .5. \qquad (5)$$

As m increases, the value of L_m increases rapidly relative to m based on the series 4, 38, 728, 26704, ... (Sloane and Plouffe 1995, M3671), so the $\frac{L_m - m}{L_m - 1}$ term rapidly approaches one. In the limit as $m \to \infty$, Equation (5) simplifies to $\bar{\rho} \geq (.5)^{\frac{1}{n-1}}$. ■

Discussion

Nearly forty years ago Thomas Schelling proposed the idea of a focal point as an explanation of how players coordinate in an empirical setting. His insight was that a mental organization of the salient features of a decision context could potentially solve coordination problems even in the absence of communication or repeated play. This idea is so compelling that it is invoked to explain coordination outcomes in settings including collective action, third-party voting, campaign contributions, government regulations, international systems, and tariff policies (e.g., Chong 1991; Ainsworth and Sened 1993; Lohmann 1994; Spruyt 1994; Cox 1997; Scholz and Gray 1997; McGillvray 1997; O'Neill 1999). However, whereas a focal-point solution incorporates a payoff-independent labelling of the alternatives, a knowledge-induced equilibrium emphasizes the extent to which mental models are consistent with and constrain players' preferences. This shared restriction on the set of admissible preferences in turn facilitates coordination, despite conflicting ideal points among players.

The advantage of developing a formal model of the intersection of knowledge representations and choice is that it allows for a reexamination of the conditions that influence successful coordination. The model suggests some new implications for old variables and some new variables for future inquiry. Four variables immediately emerge from the model and results: (1) the number of choices in the action set, (2) the number of players, (3) the form of a knowledge structure, and (4) the extent to which mental models are shared. All these variables are empirically measurable.[16]

The number of choices and the number of players are frequently mentioned in association with coordination and bargaining. Since the combinations of possibilities explode with both these variables, it is typically inferred that both more choice and more players exacerbate a coordination problem. However, *if* players hold similar abstract mental models, then coordination over more choices is not necessarily more difficult, as seen in the Monte Carlo results. In fact, the coordination indeterminacy was significantly worse with only four choices than with six or ten or even twenty choices. If players put some nonzero probability on others holding a different mental model, then more choices still did not result in drastic changes to coordination prospects. For example, with five players, regardless of whether the coordination problem was over three or 100 choices, a probability of shared knowledge of .71 guaranteed a 50 percent chance of coordination if a player chose the knowledge-induced equilibrium of his own knowledge structure.

However, the number of players had a different effect. *If* players have shared mental models, then since a knowledge-induced equilibrium is defined from the knowledge structure, including more players makes little difference to the cognitive complexity of the problem since the maximum-likelihood winner is the same for all types of players. A puzzle in empirical political coordination is how large groups of players, as in collective action problems or third-party voting, manage to coordinate their actions given the exploding combinatorics on strategy vectors and beliefs. This model suggests that *if* there is some level at which players meta-organize their understandings of the choices (despite conflicting preferences over these choices), then coordination can still occur in large groups.

However, as the number of players increases presumably the empirical probability of shared mental models decreases, although some cultural anthropologists and evolutionary psychologists make arguments to the contrary (e.g., Barkow, Cosmines, and Tooby 1992; Romney et al. 1996). Regardless of the debates on culture or cognition, the formal results did show that larger numbers of players require a higher threshold for shared knowledge in order to achieve coordination. Furthermore, this threshold probability increases relatively rapidly. For example, although with only three players the

[16] The latter two variables are measurable using statistical techniques based on multidimensional scaling (e.g., Romney, Batchelder, and Weller 1987 in anthropology; Richards and Koenderink 1995 in cognitive science).

probability of shared knowledge must be only .71 or greater, for six players the threshold increased to .87, and by twenty players was .965.

The results also emphasize that the *form* of knowledge in a given empirical context is an important variable in coordination. Is knowledge represented as a tree or in rings or is there high or low symmetry in the organization? These qualitative attributes of the empirical context may be more important factors in coordination than simply the number of choices. Although the idea of structures of knowledge is not new in political science, the qualitative characteristics of the form of knowledge has not been considered as an important variable in predicting collective outcomes. In some cases, a knowledge structure suggests a unique prescription; in other cases it has little coordinating power. An example of empirical variations in the form of knowledge is evident in collective action problems of mass mobilization (e.g., Chong 1991; Lohmann 1994). My model formalizes how characteristics of a shared mental model can overcome the effects of large numbers of players. Most successful mass mobilizations have shared knowledge structures with strong uniqueness properties. Probably the best empirical example is the spontaneous gathering of East Germans on the Ringstrasse encircling Karl-Marx-Platz—a stark unique knowledge-induced equilibrium of probably every resident's mental map of central Leipzig (Lohmann 1994, 67–8).

Finally, although the model initially assumed a shared mental model among players, the extent to which a shared representation exists is an important variable for empirical applications of the model. For example, multiparty democracies involve coordination problems in that voters must coordinate on which of the smaller parties reaches the threshold for third-party representation (Cox 1997; Myerson and Weber 1993). If voters have shared cognitive organizations of these parties then the prospects for coordination among minority parties should be much greater than suggested by a traditional game-theoretic analysis. In contrast, where there is no shared cognitive organization of the parties, as in the case of emerging democracies with numerous newly formed political parties, coordination on third parties would be expected to be more difficult, leading to frequent changes in party representation and coalition structure.

As another example, there may be differences in the extent of structure between different subgroups of agents, as in Converse's (1964) assertion that the belief structure of elites is tighter than the belief structure of the masses. Groups with tightly structured shared mental models would be expected to coordinate more easily based on this additional information, whereas groups with loosely structured shared models would have more difficulty coordinating and would need to rely on other sources, such as institutions, to solve their coordination problems. Similarly, in conflict resolution bargaining (e.g., Schelling 1960; Legro 1995), the model suggests that coordination will be much more successful if (or when) a shared mental model emerges among the participants. Coordination can be facilitated by creating a shared understanding of the relationships among the feasible negotiation proposals. This implies that even nonenforcing third-party mediators (such as an individual diplomat, a nongovernmental organization, or an institution such as the United Nations) can potentially have great influence over the ability of participants to reach a coordinated agreement—if they can facilitate a convergence of the participants' representations of the bargaining context.

If the form and existence of shared knowledge matter in coordination, this points to the power of manipulating knowledge representations to influence coordination outcomes, as in Riker's (1983) concept of "heresthetic" (the art of structuring a choice context to one's benefit) or Myerson and Weber's (1993) "focal arbiter." Rather than sending a cheap talk signal of the form "I am going to choose A," one can expand the communication aspects to signals of the form "A is like B," where players communicate representations rather than intentions (Banks and Calvert 1992; Palfrey and Rosenthal 1991). Since a knowledge-induced equilibrium is a global property of players' structure of the set of alternatives, small changes in this structure can have a big effect on the attributes of the equilibria. One way knowledge structures can be influenced is by adding or subtracting new alternatives or by changing perceptions of how alternatives are related (e.g., Riker 1983; Sebenius 1983; Myerson and Weber 1993; Jones 1994; Bawn 1999). For example, let $\mathcal{M}$ be the graph with $E = (a_1a_2, a_3a_2, a_3a_4, a_4a_1)$. The addition of a new action to the choice set: a_5 with $a_5a_2 \in E$, transforms the game from one where the full set of outcomes are knowledge-induced equilibria to a game with a unique equilibrium at all players choose a_2. Arguing relationships among existing alternatives can also radically shift the coordination outcome. Let $\mathcal{M}$ be the graph with $E = (a_1a_2, a_3a_2, a_2a_4, a_4a_5, a_5a_6, a_6a_7)$. The addition of an edge linking a_7 with a_2 in a knowledge structure shifts the outcome from all players choosing a_4 in equilibrium to all players choosing a_2. Manipulation of the knowledge structure can occur in simple ways such as arguing relationships and categories (such as assertions that "proposal A is like proposal B so if you like A you should like B"). It can also occur in more complex ways such as the framing effects of priming, or invoking relevant comparison dimensions as in issue-saliency tactics in elections (e.g., Jones 1994; Cox 1997, 255–61), or making analogical arguments (e.g., Khong 1992). These simple examples highlight the importance of communication—not only of preferences or intent—but of representations of an empirical context.

Conclusion

I have presented a formal model of how players' mental models of an empirical decision context intersect with strategy choice to influence the prospects for a coordination equilibrium. This approach emphasizes that cognitive elements, particularly the representation of an empirical context through mental models, are an important structural factor that mediates between the interactions of agents and between agents and their collective environments. The representation of knowledge used here—namely categories and relationships between categories—is probably the simplest possible representation. Much richer representations are possible, including analogies, cause-and-effect models, schemas, classifications of sets, logical systems, spatial representations, and so on. The model outlined here can be extended to take into account these more complex organizations of knowledge. Furthermore, a knowledge structure need not be modeled as static, but can be viewed as a dynamic structure that emerges through the interactions of agents (such as through deliberation, manipulation, or signaling (Richards 2000)). This would bring in communication and emergent structures, although the intent of this article is to demonstrate the effects of shared knowledge even in the absence of communication. The set of categories also need not be fixed in advance but can evolve as a representation is constructed.[17] Finally, this is a formal model where the empirical context matters.

It is worthwhile clarifying the relationship between the knowledge-induced equilibrium defined here and the related equilibrium concept of a structure-induced equilibrium defined for cooperative games (Shepsle 1979). A knowledge-induced equilibrium is based on general maximum-likelihood procedures and thus can be defined for both social choice (where the maximum-likelihood procedure becomes plurality rule as in Richards, McKay, and Richards [1998, 2000]) and for noncooperative games as shown here. In addition, a knowledge-induced equilibrium is based on the organization of preferences as mutually consistent with players' mental representations of the context, rather than on procedural rules and jurisdictions. But both a knowledge-induced equilibrium and a structure-induced equilibrium incorporate structure from the empirical context (whether it is an institutional or cognitive organization) to constrain and organize the set of players' preferences and induce equilibrium properties.

The idea of a knowledge-induced equilibrium reminds us that there are other sources of structure that induce stability in collective choice. Furthermore, these structures may have interesting relationships between them, such as the relationship between institutions and mental models, or mental models and cultural or social norms. However, the almost exclusive emphasis on institutional rules and procedures as the mechanism for facilitating collective agreements implies that without strong institutions the prospects for stable agreements are slight. This is not the case. Even in realms without institutional restrictions on preferences, such as in international politics or politics outside of parliaments or committees, preferences may be structured by shared cognitive interpretations of the empirical context which are sufficient to induce stable collective choices.

Manuscript submitted February 29, 2000.
Final manuscript received September 25, 2000.

Appendix
Maximum Likelihood and Borda Winner

Lemma 8 *The alternative that is the maximum-likelihood winner of* Γ *is the action with the greatest pairwise wins minus losses, which is equivalent to the action that is the Borda winner.*

Lemma 8 is based on the results of Young (1986, 1988, 1995), who examines optimal aggregation rules as a problem of statistical inference using maximum-likelihood estimation. Let $A = \{a_1,..., a_m\}$ be m distinct actions. Assume that the goal is to correctly identify the "best" action(s) after a series of independent pairwise comparisons. A series of pairwise comparisons on a set of m alternatives A is represented by an $m(m-1)$-dimensional vector $\mathbf{x} = (x_{jk})$, 1 j k m, consisting of nonnegative integer entries where x_{jk} is the number of comparisons in which a_j is "better than" a_k. If two actions are equally "best" in any single pairwise comparison, then it is assumed that each is selected with probability 1/2. It is assumed that pairwise trials are independent and that every action is involved in an equal number of

[17] Most decision theories must assume an *a priori* fixed set of alternatives (called the "small world assumption"), which is often a hindrance in applying formal models to dynamic empirical contexts.

comparisons c. The sampling distribution within the c comparisons corresponds to the distribution of combinations of the players' types given $\varnothing$.

Let $g(\mathbf{x}|a_k)$ be the probability of observing $\mathbf{x}$ given that a_k is the "best" action, where

$$g(\mathbf{x}|a_k) = p^{\sum_{j\neq k} x_{kj}} (1-p)^{\sum_{j\neq k} x_{jk}}. \tag{6}$$

Since for all $j \quad k$, $x_{jk} + x_{kj} = c$, the maximum-likelihood decision rule is (Young 1986, 116):

$$f(x) = \left\{ a_k \in A: \sum_{j\neq k} x_{kj} \geq \sum_{j\neq k} x_{lj}, \text{ for } 1 \leq l \leq m \right\}. \tag{7}$$

Equation 7 corresponds to choosing the action with the maximum number of wins minus losses over all pairwise comparisons, which corresponds to the action that is the Borda winner with probability weights $\varnothing$ (Young 1986, 116; Black 1958).

Proof of Theorem 5

Let v_1^* and v_3^* denote two knowledge-induced equilibria in a tree graph $\mathcal{M}$. Assume v_1^* and v_3^* are not adjacent so that there must be at least one vertex (or set of vertices), denoted v_2, with $e(v_2, v_1^*)$ and $e(v_2, v_3^*) \in \mathcal{M}$. Define a branch at a vertex u of a tree as the maximal subtree containing u as an endpoint. Let $\sum p(u)_{-x}$ denote the sum of the path lengths from root node u to all vertices descending from u excluding vertex x and its descendants. Let n_1 denote the number of vertices following from v_1^* as the root node excluding v_2 and its branches, n_2 be the number of vertices following from v_2 excluding v_1^* and its branches and v_3^* and its branches, and n_3 be the number of vertices following from v_3^* excluding v_2 and its branches. Since v_1^* and v_3^* are both knowledge-induced equilibria it must be that $\sum_{u\in\mathcal{M}} d(u, v_1^*) = \sum_{u\in\mathcal{M}} d(u, v_3^*)$, or that

$$\begin{aligned} &\sum p(v_1^*)_{-v_2} + 1 + 2 + \sum p(v_3^*)_{-v_2} + 2n_2 + \sum p(v_2)_{-v_1^*,-v_3^*} + n_3 \\ &= \sum p(v_3^*)_{-v_2} + 1 + 2 + \sum p(v_1^*)_{-v_2} + 2n_1 + \sum p(v_2)_{-v_1^*,-v_3^*} + n_3, \end{aligned} \tag{8}$$

which implies that $n_1 = n_2$. In order for v_2 not to be a knowledge-induced equilibrium, it must be that $\sum_{u\in\mathcal{M}} d(u, v_2)$ is not a minimum, or that

$$\begin{aligned} &\sum p(v_1^*)_{-v_2} + n_1 + 1 + \sum p(v_3^*)_{-v_2} + n_2 + 1 + \sum p(v_2)_{-v_1^*,-v_3^*} \\ &> \sum p(v_1^*)_{-v_2} + 1 + 2 + \sum p(v_3^*)_{-v_2} + 2n_2 + \sum p(v_2)_{-v_1^*,-v_3^*} + n_3, \end{aligned} \tag{9}$$

which results in the contradiction $n_3 < -1$. Hence v_1^* and v_3^* must be adjacent if they are knowledge-induced equilibria.

It remains to be shown that the set of adjacent equilibria cannot be greater than two. For v_1^*, v_2^*, and v_3^* to be knowledge-induced equilibria it must be that the sum of path lengths is minimal and equal:

$$\begin{aligned} &\sum p(v_1^*)_{-v_2} + 1 + n_2 + \sum p(v_2^*)_{-v_1^*-v_3^*} + 2 + 2n_3 + \sum p(v_3^*)_{-v_2^*} \\ &= 1 + n_1 + \sum p(v_1^*)_{-v_2^*} + \sum p(v_2^*)_{-v_1^*-v_3^*} + 1 + n_3 + \sum p(v_3^*)_{-v_2^*} \\ &= 2 + 2n_1 + \sum p(v_1^*)_{-v_2^*} + 1 + n_2 + \sum p(v_2^*)_{-v_1^*-v_3^*} + \sum p(v_3^*)_{-v_2^*}. \end{aligned} \tag{10}$$

In order for Equations (10) to be satisfied, it must be that $n_2 = -1$, which is a contradiction. ■

References

Ainsworth, Scott, and Itai Sened. 1993. "The Role of Lobbyists: Entrepreneurs with Two Audiences." *American Journal of Political Science* 37:834–866.

Austen-Smith, David, and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review* 90:34–45.

Axelrod, Robert. 1973. "Schema Theory: An Information Processing Model of Perception and Cognition." *American Political Science Review* 67:1248–1266.

Axelrod, Robert. 1986. "An Evolutionary Approach to Norms." *American Political Science Review* 80:1095–1111.

Bacharach, Michael. 1993. "Variable Universe Games." In *Frontiers of Game Theory,* ed. K. Binmore, A. Kirman, and P. Tani. Cambridge, Mass.: MIT Press.

Bacharach, Michael, and Michele Bernasconi. 1997. "The Variable Frame Theory of Focal Points: An Experimental Study." *Games and Economic Behavior* 19:1–45.

Banks, Jeffrey S., and Randall L. Calvert. 1992. "A Battle-of-the-Sexes Game with Incomplete Information." *Games and Economic Behavior* 4:347–372.

Barberà, Salvador, Faruk Gul, and Ennio Stacchetti. 1993. "Generalized Median Voter Schemes and Committees." *Journal of Economic Theory* 61:262–289.

Barkow, Jerome H., Leda Cosmides, and John Tooby (ed.). 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* New York: Oxford University Press.

Bawn, Kathleen. 1999. "Constructing 'Us': Ideology, Coalition Politics, and False Consciousness." *American Journal of Political Science* 43:303–334.

Bernheim, B. Douglas, Bezalel Peleg, and Michael D. Whinston. 1987. "Coalition-Proof Nash Equilibria: Concepts." *Journal of Economic Theory* 42:1–12.

Black, Duncan. 1958. *The Theory of Committees and Elections.* Boston: Kluwer Academic Publishers.

Chong, Dennis. 1991. *Collective Action and the Civil Rights Movement.* Chicago: University of Chicago Press.

Clausen, Aage R. 1973. *How Congressmen Decide: A Policy Focus.* New York: St. Martin's Press.

Conover, Pamela Johnston, and Stanley Feldman. 1984. "How People Organize the Political World: A Schematic Model." *American Journal of Political Science* 28:95–126.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent,* ed. David Apter. New York: Free Press.

Corter, James E. 1996. *Tree Models of Similarity and Association.* Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07–112. Thousand Oaks, Calif.: Sage.

Cosmides, Leda, and John Tooby. 1992. "Cognitive Adaptations for Social Exchange." In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture,* ed. Jerome H. Barkow, Leda Cosmides, and John Tooby. New York: Oxford University Press.

Cox, Gary. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems.* New York: Cambridge University Press.

Crawford, Vincent P., and Hans Haller. 1990. "Learning How to Cooperate: Optimal Play in Repeated Coordination Games." *Econometrica* 58:571–595.

Denzau, Arthur T., and Douglass C. North. 1994. "Shared Mental Models: Ideologies and Institutions." *Kyklos* 47:3–31.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

Ferejohn, John. 1991. "Rationality and Interpretation." In *The Economic Approach to Politics,* ed. Kristen Renwick Monroe. New York: Harper Collins.

Gauthier, David. 1975. "Coordination." *Dialogue* 14:195–221.

Haas, Peter M. 1990. *Saving the Mediterranean.* New York: Columbia University Press.

Harary, Frank. 1969. *Graph Theory.* New York: Addison-Wesley.

Hardin, Russell. 1982. *Collective Action.* Baltimore: Johns Hopkins University Press.

Hinich, Melvin, and Michael Munger. 1994. *Ideology and the Theory of Political Choice.* Ann Arbor: University of Michigan Press.

Jervis, Robert. 1978. "Cooperation under the Security Dilemma." *World Politics* 30:167–214.

Jones, Bryan D. 1994. *Reconceiving Decision-making in Democratic Politics: Attention, Choice, and Public Policy.* Chicago: University of Chicago.

Khong, Yuen Foong. 1992. *Analogies at War.* Princeton: Princeton University Press.

Knill, David C., and Whitman Richards (ed.). 1996. *Perception as Bayesian Inference.* New York: Cambridge University Press.

Kollman, Ken, John H. Miller, and Scott E. Page. 1992. "Adaptive Parties in Spatial Elections." *American Political Science Review* 86:929–937.

Kreps, David M. 1991. "Corporate Cultures and Economic Theory." In *Perspectives on Positive Political Economy,* ed. James E. Alt and Kenneth A. Shepsle. New York: Harper Collins.

Legro, Jeffrey. 1995. *Cooperation Under Fire: Anglo-German Restraint During World War II.* Ithaca: Cornell University Press.

Levi, Margaret. 1988. *Of Rule and Revenue.* Berkeley: University of California Press.

Levi, Margaret. 1997. *Consent, Dissent, and Patriotism.* New York: Cambridge University Press.

Lohmann, Susanne. 1994. "Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–1991." *World Politics* 47:42–101.

Luce, R. Duncan, and Howard Raiffa. 1985. *Games and Decisions.* New York: Dover.

McGillvray, Fiona. 1997. "Party Discipline as a Determinant of the Endogenous Formation of Tariffs." *American Journal of Political Science* 41:584–607.

Mehta, Judith, Chris Starmer, and Robert Sugden. 1994. "The Nature of Salience: An Experimental Investigation of Pure Coordination Games." *American Economic Review* 84:658–673.

Miller, Nicholas. 1986. "Information, Electorates, and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem." In *Information Pooling and Group Decision Making,* ed. Bernard Grofman and Guillermo Owen. Greenwich, Conn.: JAI Press.

Minsky, Marvin. 1985. *Society of Mind.* New York: Simon and Schuster.

Myerson, Roger B., and Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review* 81:102–114.

Nozick, Robert. 1993. *The Nature of Rationality.* Princeton: Princeton University Press.

O'Neill, Barry. 1999. *Honor, Symbols, and War.* Ann Arbor: University of Michigan Press.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* New York: Cambridge University Press.

Palfrey, Thomas, and Howard Rosenthal. 1991. "Testing for Effects of Cheap Talk in a Public Goods Game with Private Information." *Games and Economic Behavior* 3:183–220.

Poole, Keith T., and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* 35:228–278.

Richards, Diana, Brendan D. McKay, and Whitman Richards. 1998. "Collective Choice and Mutual Knowledge Structures." *Advances in Complex Systems* 1:221–236.

Richards, Diana. 2000. "Strategic Persuasion and the Manipulation of Knowledge Structures in Social Choice." Presented at the American Political Science Association Meeting, Washington, D.C. (http://PRO.harvard.edu).

Richards, Diana. 2001. "Reciprocity and Shared Knowledge Structures in the Prisoner's Dilemma Game." Unpublished manuscript. University of Minnesota.

Richards, Whitman, and Jan J. Koenderink. 1995. "Trajectory Mapping: A New Nonmetric Scaling Technique." *Perception* 24:1315–1331.

Richards, Whitman, Brendan D. McKay, and Diana Richards. 2000. "The Probability of Collective Choice with Shared Knowledge Structures." M.I.T. A.I. Technical Report 1690 (http://www.ai.mit.edu/publications/pubsDB/pubsDB.onlinehtml).

Riker, William H. 1983. "Political Theory and the Art of Heresthetics." In *Political Science: The State of the Discipline,* ed. Ada W. Finifter. Washington, D.C.: American Political Science Association.

Romney, A. Kimball, William H. Batchelder, and Susan C. Weller. 1987. "Recent Applications of Cultural Consensus Theory." *American Behavioral Scientist* 31:163–177.

Romney, A. Kimball, John P. Boyd, Carmella C. Moore, William H. Batchelder, and Timothy J. Brazill. 1996. "Culture as Shared Cognitive Representations." *Proceedings of the National Academy of Sciences* 93:4699–4705.

Saari, Donald G. 1994. *Geometry of Voting.* Berlin: Springer-Verlag.

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

Schiemann, John W. 2000. "Meeting Halfway between Rochester and Frankfurt: General Salience, Focal Points, and Strategic Interaction." *American Journal of Political Science* 44:1–16.

Scholz, John T., and Wayne B. Gray. 1997. "Can Government Facilitate Cooperation? An Informational Model of OSHA Enforcement." *American Journal of Political Science* 41:693–717.

Sebenius, James K. 1983. "Negotiation Arithmetic: Adding and Subtracting Issues and Parties." *International Organization* 37:281–234.

Shepsle, Kenneth. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science* 23:27–59.

Sloane, N.J.A., and Simon Plouffe. 1995. *The Encyclopedia of Integer Sequences.* New York: Academic Press.

Spruyt, Hendrik. 1994. *The Sovereign State and Its Competitors: An Analysis of System Change.* Princeton: Princeton University Press.

Sugden, Robert. 1995. "A Theory of Focal Points." *The Economic Journal* 105:533–550.

Sullivan, John L., James E. Piereson, and George E. Marcus. 1978. "Ideological Constraint in the Mass Public: A Methodological Critique and Some New Findings." *American Journal of Political Science* 22:233–249.

Taylor, Michael. 1987. *The Possibility of Cooperation.* New York: Cambridge University Press.

Tversky, Amos. 1977. "Features of Similarity." *Psychological Review* 84:327–352.

Ullman, Shimon. 1996. *High-level Vision: Object Recognition and Visual Cognition.* Cambridge: M.I.T. Press.

Weingast, Barry. 1995. "A Rational Choice Perspective on the Role of Ideas." *Politics and Society* 23:449–464.

Wendt, Alexander E. 1992. "Anarchy is What States Make of It: The Social Construction of Power Politics." *International Organization* 46:391–425.

Wendt, Alexander E. 1999. *Social Theory of International Politics.* New York: Cambridge University Press.

Wittkopf, Eugene R., and Michael A. Maggiotto. 1983. "Elites and Masses: A Comparative Analysis of Attitudes Toward America's World Role." *Journal of Politics* 45:303–334.

Young, H. Peyton. 1986. "Optimal Ranking and Choice from Pairwise Comparisons." In *Information Pooling and Group Decision Making,* ed. Bernard Grofman and Guillermo Owen. Greenwich, Conn.: JAI Press.

Young, H. P. 1988. "Condorcet's Theory of Voting." *American Political Science Review* 82:1231–1244.

Young, H. Peyton. 1995. "Optimal Voting Rules." *Journal of Economic Perspectives* 9:51–64.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure.* Princeton: Princeton University Press.

Coordination and Shared Mental Models

Diana Richards

In "Coordination and Shared Mental Models" (*American Journal of Political Science* 45(2): 259–276), a printer's error led to the omission of some inequality and "not equal to" symbols at several points in the text. The following are corrections.

The footnote on page 269 should read:

"Five hundred graphs were generated for cases with $m \leq 20$."

On page 270, every occurrence of "$j \quad i$" (lines 9, 11, and 13 in paragraph 2, the statements of Proposition 6 and 7, and the heading of Table 1) should read:

"$j \neq i$"

Also on page 270, the sixth line in the proof of Proposition 6 omits a "not equal to" symbol and should read:

"$\mathcal{M}_j \neq \mathcal{M}_i$"

Also on page 270, the first sentence in the heading of Table 1 should read:

"Confidence in Shared Models for prob(coordination) ≥ 50%."

On page 273, the eighth and ninth line after Lemma 8 should read:

"by an $m(m-1)$-dimensional vector $\mathbf{x} = (x_{jk})$, $1 \leq j \neq k \leq m$, ..."

Department of Political Science
University of Minnesota
(richards@polisci.umn.edu)