

Chicago-Kent Law Review

Volume 79

Issue 3 Symposium: "Law &": Philosophical,
Psychological, Linguistic, and Biological Perspectives on
Legal Scholarship

Article 30

October 2004

Common Knowledge, Communication, and Public Reason

Bruce Chapman

Follow this and additional works at: <https://scholarship.kentlaw.iit.edu/cklawreview>

 Part of the [Law Commons](#)

Recommended Citation

Bruce Chapman, *Common Knowledge, Communication, and Public Reason*, 79 Chi.-Kent L. Rev. 1151 (2004).
Available at: <https://scholarship.kentlaw.iit.edu/cklawreview/vol79/iss3/30>

This Article is brought to you for free and open access by Scholarly Commons @ IIT Chicago-Kent College of Law. It has been accepted for inclusion in Chicago-Kent Law Review by an authorized editor of Scholarly Commons @ IIT Chicago-Kent College of Law. For more information, please contact dginsberg@kentlaw.iit.edu.

COMMON KNOWLEDGE, COMMUNICATION, AND PUBLIC REASON

BRUCE CHAPMAN*

INTRODUCTION

It may not be a particularly flattering metaphor for the lawyer, but when that unsavory gangster on the television cop show voices his reluctance to answer any questions from the police “until his mouth-piece gets here,” he immediately reminds us about how important *talk* is to the law and legal processes. Conversations with the police may begin as basic exchanges of information, but with the arrival of the lawyer on the scene the facts exchanged will take on the special shape that only the ritual of legal talk can provide. Some questions will be answered, others met with a measured silence. Further, if the case proceeds to trial, or on to an appeal, the different lawyers will shape the facts even further with their structured and sequenced arguments, the one framing a specific claim, the other offering an objection or a defense in response to that claim, the first replying with a rejoinder to that response, in an ever increasing specification of the legal issues.¹

Ultimately, the relative success of each lawyer in this to and fro process will be assessed by a judge, or perhaps by a jury carefully advised by a judge as to the precise structure of the legal issues properly before it. Further, the judge will not generally be content merely to provide the litigants with what she thinks is a just result for their case. She will also offer more talk of her own, providing a set of publicly

* Faculty of Law, University of Toronto. I am grateful to Alan Brudner, Abraham Drassinower, Joseph Heath, Ed Iacobucci, Mikael Priks, Arthur Ripstein, Ariel Rubinstein, and Yanis Varoufakis for very helpful discussions in the course of writing and revising this Article, and to the participants at both the Symposium on “Law &” at the Chicago-Kent College of Law in October 2003 and the Annual Meeting of the European Public Choice Society in Berlin, April 2004, for their comments. Research funding from the Social Science and Humanities Research Council of Canada is also gratefully acknowledged.

1. For a discussion of how these back and forth legal conversations can provide a conceptually disciplined method of accommodating what are otherwise incommensurable values, see Bruce Chapman, *Law, Incommensurability, and Conceptually Sequenced Argument*, 146 U. PA. L. REV. 1487 (1998).

accessible reasons for the result which, in their own turn, might set off further legal conversations in appeal courts down the road. Indeed, the ultimate legal authority of the result she has determined for the case may well depend on how it survives these ensuing discussions.

So talk is of the essence in legal adjudication and the development of law. In this respect, law is like social life more generally. Its conversations may be more stilted and stylized, and its persuasive effects often may be enforceable by the state, but, just as for other systems of social interaction, linguistic communications are central. The simple fact of the matter is that individuals talk to one another about what they do and what they want to do, and they do this as if it really mattered.

Thus, it must surely come as a surprise to some that in game theory—probably the most theoretically sophisticated account of social interaction that has been developed in the social sciences in recent years—the role of language has largely been ignored and not accounted for.² Of course, in the theory of games, agents interact with one another, often thinking, silently, in a very sophisticated way, about each others' actions and thoughts. Seldom, however, does what they *say* to each other have much obvious impact on what they do. Indeed, the term of art for these verbal exchanges within game theory is *cheap talk*, a phrase that hardly instills confidence that the theory—in contrast to our aforementioned gangster—has appreciated the significance, at least for one's legal interactions, of investing in a lawyer to help with the legal conversations one might have.

In this Article, I hope to explain why game theory has been so unsuccessful in accounting for the role of language in social interaction. I will begin by exploring some of its most basic difficulties in this respect, focusing on games of pure coordination, and trace these difficulties back to the most fundamental organizing concepts in the theory of games, namely, Nash equilibrium and common knowledge of rationality. Nash thinkers and Nash actors, I shall argue, are doomed to have very impoverished conversations as Nash talkers. The sorts of conversations they will have will leave them paralyzed in games of pure coordination, and largely uncooperative in games

2. See, e.g., Joseph Farrell, *Communication, Coordination and Nash Equilibrium*, 27 ECON. LETTERS 209, 213 (1988) ("The role of talk in games is still little understood."); Ariel Rubinstein, *Comments on the Interpretation of Game Theory*, 59 ECONOMETRICA 909, 921 (1991) ("It is my impression that although language plays a crucial role in resolving conflicts, game theory has so far been unable to capture this role.").

where their interactions are at least partially characterized by conflicts of interest. These conversations are impoverished, we will see, because they attempt to forge only a causal connection across the verbal exchanges between rational actors, not a conceptual one. What is needed, I shall argue, is the richer sort of conversation that is idealized by law, that is, one where there is an interpenetration of concepts and commitments in the use of language between rational actors, the sort of thing we see under a truly shared or public reason.

I have organized the paper as follows. Part I introduces a simple game of coordination and shows how two fundamental concepts of game theory, namely, common knowledge of rationality and the idea of Nash equilibrium, combine to limit the players in these games to some very unsatisfactory results. I show that neither cheap talk nor costly signaling provides an obvious way around these problems; Nash talkers seem destined to replicate the difficulties of Nash thinkers in their only slightly more public interactions. Part II introduces the law's special sense of an objectively reasonable interaction. Law's reasonable thinkers, I argue, are more capable of coordinating, and law's reasonable talkers more capable of cooperating, than their Nash counterparts because, under reasonableness, they are committed to a more public conception of their conduct shaping what they do together. I finish with some brief concluding comments.

I. STRATEGIC INTERACTION IN THE THEORY OF GAMES

According to one of its most eminent practitioners, Ariel Rubinstein, game theory should be "viewed as an abstract inquiry into the concepts used in social reasoning when dealing with situations of conflict and not as an attempt to predict behavior."³ Some will be tempted to add quickly "and just as well." For, as a predictive tool, the theory of games does not appear to be all that successful. In a way that is consistent with a growing body of experimental research in behavioral economics, people in strategic situations appear to be less "rational" than game theory presupposes.⁴ This is typically taken

3. Rubinstein, *supra* note 2, at 909.

4. Some of these experimental results on game theoretic situations (*e.g.*, the prisoner's dilemma, coordination games, the ultimatum game, etc.) are nicely surveyed and analyzed in JOHN H. KAGEL & ALVIN E. ROTH, *THE HANDBOOK OF EXPERIMENTAL ECONOMICS* (1995). A more recent and more sustained analysis of behavioral game theory can be found in COLIN CAMERER, *BEHAVIORAL GAME THEORY: EXPERIMENTS IN STRATEGIC INTERACTION* (2003). For discussion of behavioral analyses relating to legal matters more specifically, see *BEHAVIORAL LAW AND ECONOMICS* (Cass R. Sunstein ed., 2000).

to be an indication, at least among behaviorists themselves, that people are either less "good" (e.g., less competent in their computational abilities to work out their most rational strategies) or less "bad" (e.g., less self-interested or less uninterested in fairness) than the theory suggests. That is, the gap between the theory's predictions and the behaviorists' empirical results is to be explained somehow on the empirical side, either as a deficiency of individual capabilities or as a difference in individual inclination from that which is conventionally supposed.

However, I want to emphasize something different in Rubinstein's comment, something that might encourage us to think that the predictive deficiencies of game theory are more usefully explored on the conceptual side than the empirical. Rubinstein suggests that the subject matter for game theory is the concepts used by agents in "social reasoning." At first blush, this is an odd phrase. We are inclined, I think, to take this to refer to *individual reasoning* in the context of a *social interaction*, that is, where more than one individual is involved and where the outcome of the social interaction reflects the accumulation of choices that different individuals make, based on their individual thinking about one another, and one another's thinking. Reasoning, after all, is something that takes place in some individual's mind and there is no group mind that could provide the locus for something different, something that we might call *social reasoning*.⁵

And yet Rubinstein, a careful thinker, uses the phrase "social reasoning" at a moment when he is contemplating a very self-conscious interpretation of the fundamentals of game theory. Thus, it is tempting to make some sense of the phrase and, more particularly, of how the concepts used in social reasoning might differ from those used in individual reasoning in a social context. Is there a form of reasoning that is somehow more joined than what is provided by a mere accumulation of different moments of individual reasoning, but which falls short of invoking the dubious metaphysics of a group

5. Philosophers are typically very careful to avoid any suggestion of a shared mental state, even when their arguments appear to bring them close to the possibility. For example, in a paper discussing "shared intention," Michael Bratman cautions that a "shared intention . . . [is not] an attitude in the mind of some superagent." See Michael E. Bratman, *Shared Intention*, 104 ETHICS 97, 99 (1993); see also John R. Searle, *Collective Intentions and Actions*, in INTENTIONS IN COMMUNICATION 401, 404 (Philip R. Cohen et al. eds., 1990) ("I find this talk [of group minds] at best mysterious and at worst incoherent. Most empirically minded philosophers think that such phenomena must reduce to individual intentionality.").

mind? I want to argue that there is, and that, somewhat ironically for Rubinstein, the apparent failures of “social reasoning” in game theory to determine reasonable solutions for even the simplest games of coordination (together with the implausibility of certain technical devices that are sometimes wheeled out to provide these solutions when they are otherwise not readily available) manifest failures of individual, not social, reasoning. The reasoning is individual because it is allowed to have its impact only causally through one individual’s rational acts on the (equally rational) acts of another individual, something (we shall see) that shapes the sort of conversation that one can have, often rendering it mere cheap talk. But true social reasoning occurs only when (for each individual) there is a joint commitment first to think and talk, and finally to act, under a shared or public conception of one’s conduct. This calls for a quite different sort of conversation, one that is not easily articulated within the conventions of game theory given the difficulties that the agent there confronts with rational commitments of any kind.⁶

We can begin to see the nature of the difficulty if we consider the following pure coordination game. Two friends, Row and Column, want to meet for lunch at some restaurant but, as each needs to run some personal errands on the way, each must travel to the restaurant alone. There are two restaurants available, *Andy’s* and *Bob’s*. Row and Column agree that *Andy’s* is a much better place to lunch together than *Bob’s*, but they also agree that it would be better to lunch together at *Bob’s* than to lunch at either place alone. Figure 1 summarizes their situation in what game theorists call *normal* or *strategic form*. Row must decide between row A (“go to *Andy’s*”) and row B (“go to *Bob’s*”), and likewise Column must decide analogously between column A and column B.⁷ The preferences each friend has for each of the four possible outcomes (reflecting the preference each has for lunch together at *Andy’s*, and their reluctance to lunch alone) are represented by the numbers in each of the four boxes, with higher numbers representing more preferred outcomes, and the first number in each pair representing the preferences of Row, the second the

6. The tension that exists between the idea of a rational commitment and the game theorist’s account of rational choice is explored in some detail in Bruce Chapman, *Legal Analysis of Economics: Solving the Problem of Rational Commitment*, 79 CHI.-KENT L. REV. 471 (2004).

7. This is to say that Column and Row must choose between one of his or her *pure* strategies. In the discussion that follows, I ignore the possibility that each friend might adopt a *mixed* strategy, *i.e.*, one that chooses according to some randomization between each of the pure strategies.

preferences of Column. The game is called a *pure* coordination game because there is complete agreement (or no conflict of interest) between the players on the ordering of the four different possible outcomes.

Figure 1:

		Column	
		A	B
Row	A	4,4	0,0
	B	0,0	1,1

A. Nash Thinkers and Common Knowledge of Rationality

Suppose first that Row and Column have made no prior arrangement as to where they should meet for lunch. Each is simply finishing off those personal errands and now wants to meet the other for lunch at one of the two restaurants. Further, each knows the facts as summarized in Figure 1. It might seem that there is a reasonably obvious thing for each to do. Since they both much prefer to lunch together at *Andy's* to lunching together at *Bob's*, then each should choose A. But, strictly speaking, that is not what the game form matrix shows. Row should indeed choose A, *but only if Column also chooses A*. Otherwise, she should choose B. For her to choose A when Column chooses B would result in one of her two *least* satisfactory outcomes. The problem, of course, is that she does not know what Column has chosen to do. Indeed, as she thinks about it a little more, she will conclude that Column will not yet have chosen a restaurant himself. For he is working out the same problem at his end and he too is stymied. Alas, until she chooses, he cannot choose either!

The point, as Rubinstein has reminded us, is not to predict that Row and Column will *in fact* be stymied. That seems unlikely; it seems much more likely that each will choose A without much trouble and meet for lunch. The point, rather, is to isolate what it is about an agent's reasoning in this situation that generates what appears to be an *artificial* problem, that is, a problem arising out of the reasoning of each agent as game theory understands it. Recall that Rubinstein characterized this reasoning as "social." How might this term be interpreted here? We have two individuals, each isolated from the

other, working out, row by row, or column by column, what each should do. This does not appear to be social at all.

But notice that there are views that Row could have about Column which would avoid any problem for Row in working out what to do. For example, if it was Row's view that Column would simply, and somewhat thoughtlessly (compared to Row), go to the restaurant where he most preferred to lunch with Row, namely *Andy's*, then Row would have no difficulty conditioning on Column's choice of A and would choose A herself. But Row gives Column more credit as a rational actor than that. Row recognizes that Column is every bit as rational as she is and, therefore, in this mirror image situation, she imagines him to be working out a mirror image problem. In this limited respect, therefore, Row thinks of her own reasoning as "social" (or at least "socialized"); it is replicated by other rational agents in the same situation, and that replication is reintroduced into her own thinking as something that she thinks about. Indeed, that is at least partially responsible for what is so paralyzing here. For when she thinks that like-minded Column is thinking about what she is thinking and, further, is thinking about what she is thinking that he is thinking, she realizes that there can, as yet, be no choice by Column (for example, to choose A) upon which she can condition her own choice (to choose A).

It is worth noticing that Row's reasoning might not be perfectly replicated by Column and, therefore, might not be "social" in this somewhat limited sense, even if the two friends, in identical situations, were like-minded thinkers or reasoners. This might happen if they began with different "priors" or assumptions about the situation they were in. For example, suppose that each friend thinks that he or she should only choose A if the other chooses A, and likewise for B. Then the two are like-minded thinkers in this respect. But suppose that, contrary to fact, Row does not think that Column is such a thinker (even though he is). Then it might be possible for the two friends to each choose A and end up at *Andy's* for lunch together. For even if Column is such a thinker (only conditioning his choice of A on Row's choice of A, and likewise for B), so long as Row *thinks* that Column is, say, the sort of thinker who chooses only to go to the restaurant in which he prefers most to have lunch together with Row, and if Column thinks that Row thinks that, then he would think (correctly) that Row (on the basis of Row's mistaken assumption about Column) would choose A and Column too could choose A. So the

important point is that there be some sort of asymmetry in the *reasoning* between the two (*i.e.*, such that the reasoning is not perfectly replicated one by the other), not that there be only one such *reasoner*. Two like-minded thinkers, or reasoners, might reason differently about their identical situations if they begin with different understandings or assumptions about the sort of situation they are in. This assumption of “common priors” figures prominently in game theoretic accounts of strategic interaction, although usually to provide solutions to problems, not, as here, to generate a problem.⁸

I said that Row’s realization that her own reasoning is *replicated* by the reasoning of Column (and, of course, vice versa) is only “partially responsible” for the problem of paralysis in the pure coordination game. It is also essential that the *reasoning* that is replicated, or that is “social” in the quite limited sense that we are now exploring, be the sort of reasoning that has been discussed so far. But what sort of reasoning is that? It is the sort of *strategic* reasoning that, for each of the two friends, conditions a choice of action on a choice of action by the other. Put differently, but equivalently for the game form matrix in Figure 1, the rationality of an action, say, the choice of a row by Row (or a column by Column), is the rationality of that choice of row (or column) *given* a certain column choice by Column (or row choice by Row). And, surely, it is tempting to ask: How could it be otherwise? Rationality for each of the two friends goes to the rationality of their individual actions, where Row chooses row by row, and Column chooses column by column. There can be no question of what it is to choose rationally in anything other than a strictly vertical (row by row within a given column) or horizontal (column by column within a given row) direction. Choices in a diagonal direction (or choices for Row outside a given column or for Column outside a given row) are simply not available *as choices* for any one individual.⁹ Thus, if rationality or reasoning is to govern choices, and choices are ultimately made by individuals, then it is natural to think that the rationality of an action is the strategic rationality of that action *given* the choice of action by the other.

8. See Robert J. Aumann, *Agreeing to Disagree*, 4 ANNALS STAT. 1236 (1976), for a defense of common priors. Common priors are needed to support the sort of consistent alignment of beliefs that is ultimately required if the idea of a Nash equilibrium is to be useful for predicting how players will play a game. For a helpful discussion, see SHAUN P. HARGREAVES HEAP & YANIS VAROUFAKIS, *GAME THEORY: A CRITICAL INTRODUCTION* 23–28, 51–79 (1995).

9. See Christopher Woodard, *Group-Based Reasons for Action*, 6 ETHICAL THEORY & MORAL PRAC. 215, 217 (2003).

However, while natural enough, strategic rationality is not the only way to think about how rationality and reasoning might govern individual choices in cases of social interaction. An alternative form of reasoning, while conceding that choices are ultimately made by individuals, does not assess the rationality of these individual actions directly, but derives their rationality from a prior assessment of the rationality of a *pattern* of actions for the group of individuals taken as a whole. In other words, the rationality of a part—some individual action—is derived from the rationality of the larger whole of which it is a part.¹⁰ Moreover, this alternative form of reasoning about individual choices, even when replicated across all the players in identical situations with identical priors, does not result in the sort of coordination problem we earlier observed of our two friends, Row and Column. Under this alternative approach, Row and Column each ask themselves, first: What is it that *we* should do? (Each answers easily: We should meet at *Andy's*.) And then second, they ask themselves: What is it that *I* do when we do that? (Just as easily each answers: I should choose A.)¹¹ Of course, the first question *does* require each of the two to make a *diagonal* comparison in Figure 1, a comparison that neither can immediately follow through on as a matter of individual choice. But the second question returns each to the issue of what action he or she should choose as an individual. In particular, each friend should choose that row or column which allows each to do his or her part in the achievement of the outcome deemed best for the group as a whole. However, while individual actions must *ultimately* follow the strict vertical and horizontal contours of the game in this way, a *prior* assessment of what is rational in individual action need not. It is a peculiar feature of the sort of reasoning that game theory contemplates of interacting agents that it simply assumes that the empirical requirements of individual action must somehow be reproduced in an individual's assessment of that action's rationality. But there is no obvious reason why our conceptual world should track

10. *Id.* at 219.

11. Michael Bacharach refers to this sort of approach as thinking in a “we” frame as opposed to the more conventional “I/he” frame that is assumed within game theory. See Michael Bacharach, “We” Equilibria: A Variable Frame Theory of Cooperation 5 (June 24, 1997) (unpublished manuscript, on file with the author). Robert Sugden's “team reasoning” has a similar structure. See Robert Sugden, *Thinking as a Team: Towards an Explanation of Nonselfish Behavior*, 10 SOC. PHIL. & PUB. POL'Y 69 (1993); Robert Sugden, *Team Preferences*, 16 ECON. & PHIL. 175 (2000). For discussion of both Bacharach and Sugden in the context of a more general discussion of categorical reasoning and rational choice, see Bruce Chapman, *Rational Choice and Categorical Reason*, 151 U. PA. L. REV. 1169, 1205–09 (2003).

what is possible in the causal world in just this way. The causal world or, more particularly, how individuals act upon, or interact with, each other in some shared physical environment, could be informed by, or track, some independent judgment that we make of our actions as a purely conceptual matter.¹²

The game theorist might object that, rather than haul out all this heavy artillery on the causal and the conceptual, only a minor addition to the theory is required to solve what is, after all, a peculiarity of the pure coordination problem.¹³ In that particular problem, the pay-offs for the individuals are the same in all the possible outcomes.¹⁴ Thus, the difficulty is not like the more usual one faced by the game theorist where there is a conflict of interests; it really is only the more special difficulty of *pure* coordination. Perhaps, for this special case, the argument might go, we should simply add some minor principle like the following: If one outcome dominates another—that is, is strictly preferred to another by every individual—then rationality requires each individual to choose the strategy that is consistent with achieving that dominant outcome rather than the strategy that is consistent with achieving the dominated outcome.¹⁵ After all, is this not the principle (and nothing more) that we are feeling when we feel it likely that Row and Column will solve their coordination problem and end up at *Andy's* together? Why would they *not* each choose A? There is no outcome that could possibly tempt either away from the strategy that the principle prescribes. Call this principle the *joint dominance principle*.¹⁶

However, this response really does underestimate what is at stake in such an ad hoc adjustment. First, it has to be re-emphasized that the temptation for either of the friends to choose B is only really removed under this principle because each can *expect* the other, now as a matter of rationality, to choose A. Otherwise, again, it is simply

12. This argument is developed more fully under the idea of “categorical reason” in Chapman, *supra* note 11, *passim*.

13. This objection is considered in Martin Hollis & Robert Sugden, *Rationality in Action*, 102 MIND 1, 11 (1993).

14. Strictly, all that is required is that the players in the game order the outcomes identically; their individual payoffs within each outcome need not be identical.

15. This principle is advanced in David Gauthier, *Coordination*, 14 DIALOGUE 195 (1975); see also DAVID M. KREPS, GAME THEORY AND ECONOMIC MODELING 31 (1990), for a cautious endorsement of a similar argument.

16. This is the sort of terminology used to capture this idea by R. DUNCAN LUCE & HOWARD RAIFFA, GAMES AND DECISIONS: INTRODUCTION AND CRITICAL SURVEY 107 (1957).

not individually rational to choose A; one should choose B. But it is difficult to know what it is in the standard account of rational choice that allows this easy addition to the theory. What is it about rationality that suddenly allows one friend to expect the other to virtually commit to choosing A under this principle?¹⁷ This suggests that solving the pure coordination problem by using this principle is a bit like putting a rabbit in the hat, pulling it out again, and expecting a round of applause. The tools really are not in place for this sort of move. Perhaps all that fundamental artillery on the causal and the conceptual really does have to be hauled out after all.

Second, the ad hoc addition of such a principle to solve the pure coordination problem is not merely ungrounded; it is also fundamentally at odds with what game theory requires elsewhere. Specifically, the addition of such a principle goes against the basis for thinking that the concept of Nash equilibrium provides the best way to think about how to play games more generally.¹⁸ To see this, consider the following variation on our lunchtime scenario. Imagine that there is a third restaurant that Row and Column might consider, *Charlie's*, which would involve each choosing action C if they were to get there together. *Charlie's* is as good a place to lunch together as *Andy's*, and better than *Bob's*. Row would be perfectly content to lunch there with Column. However, so long as Row is not lunching alone at *Bob's*—where Row tends to get accosted by a somewhat boring colleague who always lunches there—Column would most prefer to lunch alone at either *Andy's* or *Charlie's*. (Thus, the truth of the matter is that Column really does like to lunch alone, so long as he can be sure that Row, his friend, is not bothered by this other colleague. Row, on the other hand, prefers to lunch together with Column at any of the restaurants to lunching alone, with lunches at *Andy's* and *Charlie's* most preferred.) Figure 2 summarizes their new “three restaurant” problem.

17. Hollis & Sugden, *supra* note 13, at 12.

18. This point is explicitly recognized by Prashant Parikh when he introduces such a principle to solve these sorts of coordination games; see Prashant Parikh, *Communication, Meaning, and Interpretation*, 23 LINGUISTICS & PHIL. 185, 205 (2000). See also discussion *infra* note 38.

Figure 2:

		Column		
		A	B	C
Row	A	4,4	0,0	0,5
	B	0,0	1,1	0,0
	C	0,5	0,0	4,4

Now it might appear that the addition of a third restaurant severely complicates any coordination problem that Row and Column might have. One is tempted to say (again) that under the joint dominance principle they will obviously and easily eliminate B as a choice. But there will be some difficulty coordinating on either A or C as possible choices since *Andy's* and *Charlie's* are equally good as a place to lunch together. Moreover, there is now the complication that Column would prefer most to lunch alone, so long as he can be sure that Row has not gone to *Bob's*. But can he *not* be sure of that? After all, did we not just conclude that *Bob's* was obviously not the place to go for lunch? The answer, according to game theory, is "No." Indeed, it turns out that lunch together at *Bob's*, where each friend has chosen B, is the *only* rational thing that can happen. Far from complicating the coordination problem, the addition of the third restaurant has made it disappear.

To see this, one must first appreciate that the outcome where each friend chooses B is the unique Nash equilibrium in this new three restaurant game. "A Nash equilibrium is an array of strategies, one for each player, such that no player has an incentive (in terms of improving his own payoff) to deviate from his part of the strategy array."¹⁹ In Figure 1 there were two such Nash equilibria, each one representing an outcome where the two friends were lunching together at one of the two restaurants. Having found themselves together at either of these two restaurants, each friend would have no incentive to get up and leave on his or her own. This is what it means to be in Nash equilibrium. But in Figure 2 this would only be true if both friends found themselves together at *Bob's*! If the two friends were to find themselves together at either *Andy's* or *Charlie's*, they would not be in equilibrium; Column would immediately regret that choice and want to get up to go, leaving Row to lunch on her own. Of

19. KREPS, *supra* note 15, at 28.

course, if Column were actually to find himself alone at either *Andy's* or *Charlie's*, that too would not be an equilibrium; Row, eating alone at the other restaurant, would seek to join him there. Thus, since lunching *alone* at *Bob's* is also not a Nash equilibrium for either player, the only equilibrium outcome is the one where they lunch together there.

Now all this talk of equilibrium might still be puzzling. If each of the friends has only one choice to make, then what does it mean to say that, having made that choice, he or she might not be in equilibrium? Not being in equilibrium suggests that there is something else to be done. One can appreciate, perhaps, that one or the other (or both) of the friends might regret the choice that they have made (as compared to another that they could have made, and which was preferred given the choice made by the other), but how is this *ex post* regret, this feeling of "disequilibrium," operationally significant for determining a way for each to act *ex ante* if there is only one choice to make? This is where game theory's peculiar form of "social reasoning" again comes in. For, if it is assumed that each friend is rational and, further, this fact of rationality is *common knowledge* in the game, then the only outcome that can survive this assumption *ex ante* is lunch together at *Bob's*. A proposition p is common knowledge for any two players in a game if each player i knows that p , each i knows that the other knows that p , each i knows that the other knows that i knows that p , and so *ad infinitum*.²⁰ Thus, the assumption of *common knowledge of rationality* ("CKR") is the assumption that lies behind the game theorist's interest in Nash equilibrium as a way of understanding how a rational player can anticipate how other rational players will play the game and, therefore, come to understand how he should the play the game himself. Ken Binmore has referred to it as the "implicit axiom of game theory."²¹

To see how CKR works in the context of Figure 2, imagine either friend thinking of any possible outcome other than lunch together at *Bob's*. Suppose, for example, that Row thought that Column thought that Row was going to avoid *Bob's* and go to *Andy's*. Then Row would think that Column, having thought this of Row, might think to go to *Charlie's* to eat alone. This, after all, is one of *his* most preferred outcomes. Of course, if Row thought all this of Column (and of Column's thinking), then she would think to go to *Charlie's* and meet him

20. MARTIN J. OSBORNE & ARIEL RUBINSTEIN, *A COURSE IN GAME THEORY* 73 (1994).

21. Ken Binmore, *Modeling Rational Players, Part I*, 3 *ECON. & PHILO.* 179, 193 (1987).

there, thereby securing *her* most preferred outcome. And, of course, it is not a bad outcome for Column either, at least as compared to eating at *Bob's*. The problem is that this outcome cannot survive the CKR assumption. For as soon as all this thinking crosses Row's mind, then Row must also think that Column has the same thought that this is what Row might be thinking. After all, this is what the various levels of CKR permit. And so Column, thinking that Row might be thinking this, thinks to go to *Andy's*. Or so Row thinks; and so she thinks to go to *Andy's*. But, of course, now Column will think that Row thinks *that . . .*; and Row will think that Column thinks that Row thinks that . . .; and so on. So, just as a non-Nash equilibrium outcome *cannot be stable* as a resting point for any action *ex post* (in that one of the friends will want to change his or her strategy on observing the strategy of the other), so too *ex ante* a non-Nash equilibrium outcome *cannot be sensibly thought through* as a possible outcome for the friends as a matter of rational choice. Given that they are rational, and the fact that this rationality is common knowledge, the only possible outcome that these friends can thoughtfully or rationally settle on is lunch together at *Bob's*. Thus, the joint dominance principle, which would reject lunch together at *Bob's* as a possible solution for the game, is in tension with the only solution that the friends could come up with as Nash thinkers aware of each other's rationality.

B. Nash Talkers

1. Cheap Talk

Someone might still object that our problem is very artificial. In all likelihood the two friends will have talked about having lunch together and, as a result, will have agreed to lunch together at, say, *Andy's*. Certainly in the two restaurant problem represented in Figure 1 this seems to be the likely outcome of any prior conversation that might have occurred. ("So, where's it to be?" "*Andy's*?" "Sure.") Further, since such an outcome is so predictable of any actual conversation, then our two friends, each thinking about what they *would have agreed to* had they thought to have this conversation in advance (even though they did not), would think that they would have agreed to go to *Andy's* and would, therefore, each go there. After all, if an actual conversation can bring about this result, then surely each player thinking about such a hypothetical conversation with such a predictable content, can help to bring it about too. What can be so

special about actually *having* such a predictable conversation? It is not as if actually having the conversation generates any *information* that cannot as easily and certainly be predicted from merely imagining it. After all, on this issue between these two friends, what other conversation (what alternative information) could there possibly be?

Interestingly, the game theorist is inclined to accept this argument that the scenario without an actual conversation cannot be all that different from the scenario where an actual conversation takes place. However, the collapse is of the latter into the former rather than the other way round. That is, the argument typically is that such a conversation, being *payoff irrelevant*, cannot affect the outcome of the game. Having had the conversation, each of the two friends faces the *same game* as he or she faced before the conversation and, therefore, remains just as unsure as before about what to do.²² For rational players in the conventional sense, payoff irrelevant happenings, including payoff irrelevant or “cheap” talk, cannot be equilibrium relevant. Equilibrium, as we have already seen, turns only on the assumptions that the players are rational and there is common knowledge of their rationality.

2. Signaling

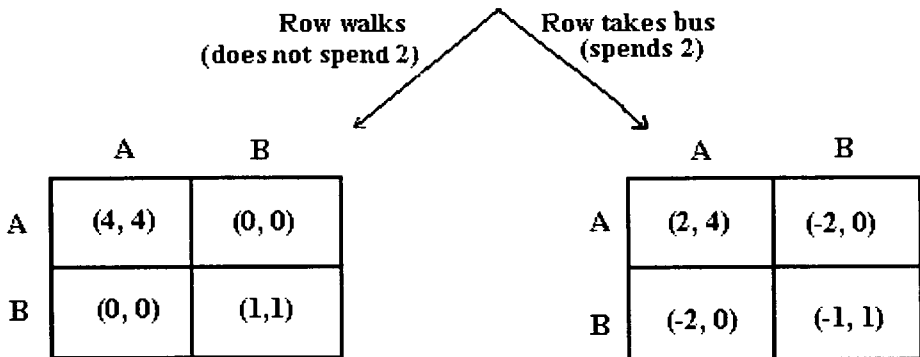
Of course, Nash talkers might have a different sort of “conversation,” one that that *is* payoff relevant. Such costly conversation is referred to as *signaling* and, the argument goes, signaling can have an impact on how the game is played.²³ Suppose, for example, that Row, who could have walked to either restaurant without any cost, takes the bus instead and that this choice of transport is observed by Column. Further, suppose the bus costs two dollars on the scale of what, for purposes of comparability, we now assume are dollar payoffs represented in Figure 1. We might imagine that the extended game (*i.e.*, an extension of the game shown in Figure 1) between Row and Col-

22. See Karl Warneryd, *Cheap Talk, Coordination, and Evolutionary Stability*, 5 GAMES & ECON. BEHAVIOR 532, 533 (1993) (“If messages are costless, it seems that they should not affect the solution of the underlying game. To rational players in the standard sense, payoff-irrelevant happenings cannot be equilibrium relevant.”); Farrell, *supra* note 2, at 210 (“The problem is that we have no link between words and actions. As a result, since only actions count, talk does nothing.”). Despite these cautionary introductory comments, in each of their papers, Warneryd and Farrell go on to provide analyses of how cheap talk might be equilibrium relevant, at least for games where the interests of players completely coincide.

23. What follows exemplifies what game theorists call “forward induction.” See Elon Kohlberg & Jean-Francois Mertens, *On the Strategic Stability of Equilibria*, 54 ECONOMETRICA 1003 (1986).

umn has the form shown in Figure 3, where Row now has the prior option either to spend two dollars to go by bus to one or the other of the two restaurants (*i.e.*, go right in Figure 3) or to spend nothing and go on foot (*i.e.*, go left in Figure 3). The net effects of having spent the two dollars to get to the different restaurants by bus are reflected in the fact that Row's payoffs in each of the four cells of the game are reduced by two on the right hand side of Figure 3.

Figure 3:



Why would Row choose to spend two dollars to go to the restaurants by bus if she could go there on foot without incurring any cost? The answer lies in the signal that such a choice sends to Column so long as Column can observe this choice. The only way that Column can make sense of Row's choice to spend the two dollars (going right in Figure 3) is that Row must be intending to go to *Andy's* rather than *Bob's*. After all, without spending two dollars (by going left in Figure 3), Row could simply choose one or the other of the two restaurants and never do worse than getting zero dollars in return, a payoff that is better than any outcome in the matrix on the right *except* the cell that

has both friends ending up at *Andy's*.²⁴ So if she chose to spend the two dollars to take the bus, it must be that she is expecting to end up at *Andy's* with Column, an expectation that requires her to go to *Andy's* herself. Thus, Column, having observed Row's choice to spend the two dollars, can now work out what he should do; he should match up with Row by also going to *Andy's*. So, somewhat oddly, where *cheap* (payoff irrelevant) talk cannot get the two friends to the restaurant together, costly signaling can. Or so the argument goes.

However, here is a wrinkle on this neat solution that arises, perhaps, from Column having one thought too many.²⁵ Suppose that Column observes that Row does *not* spend the two dollars to take the bus even though she could have chosen to do so. What is Column to make of that? Column might think, in accordance with the reasoning above, that had Row taken the bus she would have signaled credibly that she was going to go to *Andy's* and, therefore, that she would have achieved a net return of two dollars (four dollars for lunching at *Andy's*, less the two dollars spent on the bus). But now he observes that she has chosen *not* to take the bus. Column now thinks that this must mean that she expects to do *better* than the net return of two dollars that she would get by taking the bus, and that expectation is only consistent with her going to *Andy's*. After all, going to *Bob's* cannot secure more than one dollar for her. Thus, Column can now interpret the fact that Row chose *not* to take the bus as a signal that she will go to *Andy's* and, therefore, choose to match up by going to *Andy's* himself. Thus, both taking the bus (as opposed to not taking it) *and* not taking the bus (as opposed to taking it) seem to be credible signals that Row intends to go to *Andy's*. How can this be? Should Column conclude that whatever Row does, he knows that she is going to *Andy's*? Or should he conclude that the signals, even though costly, are as meaningless as cheap talk?

We can discern a sense of what might be wrong here if we realize that almost *any* option for Row to dispose of two dollars could serve

24. The game theorist would say that "not taking the bus, followed by going to any restaurant at all" is a strategy that *dominates* the strategy "taking the bus and going to *Bob's*" and, since rational players do not choose dominated strategies, it must be that *if* Row takes the bus she is intending not to go to *Bob's* and, therefore, to go to *Andy's*.

25. The following example is based on arguments found in Eric van Damme, *Stable Equilibria and Forward Induction*, 48 J. ECON. THEORY 476 (1989). Rubinstein, *supra* note 2, at 919, refers to van Damme's example as "one of the most intriguing . . . that I have ever seen in game theory."

as a signal in the ways described above. For example, if Row had the option at any point while on the walk to either of the two restaurants to simply discard the two dollars, and Column observed that she did (or did not) do that, then that too would be a signal of her intent to go to *Andy's*! Recall that there was no substantive reason for her to take the bus in the original case of signaling; what was important for the signaling was only the fact that taking the bus was costly, that is, that she discarded two dollars while on the bus. Of course, without mention of a bus there is no real connection between the act of disposal and getting to the restaurant. But, as already suggested, the *relevance* of that connection is doing no real work in the signaling argument. Any equally costly option along the way to the restaurant, no matter how irrelevant, would do just as well.

But all this must now seem very odd. For instance, suppose that Column observes Row walking towards the restaurants and wonders which of the two, *Andy's* or *Bob's*, he should go to in order to meet her. He watches closely and observes that at no point does she reach into her purse and discard two dollars. He reasons: "Aha! She could have discarded two dollars, but chose not to. Had she done so I would know that she would be signaling her intent to go to *Andy's* and, since I would have gone there too, she would have secured a net gain of two dollars. So it must be that by *not* discarding two dollars, she means to do better than that, something that is only consistent with her going to *Andy's*, where she secures a gain of four, and not with her going to *Bob's*, where she only gets one. Therefore, I will go to *Andy's*." Indeed, why limit Column's observations of Row in this way? Why not let Column watch (so far as he can) for Row's disposal of two dollars at *any* point prior to Row's arrival at the restaurants (for example, while Row was sitting through some general office meeting that morning or, perhaps, the morning before). *Any* such opportunity to discard, combined with a choice not to discard, would signal *Andy's* as Row's intended restaurant. Surely a credible signaling argument cannot be as crude as this.

Of course, consideration of the possibility of a credible signal from Row to Column cannot merely be something that goes on in Column's mind or, more specifically, in Column's thinking about Row's thinking. In the usual way that is invoked by the CKR assumption, Row must also be thinking that Column might be thinking that . . . , and so on. This will set some limits on the sorts of actions, or inactions, that might credibly be thought of as signals by Row to Col-

umn. It might occur to Column (who is, perhaps, a little too obsessed with the possibility of meeting Row for lunch) that Row's failure to discard two dollars that morning at some general office meeting amounts to a credible signal of her intention to go to *Andy's*. But if Row makes no such connection between lunch and her inaction at such a moment—if, indeed, such a connection *makes no sense* to her or is *inconceivable*—then whatever Column might think about the matter, there is no signal here. We might even say that this is what will set the limit on Column's "one thought too many" (which we considered above), which had him conjecturing that Row's choice *not* to take the bus (given that she could have taken it), just as much as her choice to take it, signaled that she was headed for *Andy's*. Likewise, and conversely, it would do Row no good to "signal" her intention to go to *Andy's* by disposing of two dollars unless Column also made the connection between this prior disposal and her choice of a restaurant. Without this connection, Column is as likely to think that Row has done something irrational or "crazy" (asking himself, "Why on earth did she throw her money away?") as he is that she is rationally signaling her future intentions.²⁶

Thus, the very possibility of effective signaling requires that the signaler and the signaled share a conceptual scheme that links, or makes sense of, certain actions together. Mere *payoff* relevance will not do the trick of distinguishing effective signaling from cheap talk; we also need an account of *payoff relevance*, that is, an account that links, or makes sense of linking, some prior costly action (or signal) to the payoffs (and actions) in the coordination game.²⁷ Moreover, the conceptual scheme that makes sense of these actions together must be one that is shared by all the players. It is not enough that the extra rationality that links actions (or, again, makes sense of them together, as parts of some whole) be private if signals are to communicate effectively. The reason that operates here must be a *public reason*. Of course, this is the sort of reason that is idealized by law, and so it is to the idea of public reason in law to which I now turn.

26. Cf. Rubinstein, *supra* note 2, at 920 ("Even if disposal gives information about 'the rationality of [Row],' a sensible conclusion might be that a player who throws the dollar out of the window is just 'crazy'.").

27. *Id.* ("If disposing of the dollar were a relevant consideration in the players' perception of the situation, then the result would (probably) make sense. However, I cannot believe that any reasonable person would consider a pre-game disposal of a dollar to be relevant in the analysis of [the game].").

II. PUBLIC REASON AND REASONABLE INTERACTIONS

A. *Reasonable Thinkers*

I began my discussion in the last section by focusing on Rubinstein's phrase "social reasoning." What made an individual's reasoning "social" in the game theoretic sense was the idea that each player, under CKR, could replicate his or her own reasoning in the reasoning of the other equally rational players in the game. It was this, for example, that precluded the possibility of non-Nash thoughts, for as soon as any one player entertained the thought that she should choose an out of equilibrium strategy (say, because it made her better off), then she would immediately realize that a Nash thinker under CKR would think her thinking that and (since the possibility being entertained is a non-Nash equilibrium) think to respond by choosing the action that makes him better off *given* that she is thinking that. Of course, she would think him thinking that and, therefore, immediately cease entertaining the thought that she could secure any such advantage after all. The only thoughts that can survive such social reasoning, it seems, are Nash equilibrium thoughts.

Now there is certainly a good deal of sophisticated reasoning and rationality being exercised here. The ideas that occur to rational players in a strategic game under the CKR assumption are very sophisticated ("I think that she thinks that I think . . .," etc.), and are often so complicated that it is sometimes difficult even to articulate them. But how "social" or public are they? A player's thoughts, once replicated, are "socialized" in a sense, in that they become the thoughts of all the other players at some level. But in another sense what is replicated or socialized is still only the stuff of *individual* thoughts. When I am thinking about what you are thinking, there is something *intersubjective* going on. But, as a point of emphasis, we have to realize as well that our mutual engagement is still only *intersubjective*. Our private thoughts are overlapping, perhaps, but there is no real rational *interaction*, at least if we mean by interaction the idea that we meet one another in some public space.

Of course, the game theorist will say that interaction is at the very core of the theory of games. The players interact when each chooses a strategy, say a row or column in one of our two person games, and these different choices combine, or interact, to produce a

final outcome, that is, some given cell in the normal game form matrix. What could be more interactive than that?

But my claim is that, however interactive this may be, it is not an account of *rational interaction*. For when there is interaction between players in the theory of games, it is causal rather than conceptual; each player simply determines a part of the world (a row or a column) as a matter of individual choice, and the determination of each part determines the whole as a causal matter. But there is no interaction in a shared conceptual space. So when there is interaction, it is not *rational* interaction. On the other hand, when there is something rational, even socially rational as in Rubinstein's "social reasoning" under the demands of CKR, the rational is, perhaps, intersubjective, but it is not interactive. The players think, of course, and even think through each other's private thoughts, but they never think together in some more public, or objective, conceptual space. Thus, when there is something rational, it is not interactive, and where there is something interactive, it is not rational. There is never, therefore, any rational interaction.

The difference between an intersubjective relationship (even a rationally intersubjective relationship) and a rational interaction is familiar enough to lawyers. Law typically concerns itself with a person's external relations or how that person interacts with other persons in the world. It is not concerned with a person's internal thoughts in so far as they do not impact upon another, that is, in so far as they are not acted upon. And when they *are* acted upon, what matters for law (in contrast, say, to ethics²⁸) is not what an individual might mean to do as a private matter,²⁹ but what she does as a public matter, that is, what she (actually) does under a publicly accessible, or (objectively) reasonable, understanding.

28. This is the contrast much emphasized by Kant in his *Doctrine of Right*, or that part of *The Metaphysics of Morals* that deals with his philosophy of law. According to Kant, law concerns itself only with what he calls the "universal principle of Right," or the coexistence of everyone's *freedom* in accordance with a universal law. IMMANUEL KANT, *THE METAPHYSICS OF MORALS* 231 (Mary Gregor trans., Cambridge Univ. Press 1991) (1797). "[A]nyone can be free," says Kant, "as long as I do not impair his freedom by my *external* action, even though I am quite indifferent to his freedom or would like in my heart to infringe it." *Id.* By contrast, "That I make [the universal principle of Right] my maxim to act rightly is a demand that ethics makes on me." *Id.*

29. *Cf.* OLIVER WENDELL HOLMES, *THE COMMON LAW* 87 (Mark DeWolfe Howe ed., Little, Brown and Co. 1963) (1881) (speaking of the defendant being judged according to an objective standard of liability in tort law: "[T]he courts . . . decline to take his personal equation into account.").

This is true for law in general, but is particularly important for contract and consent, where parties set out to do something together, that is, when they choose to enter into cooperative activities.³⁰ For example, whether two parties have a contract for the sale of, say, “new oats” or “old oats,”³¹ will not depend on whether there is a meeting (or overlap) of their (private) minds on this issue. Rather, the court will attend to the most plausible public understanding of the transaction and deem the contract to be for “old oats” if that is the most (objectively) reasonable meaning of its terms in the context in which contracting occurred. Indeed, even where subjectivity does seem to be important—for example, in the criminal law—what the accused will have to attend to as a subjective matter (as a matter of, say, honest belief implicating subjective states of mind like intent, knowledge, recklessness, etc.) will be a public or shared (objective) understanding of what the concept of right conduct requires. Thus, it will not be enough in the case of a sexual assault, for example, for the accused to say (even honestly), “I thought (in my mind) that *in her mind* she was consenting” if there was no reasonable, or public, manifestation of that consent.³² The accused’s appeal to his thoughts about her thoughts, although manifesting the same intersubjectivity that is characteristic of Nash-like thinking, is inadequate because for law the subject matter of his (subjective or honest) belief is insufficiently objective. What the accused needs to be able to establish is that he had an honest belief in a reasonable manifestation of her consent, that is, he needs to be thinking about her consent as a publicly comprehensible matter.³³

30. ARTHUR RIPSTEIN, *EQUALITY, RESPONSIBILITY, AND THE LAW* 201–17 (1999).

31. *Smith v. Hughes*, 6 L.R.-Q.B. 597 (1871) (a buyer refused oats that the seller had delivered on the grounds that he had meant to buy old oats, whereas the oats delivered were new). For excellent discussion of the case under the idea of a public rather than merely intersubjective understanding (“the meeting of minds” idea), see Brian Langille & Arthur Ripstein, “*Strictly Speaking – It Went Without Saying*”, 2 *LEGAL THEORY* 63, 76 (1996).

32. The facts of *Director of Public Prosecutions v. Morgan*, 2 All E.R. 347 (H.L. 1975), come to mind. After an evening out, Morgan, the accused, invited some of his friends to come back to his house and have intercourse with his wife. Somehow, contrary to the truth, he managed to convince them that she would likely feign her resistance, but that she would be willing nevertheless. One issue was whether the friends’ honest beliefs in her consent, no matter how unreasonable in the face of her manifest unwillingness, were enough to undermine the *mens rea* requirement for their assaults. The House of Lords, reversing the Court of Appeal, held that they were.

33. Sometimes this is expressed as a requirement that his belief in her consent be honest and reasonable. But that is problematic because it obscures too much the distinction that properly exists between the fault standards that are appropriate for criminal law and those that are appropriate for private law. It is more accurate to say that in criminal law the accused must have an honest belief in what is an objectively reasonable (public) manifestation of consent. Without

This means that, under law, two parties who are acting together will have the separate individual actions that make up their cooperative activity linked conceptually under some objective or public understanding. Thus, if each party is to understand what her separate obligations are, or what (in law, at least) she should do, she will have first to consult that shared understanding of *what it is* that they are doing together (“Is this a contract at all?” “Is this a contract for old oats?”), and only then ask what she should do under that shared understanding of the cooperative venture. This may seem obvious enough, but it is important to appreciate that, unlike for game theory, this *does* mean that the interaction of the parties is a rational interaction. Each party orders her individual action in the cooperative venture as a part of a conceptual whole that likewise also orders the individual action of the other party. Of course, it will be the individuals who ultimately act out their respective parts of the cooperative plan as understood in this public way, and at this point the interaction (part with part) might appear to be only an interaction between rational individuals (as in the usual game) and not a rational interaction. But it is precisely because each action (although individual in the causal sense) is ordered by a shared or public conceptual scheme (part with whole) that the interaction of these (rational) individuals is a rational interaction or, as law would articulate this idea, a *reasonable* one.

Notice how this interpretation of a rational or reasonable interaction is immediately useful for our two friends trying to get to lunch together. This, they realize, is a cooperative venture, even as they independently stare at the normal form game in Figure 1. So they think it through, roughly, as follows: First, each friend, for example Row, asks herself: What is it that *we* are doing here? Answer: We are trying to get to lunch together. Second: What should *we* do given that we are trying to do that? Answer: We should go to *Andy's*. And Third: What should *I* do when we do that? Answer: I should go to *Andy's*.³⁴ The first question is not, as it would be under the merely

that he is at least guilty of a reckless indifference to what right conduct requires as a matter of law. See Bruce Chapman, *Responsibility and Fault as Legal Concepts*, 12 KING'S COLLEGE L.J. 212 (2001); see also RIPSTEIN, *supra* note 30, at 202–14.

34. It should be noted that it is not necessary that the answers to this series of questions identify an *identical* action for each individual under some sort of shared conception of their actions together. For example, the individual members of a sports team (which, we will assume, is intent on winning some game) might ask themselves what they should do as a team against some other team (e.g., as an effective defensive strategy). Each member of the team, playing a different position, might then identify a quite different action as his or her special contribution

intersubjective requirements of game theory: What am *I* doing here?, a question answered with: I am trying to get to lunch with Column. This generates the problematic second question: What should *I* do given that I am trying to do that?, together with its problematic answer: I should go to *Andy's* if Column also goes to *Andy's*, otherwise I should go to *Bob's*. There is, of course, no third question here, only the paralysis of realizing, under common knowledge of such reasoning, that one really has no idea of what to do.

This is not to say that seeking to comprehend one's interactions as rational or reasonable interactions will always determine a course of action for an individual. Figure 2 provides a case in point. Here the addition of a third restaurant generates a true conflict of interest between the friends, and renders the first question without an obvious answer. In answer to the search for a common or public understanding of their interaction, each friend would not say, "We are trying to get to lunch together." Row might say that, but Column would not. And without an answer to that question—that is, without a shared conception of their individual actions together—it is less clear that they will be able to coordinate. It might seem unlikely that any shared understanding of their interaction would have the two friends agreeing to go to *Bob's*. Perhaps the shared conception of their interaction might be: We are trying to avoid Row having lunch alone at *Bob's*. (Recall that Row had this view because Row wanted to lunch together with Column; Column had this view because he did not want Row to be accosted by another colleague while she was lunching alone at *Bob's*. Thus, the friends had a common goal, but each had it for a different reason.³⁵) This could lead to a second question: What should we do given that we are trying to do that? One answer to this second question—an answer that focuses immediately and particularly on the prospect of Row's lunching *alone* at *Bob's*—is: We should lunch *together* at *Bob's*. Then, of course, each friend would answer the third question with: I should go to *Bob's*. But a more richly conceptualized answer (perhaps because it engages more categories of

to the sort of team play that is thereby required. Such a shared conception of their appropriate team strategy should help them to coordinate on the sports field. The same holds true for the quite different parties to a contract, each trying to work out their separate contractual obligations under the publicly comprehensible or reasonable interpretation of the terms of some contract that links them together.

35. It is arguable that if two individuals have a common preference for an outcome, but have it for reasons that are not fully compatible with each other, then on choosing to satisfy that common preference they are not acting under a common conception of what they are doing. See Bruce Chapman, *Rational Aggregation*, 1 POL., PHIL. & ECON. 337, 342-43 (2002).

thought in a more broadly categorical way in its intermediate steps towards a solution) might be: We should lunch *together* (and not necessarily at *Bob's* in particular), an answer that would likely precipitate the further question: And where together? A question, it must now be admitted, with no determinate answer since either *Andy's* or *Charlie's* would work equally well for the friends. Indeed, since this last conceptual sequence of questions has as a possible consequence (because of the indeterminacy) Row ending up alone at either *Andy's* or *Charlie's*, one might plausibly argue that this sequence will strike Row as *unthinkable* or *inconceivable* (certainly unacceptable). Therefore, the only conceptual ordering of their interaction that the friends might be able to “share” here is the one that has them both going to *Bob's*, the very outcome that is recommended by Nash equilibrium thoughts. However, this only goes to show that where there is no shared conception of the interaction (or where what is shared is conceptually very thin), that is, where we only have an interaction between friends who are rational (the stuff of game theory), but not a rational or reasonable interaction, we might be reduced to what Nash thinking prescribes.

B. Reasonable Talkers

1. Committed Talk

Consider again our two friends, now in conversation, over where to meet for lunch, *Andy's* or *Bob's*, where there is no conflict of interest between them (*i.e.*, the situation as depicted in Figure 1). The conversation they will be tempted to have as rational agents within game theory is the one that has each say to the other, “I will go to *Andy's* if you will go to *Andy's*; otherwise I will go to *Bob's*.” What else can they say? Anything else, and in particular anything stronger (such as announcing a commitment to go to *Andy's*) seems to be out of keeping with what rationality requires of them. Any such announced commitment would be “cheap talk” since going to *Andy's* is *not* something that either friend should do regardless of what the other player does. That is what the payoffs in the game tell us and cheap talk is, by definition, irrelevant to those payoffs.

So Row says to Column (of, say, the prospect of going to *Andy's*), “I will A if you will A.” And Column responds, “I will A if you will A.” But what sort of “willing” is this? Note that it is not an attempt at agreement, at least if an agreement means anything sub-

stantive. Again, it cannot be that if the conversation is payoff irrelevant. What it can be at most is an announcement of what the game form matrix, together with the CKR assumption, already makes clear, namely, that, “I will A if you will A because, given that you A, A is best (most rational) for me.” But this is exactly what makes the conversation so useless for speakers who assume CKR and already know the game form matrix in Figure 1. Having such a conversation does not seem to be much better for the friends than their imagining having it (as argued earlier), and imagining this conversation does not (indeed, cannot, at least if the game is an accurate representation of their position) seem to change the quandary they are in.

One immediately begins to wonder why two friends would begin such a conversation at all if nothing is changed on actually having it. Consider again an announcement by Row: “I will A if you will A.” Why does it sound so unresponsive for Column to reply, “And I will A if you will A”? There seems to be something here that *does* require a response from Column, something that was not here before Row made her announcement, and something that is not properly responded to, or recognized, when Column simply announces (as if nothing has happened) what we all knew before, namely, that he too “will A” if Row will A. Column seems to have missed the point somehow. While Column has a private understanding of what Row is saying, an understanding that he matches with his own vacuous response (“And I will A if you will A.”), this private understanding can make no sense of why Row bothered to say anything at all. *And yet she said it.* Can Column really, then, be party to a shared or public or reasonable understanding of what she said when she opened by saying, “I will A if you will A.”?

David Velleman has argued that Column’s proper response to Row’s “I will A if you will A.”—a response that reveals he truly understands what Row is saying to him—is “‘Then I will.’”³⁶ The word “then” in Column’s response, Velleman says, has a double aspect. It indicates that Column’s willing A is conditional on Row’s willing A conditional on Column’s willing A, in a way that is analogous to Row’s willing A being conditional on Column’s willing A. This much seems familiar. But Column’s saying “then” also says something more; it indicates that the condition for Column’s willing A has already been satisfied. In other words, Row’s announcement *does*

36. J. DAVID VELLEMAN, THE POSSIBILITY OF PRACTICAL REASON 216 (2000).

amount to some kind of change in the overall situation, and Velleman says that the change is that the condition for Column's willing A is now in place where it was not in place before Row actually spoke.

But, of course, Row has only spoken and not yet acted. So the condition for Column's willing A (which is mirrored in Row's conditioned willing of A) cannot be the *act* of Row actually choosing A in the way that game theory typically understands it. Rather, says Velleman, the condition for Column to will A, and the only condition that can make sense of Row speaking at all, is that Row *commits* to A (conditional on a like such commitment from Column), something that Row can accomplish simply by saying, "I will A if you will A." Thus, the conversation between the two is not the vacuous exchange of useless information (useless because it provides the friends with nothing they do not already know). Rather the only interpretation that can make sense of the conversation occurring at all is that it involves an exchange of commitments, the very commitments that provide the conditions for each to will A.³⁷

Moreover, it should be emphasized that this exchange of commitments is forceful enough to do some work even in the more problematic case of the three restaurants (see Figure 2) where there was a conflict of interests between the friends. In this situation, it will be recalled, there was a tension between what Nash equilibrium thinking demanded (namely, that each of the two friends head straight for *Bob's*) and what the joint dominance principle seemed to demand (namely, that *Bob's* be eliminated from consideration since both

37. Velleman's interpretation of the conversation between the players also solves the circularity problem that plagues two players who, on the more traditional interpretation, are each trying to condition their *actions* on the conditional *actions* of the other (*i.e.*, where each says to the other of any action "I will if you will."). Such a conversation leaves each of the players in limbo, waiting for the other to act and provide the condition for the action of the first. Under Velleman's interpretation, Column's response "Then I will" is equivalent to an *opening* statement by Column that "Column will if Row will if Column will," that is, an opening that commits Column to an action conditional on Row's *commitment* to the action conditional on action by Column. *See id.* This sort of conversation allows the commitments to action to interlock with one another. Thus, the conversation goes as follows:

1. Column (to Row): "Column will if Row will if Column will." (identifying the antecedent condition for Column's commitment to action).
2. Row (to Column): "Row will if Column will." (identifying the antecedent condition for Row's commitment to action and supplying the antecedent condition for Column's commitment to action; see 1).
3. Column (to Row): "Therefore, Column will." (supplying the consequent in response to Row's supply of the antecedent condition for Column's commitment to action and supplying the antecedent condition for Row's commitment to action; see 1 and 2).
4. Row (to Column): "Therefore, Row will." (supplying the consequent in response to Column's supply of the antecedent condition for Row's commitment to action; see 2 and 3).

friends preferred to eat together at either *Andy's* or *Charlie's* to eating together at *Bob's*). For game theorists, joint dominance simply does not have the power to resist Nash equilibrium thoughts, something that has to have us wondering, more generally, about how forceful joint dominance (or even cheap talk conversations addressing jointly dominant outcomes) can be. However, suppose now that our two friends, facing the three restaurant problem, had the more forceful conversation that involves an exchange of commitments in the way envisaged by Velleman. Suppose in particular that Row said, "I will A if you will A." and Column responded, "Then I will A." Now each is committed to going to *Andy's*. Is this commitment only forceful enough to do some work in the two restaurant case (Figure 1), where there is no conflict of interest between the friends, and totally empty of any force to do comparable work in the three restaurant case where there is such a conflict? That seems implausible. If the game theorist is tempted to allow some force (beyond cheap talk) to these sorts of conversation in the two restaurant case without any conflict of interest, then he should also allow such conversations to have some impact on cases where there is a conflict, and even where the effect might be to take the two friends away from a Nash equilibrium.³⁸

Of course, it might still be asked *what it is* in a conversation, now interpreted as an exchange of commitments rather than a mere exchange of (largely useless) information, that provides for this greater force. Some part of the answer, surely, is that the conversation involving commitments, unlike a conversation exchanging information without commitments, brings each of the friends under a shared conceptual scheme, which can rationally order each of their actions. As argued earlier, before the conversation takes place, the two friends in Figure 2 simply do not have a shared or public understanding of what they are doing. They know that neither wants to end up at *Bob's* having lunch together as compared to having lunch together at either *Andy's* or *Charlie's*. But, until they actually have their conversation, they cannot say that what they are doing is "trying to get to lunch

38. In this respect I differ from the approach adopted in Parikh, *supra* note 18, at 205, who would only allow principles like joint dominance and, presumably, the cooperative conversations that address such principles, to have a "second order" effect. That is, he would first determine the set of Nash equilibria, and only then would he allow these second order principles to work a selection across these equilibria. Under this approach, of course, the second order principles would only have an effect on the two restaurant problem (Figure 1), where there are two Nash equilibria, and would have no work to do on the three restaurant problem (Figure 2), where there is only one (jointly dominated) Nash equilibrium. But this seems to involve an almost slavish acquiescence to Nash thinking.

together.” For, given the conflict of interest in Figure 2, and before any conversation takes place, Column is simply *not* trying to do that at all. However, *after* the conversation takes place (a conversation which, although motivated by joint dominance concerns, actually involves an exchange of commitments), they do share this understanding of their interaction. Now they *are* trying to get to lunch together and, more particularly, they are trying to get to lunch together at *Andy's*. Thus, the same conversation that will do some work for them under a shared conception of their interaction in the two restaurant no-conflict-of-interest case (Figure 1) should also do some work for them in the three restaurant conflict-of-interest case (Figure 2). And it should do this work for the same action-linking reasons that are provided by a shared or public conception of their conduct.

2. *Reasonable Signaling: Rational Commitments Reasonably Understood*

The game theorist will likely remain skeptical. How can a commitment under a shared conception of their interaction support an action for some individual that is contrary to that individual's preference? For is that not what is required if the shared conception of one's conduct is to effect a non-Nash equilibrium in the sort of situation described by Figure 2? Surely this is the crucial difference between Figure 1 and Figure 2. Within game theory, just as for rational choice theory more generally, rationality is typically thought to preclude any choice contrary to preference.³⁹ So, if there is some sort of rationality (that orders individual action) in a shared or public reason, and in the commitments arising out of a conversation informed by such a shared or public reason, and if this form of rationality has this

39. See, e.g., KEN BINMORE, *PLAYING FAIR* 192 (1994) (“Cheap talk will never persuade people to act contrary to their own interests.”). The idea that cheap talk might be credible if it is “self-signaling” and “self-committing” is also dependent on the idea that individuals will not act contrary to their preference. See Joseph Farrell & Matthew Rabin, *Cheap Talk*, 10 J. ECON. PERSP. 103, 111–12 (1996). In the context of Figure 1, Column's statement that he was going to *Andy's* would be “self-signaling” because he wants to say that if and only if it is true. Further, the statement is “self-committing” because if the statement is believed by Row, then Column has an incentive actually to do what he says. But in the context of Figure 2, where eating at *Andy's* is not a Nash equilibrium, any such statement by Column that he is going to *Andy's* is neither self-signaling nor self-committing. Column does not want to tell Row this if it's true, nor only if it's true, and if, despite this, Column can get Row to believe his statement, he certainly has no incentive to follow through on it. That is the effect of following one's preference away from a non-Nash equilibrium. Farrell and Rabin recognize that insisting on messages being credible because they are self-committing is equivalent to insisting on a cheap talk equilibrium being limited to one of the Nash equilibria in the original game without cheap talk.

counter-preferential effect, then it must be the sort of rationality that is very different from (that is, not only additional to, but also contrary to) that which is conventionally assumed within game theory.

This conclusion about the sort of rationality that is required for a non-Nash equilibrium is correct, but it is hardly grounds for skepticism. When an individual makes a commitment, or announces an intention, to do some act *x*, it is not uncommon for the individual to anticipate that doing *x* will be contrary to her preferences when the time comes actually to do it. For example, an individual P might threaten or promise *ex ante* to do action *x* if individual Q does action *y*, knowing that if Q does *y*, it might then be pointless, even costly, for P actually to carry out the threatened or promised action *x ex post*.⁴⁰ Within game theory, and within decision theory more generally, it is sometimes thought that this renders the threat or promise non-credible (in that Q, knowing that P is rational, will not believe that P will carry out the threat or promise). Further, given that the threat or promise is non-credible (or, again, only cheap talk), the typical conclusion is that P is irrational even to bother making such a threat or promise *ex ante*.

But, of course, P has a *reason* to make such a threat or promise *ex ante*; the thought that such prior commitments might influence Q's behavior is the reason why P makes them. So her problem has the following structure: she has a reason to choose (or intend or commit) to do something that she knows (or can anticipate) she will have no reason actually to do, or, even more strongly, that she knows (or can anticipate) that she will have reason *not* to do. However, this is only a problem for her⁴¹ if rational action is construed as action only and

40. I ignore here the possibility that an individual might carry out the threat or promise, despite these short term costs, to secure a reputation for doing so and, therefore, to secure the larger longer term benefits of having such threats or promises believed in the future. For discussion of reputation in the context of repeated games, where such reputational concerns are most relevant, see DOUGLAS G. BAIRD ET AL., *GAME THEORY AND THE LAW* 159–87 (1994). While these reputational arguments are no doubt important, they do not explain the tendency, observed frequently in the sorts of experiments discussed in KAGEL & ROTH, *supra* note 4 and CAMERER, *supra* note 4, that players have to execute on their commitments even when long run reputations cannot be their concern, e.g., when their behavior is anonymous.

41. The economist refers to the problem that one might have (on balance) reasons *ex ante* to choose to do *x*, and (on balance) reasons *ex post* not to do it, as a problem of “dynamic inconsistency.” See R. H. Strotz, *Myopia and Inconsistency in Dynamic Utility Maximization*, 23 *REV. ECON. STUD.* 165 (1956); Peter J. Hammond, *Changing Tastes and Coherent Dynamic Choice*, 43 *REV. ECON. STUD.* 159 (1976). The economist typically resolves the dynamic inconsistency by arguing that a “sophisticated chooser” will recognize that any prior commitments to choose contrary to later preferences are without any effect and plan accordingly (e.g., by avoiding *ex post* temptations altogether, or by changing *ex post* payoffs, for example, by providing for enforcement devices or penalties, so that one is not tempted to defect from the *ex ante* commit-

always for (undefeated⁴²) reasons. Certainly, such a construal of rational action is common, and rational choice as choice according to preference (where preference can be a reflection of any number of different sorts of reasons) is only a special case of it. But it seems to provide for a very limited interpretation of rational action. For it seems clear that good reasoning can link particular actions together as a coherent or rational whole without transferring substantive reasons from one action to the other. Thus, it would seem that it might be rational (or “make sense”) to perform some action x because it stands in some rational relationship to another action y (which one might have some substantive reason to do) without there being any substantive or independent reason to do x . Indeed, this rational relationship between actions might even mean that it is rational to do something that one has reason not to do; and if that is so, then rational action is not action only and always for (undefeated) reasons.

An example might help to clarify the point. Suppose that I have a reason to visit Canberra and I announce my intention to do so. If I also believe that to visit Canberra, I need to fly to Sydney, then it might seem that good practical reasoning now gives me a reason to intend to fly to Sydney. In other words, the good practical reasoning that links (1), my prior intention to visit Canberra with (2), my belief that to visit Canberra I need to fly to Sydney, appears to give me a *reason* for (3), a derivative intention to fly to Sydney. But that cannot be quite right. For suppose that I have (1), the same intention to visit Canberra, and that I believe (erroneously) (2*), that to visit Canberra I need to fly to Chicago. Then the same logically sound practical reasoning seems to require me to have (3*), the derivative intention that I fly to Chicago. Certainly, the reasoning from (1) and (2*) to (3*) is no less powerful than the reasoning from (1) and (2) to (3). But, in truth, I have no *reason* to fly to Chicago at all, or at least not one in virtue of my prior intention to visit Canberra and my belief that to visit Canberra I need to fly to Chicago.

ment). For a philosopher's discussion of the economist's approach, and for an argument that a prior chooser can be more “resolute” than these arguments suggest, see EDWARD F. MCCLENNEN, *RATIONALITY AND DYNAMIC CHOICE: FOUNDATIONAL EXPLORATIONS* (1990).

42. This qualifier is necessary so as to recognize that conflicting reasons might apply to some choice between actions. The rational choice theorist is not committed to the idea that there cannot be such conflicts of reasons, only that something like the “balance of reasons,” or reasons “all things considered,” or, finally, “undefeated reasons,” is what ultimately determines the rationality of a choice.

So there must be something in the rational relationship that exists between actions (for example, those derived from practical reasoning), something that can make it rational to have the intention in (3) given the reasons that I have for my intention in (1) (and given my belief in (2)), but which does *not* have the effect of transferring reasons for the prior intention in (1) to the derivative intention in (3). What might that rational relationship be? In some important recent work, John Broome suggests a convincing answer:⁴³ (1), my prior intention to visit Canberra, together with (2), my belief that to visit Canberra I need to fly to Sydney, *normatively require* me to have (3), the derivative intention to fly to Sydney, but they do not give me a *reason* for that derivative intention. This difference between normative requirements and reasons is a logical one. Where normative requirements are *strict* and *relative*, reasons are in one sense weaker in being only *pro tanto* (not strict) and in another sense stronger in being *independent* (not relative). A normative requirement is strict in the sense that if I have the prior intention in (1) and the belief in (2), then I am strictly required to have the derivative intention in (3); if I do not, then I am not entirely as I should be. However, the normative requirement is relative in that I am *only* strictly required to have the derivative intention in (3) *if* I have the prior intention in (1) and the belief in (2). So I can give up the derivative intention in (3) (perhaps because I have an independent reason to do so) if either I repudiate the prior intention in (1) or alter the belief in (2). But reasons are not relative in this way. Rather, they detach and provide independent grounds for, say, intending (or not intending) to fly to Sydney. It is in this way that they are stronger than merely relative normative requirements; they cannot be so easily given up. However, reasons are weaker in being merely *pro tanto* not strict; I can have reasons both for and against intending to fly to Sydney and still be entirely as I should be. There is nothing like a contradiction here that needs to be resolved by giving up one of the reasons, although it might be that my

43. See John Broome, *Normative Requirements*, 12 *RATIO* 398 (1999); John Broome, *Are Intentions Reasons? And How Should We Cope with Incommensurable Values?*, in *PRACTICAL RATIONALITY AND PREFERENCE: ESSAYS FOR DAVID GAUTHIER* 98–120 (Christopher W. Morris & Arthur Ripstein eds., 2001); John Broome, *Normative Practical Reasoning* (Part I), *Supp. Vol. 75 ARISTOTELIAN SOC'Y* 175 (2001); John Broome, *Practical Reasoning*, in *REASON AND NATURE: ESSAYS IN THE THEORY OF RATIONALITY* (J. Bermudez & A. Millar eds., forthcoming); John Broome, *Reasons*, in *REASON AND VALUE: ESSAYS ON THE MORAL PHILOSOPHY OF JOSEPH RAZ* (P. Pettit & S. Scheffler eds., forthcoming). For an interpretation and application of Broome's analysis of rationality to a problem in law and economics, see Chapman, *supra* note 6.

reasons will have to be weighed against one another to determine (on balance, or all things considered) what they recommend.

If one accepts this richer conception of practical rationality, which recognizes that both reasons *and* normative requirements can inform rational action for some individual, then one can make sense of the idea that an individual, having made a commitment, or having announced a prior intention, to do some action x , will feel some rational pull (based on normative requirements) actually to do x even though there might be reasons (or preferences based on reasons) not to do x when the time comes actually to do it.⁴⁴ The importance of this point for our discussion of committed talk and rational signaling is twofold. First, and most obviously, it now appears, having had the sort of (publicly comprehensible) conversation that exchanges commitments (rather than useless information), that it can be rational for the interlocutors actually to do what they say they will do, even if doing so runs contrary to some preference. Thus, committed talk can rationally support a non-Nash equilibrium of the sort required in Figure 2 if lunch together at *Bob's* is to be avoided. Of course, without the prior talk, and the exchange of commitments that goes with it, we would only have the possibility of action contrary to reason or preference, that is, something which on its own seems quite irrational.

Second, just as the idea of a rational commitment orders different actions for an individual across time (*e.g.*, a commitment today to do something tomorrow) under a coherent scheme for that individual, so too a publicly comprehensible conversation that involves an exchange of such commitments orders different actions by different individuals at different points of time under a shared or reasonable conception of their overall interaction. It will be recalled that it was precisely this richer sense of an interpersonal, inter-temporal, reasonable interaction that was required (but not much attended to) in the rational signaling scenario discussed above, where Column (under one individually rational thought too many) seemed poised to construe any decision to dispose of two dollars, or even any decision *not* to dispose of two dollars, as a prior signal by Row that she was intending to go to *Andy's*. This ambiguity could only be resolved, it seemed, if some interpersonally significant conceptual discipline could be brought to bear upon Column's tendency to link Row's ac-

44. I have argued that the rational pull of normative requirements to do x is particularly strong against reasons not to do x when these reasons have been fully anticipated at the time the prior commitment or announcement is made. See Chapman, *supra* note 6.

tions inter-temporally—that is, where an earlier action by Row could be construed, *reasonably*, as a (costly) signal implicating her later decision to choose a particular restaurant. In other words, it was not enough, if the signaling argument was to go through successfully, that this prior conduct be *payoff* relevant rather than cheap talk in the way conventionally claimed by the game theorist. It was also crucial, if these ambiguities were to be avoided, that this prior (costly) conduct be payoff *relevant* in an interpersonally significant way. Rational signaling, in other words, had to be *reasonable* signaling.

This last point suggests that the theories of committed talk and reasonable signaling, although importantly different in many respects, nevertheless share a comparable structure. Each involves an action by some individual that, if viewed in isolation from other actions that the individual undertakes, will simply appear irrational to another party who, unreasonably or reasonably, fails to link the actions conceptually. In committed talk, what is irrational if done on its own *ex post* (e.g., expending the two dollars one has promised or threatened after the other party has completed her action) becomes reasonable when it is rationally connected (under the normative requirements of practical reasoning) in an interpersonally significant way as an *ex post* fulfillment of an *ex ante* commitment. In reasonable signaling, what is irrational if done on its own *ex ante* (e.g., spending the two dollars on the bus when one could just as easily walk) becomes reasonable when it is rationally connected (under the normative requirements of payoff relevance) in an interpersonally significant way as an *ex ante* signal of *ex post* behavior. It should not be surprising, perhaps, that game theory, armed with a reductive theory that can only accommodate a choice as rational if it accords with the demands of a particular preference, cannot account for what is reasonable or publicly significant in an action the rationality of which is to be found in its inter-temporal connection to another action. However, without such an account, game theory is unable to comprehend not only the importance of committed talk for cooperation and coordination, but also the mechanisms that drive its own theory of rational signaling.

CONCLUDING REMARKS

I introduced this Article by pointing to the special significance that conversation has within the law. An accused feels compelled to employ a lawyer as a special sort of interlocutor, not because a legal conversation exchanges information, but because the conversation

has a public significance that shapes the obligations of those who rationally interact under the shared meaning of its terms. One must be careful about what one says precisely because talk is not cheap; it has practical implications for what one must reasonably do, and for what others will reasonably expect one to do.

This insight from the law is missing in the game theorist's account of strategic interactions between rational individuals each aware of each other's rationality. It is missing first when these individuals simply think about each other and about each other's thinking—that is, before any interaction, verbal or non-verbal, takes place. Game theory's rational actors can replicate each other's reasoning, often to very sophisticated levels, but they never think together under shared concepts. The result is an apparent inability to coordinate in even the simplest of social interactions where there is no conflict of interest.

This absence of any true social reasoning, to use Rubinstein's ironic phrase, is simply reproduced when these actors go on, first, to talk to one another and, finally, to interact. Talk can help by providing information, but it cannot commit the speaker to conduct under any sort of shared or public understanding of their interaction. Certainly it cannot do so if such a commitment requires conduct contrary to an individual's own preference (as it often will when there is a conflict of interest), even if the effect of this restriction is to limit the interlocutors (and interactors) to outcomes that make each of them worse off than they could otherwise be under more committed talk.

It is difficult to see how game theory, confronted with these sorts of results, can plausibly claim to be a theory of rational interaction. Law's idea of a reasonable interaction, understood as an objectively intelligible ordering of the different acts of the parties under some shared public conception, one that will often require the parties to engage in committed talk rather than the exchange of largely useless information, seems to advance a more credible claim. Of course, this claim has received only the most general kind of support in the argument provided here; a more rigorous articulation of law's sense of a rational interaction must await another occasion. However, I hope that enough has been said about the limits of the fundamental concepts of game theory here to encourage both game theorists and legal theorists alike to explore the special contribution that law can make to the study of rational interaction.

