Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Comment—Subjective Probability and the Theory of Games: Comments on Kadane and Larkey's Paper

John C. Harsanyi,

50. VON NEUMANN, J., "On the Theory of Games of Strategy," *Contributions to the Theory of Games*, Princeton Univ. Press, Princeton, N.J., pp. 13–42.
51. WALD, A., *Statistical Decision Functions*, Wiley, New York, 1950.
52. YOUNG, ORAN R., *Bargaining: Formal Theories of Negotiation*, Univ. of Illinois Press, Chicago, Ill., 1975.
53. ZAMIR, S., "On the Relation Between Finitely and Infinitely Repeated Games with Incomplete Information," *Internat. J. Game Theory*, Vol. 1, No. 1 (1971), pp. 179–198.
54. ZELLNER, A., ed., *Bayesian Analysis in Econometrics and Statistics, Essays in Honor of Harold Jeffreys*, North-Holland, Amsterdam, 1980.

COMMENT*

MANAGEMENT SCIENCE
Vol. 28, No. 2, February 1982
Printed in U.S.A.

SUBJECTIVE PROBABILITY AND THE THEORY OF GAMES: COMMENTS ON KADANE AND LARKEY'S PAPER†

JOHN C. HARSANYI‡

The normative solution concepts of game theory try to provide a clear mathematical characterization of what it means to act rationally in a game where all players expect each other to act rationally. Kadane and Larkey reject the use of these normative solution concepts. Yet, this amounts to throwing away an important piece of information to the effect that the players are rational and expect each other to be rational. Even in situations where the players do not expect each other to act with complete rationality, normative game theory can help them heuristically to formulate reasonable expectations about the other players' behavior.
(GAME THEORY; RATIONALITY; BAYESIANISM; SUBJECTIVE PROBABILITIES)

According to some textbooks, there are two versions of Bayesian decision theory. The *subjectivist* version supposedly permits the decision-maker to choose his subjective probabilities in any arbitrary way (at least as long as they obey the addition and the multiplication laws of the probability calculus). In contrast, the *necessitarian* version uniquely specifies the subjective probabilities which a rational decision-maker can use in any given situation.

A little reflection will show that both textbook versions of Bayesian theory are empty caricatures. Admittedly, the verbal pronouncements of some Bayesian statisticians often come quite close to one or the other of these two extreme views, but I have yet to see a working statistician whose statistical practice has actually been governed by either view. When confronted with real-life statistical problems, all competent Bayesian statisticians will recognize that in some situations there is only *one* rational prior distribution, whereas in other situations we have a more or less free choice among many *alternative* priors.

Even Leonard Savage, whose views come closest to extreme subjectivism among distinguished Bayesian statisticians, has always admitted that in some situations (viz., those involving random devices with suitable physical symmetries, such as fair coins or fair dice, etc.) *all* reasonable people will use the *same* (rectangular) probability distributions. More recently, many Bayesian physicists have argued convincingly that in thermodynamics and in some other branches of physics one can derive the classical

*This Comment has been refereed.
†Accepted by Ambar G. Rao; received June 1981.
‡University of California, Berkeley.

0025-1909/82/2802/0 $01.25

prior probability distributions in a uniform and systematic way by using the maximum entropy principle in conjunction with certain physical invariance principles (cf. especially, the brilliant work of Jaynes, [3], [4]).

The basic reason why many physicists now favor this approach is simply that it *does the job*: it permits derivation of the usual theorems of thermodynamics in a particularly simple and elegant way and also permits their generalization to nonequilibrium situations when alternative approaches fail. Would Kadane and Larkey propose that our physicist colleagues should give up this extremely fruitful approach merely in order to conform to some arbitrary dogmatic subjectivist orthodoxy for no good reasons whatever?

From a Bayesian point of view, rational strategy choice by a given player in any game always amounts to choosing a strategy maximizing his expected payoff (expected utility) in terms of a subjective probability distribution over the strategy combinations available to the other players. But this immediately poses the question of *how this probability distribution is to be chosen* by a rational player—more specifically, how this distribution is to be chosen by a rational player who expects the *other players to act rationally*, and also expects these other players to entertain similar expectations about him and about each other.

Most game theorists answer this question by constructing various normative "solution concepts" based on suitable rationality postulates and by assuming that the players will act, and will also expect each other to act, in accordance with the relevant solution concept. In contrast, Kadane and Larkey oppose any use of normative solution concepts and oppose imposing any rationality criteria on the players' choice of subjective probabilities. They do not seem to realize that their approach would amount to *throwing away essential information*, viz., the assumption (even in cases where this is a realistic assumption) that the players will act rationally and will also *expect* each other to act rationally. Indeed, their approach would trivialize game theory by depriving it of its most interesting problem, that of how to translate the intuitive assumption of mutually expected rationality into mathematically precise behavioral terms (solution concepts).

For example, what makes von Neumann's theory of two-person zero-sum games so important and interesting is precisely the very specific solution concept it yields: it is the fact that his theory convincingly demonstrates that in a two-person zero-sum game *against a rational opponent* the only sensible policy is to use a maximin strategy. For example, consider the following game:

	X	Y	Z
A	$3, -3$	$-7, 7$	$-5, 5$
B	$8, -8$	$5, -5$	$-1, 1$
C	$10, -10$	$-8, 8$	$-6, 6$

In this game, player 1 can easily see that player 2 by using his minimax strategy Z can always ensure that player 2's payoff will not fall below 1 unit and that, therefore, player 1's payoff will not rise above -1 unit. Hence, player 1 will know that, *if player* 2 *acts rationally*, player 1 cannot expect to obtain a payoff *above* -1. Consequently, player 1 must concentrate on ensuring that his payoff will at least not fall *below* -1. Yet, the only way he can ensure this is by using his *maximin* strategy B. Therefore, the only rational thing for him to do is to use this particular strategy.

All this argument assumes is that it is rational for player 2 to ensure that he will win *at least* $1, if he cannot reasonably expect to win more than that, and that it is rational for player 1 to ensure that he will lose *at most* $1, if he cannot reasonably expect to lose less than that.

Kadane and Larkey have obviously never understood the logical force of this argument basic to von Neumann's theory because they argue that use of a maximin strategy is rational only if our opponent has *committed* himself specifically to use a minimax strategy—as if his being committed simply to maximizing his payoff from the game were not sufficient.

What could Kadane and Larkey put in place of von Neumann's theory? Only the highly uninformative statement that in a two-person zero-sum game, just as in any other game, each player should try to maximize his expected payoff in terms of his subjective probabilities. If we are not told how he should choose his subjective probabilities, this statement amounts to no more than saying that he should do whatever he thinks is best—without telling him in any way what he *should* think was best for him to do.

To be sure, von Neumann's argument is restricted to the case where the two players *expect each other to act rationally*. For instance, if player 1 thinks that player 2 is foolish enough to use strategy X (or even if he assigns a high-enough probability to this possibility), it will become rational for him to employ strategy C—in the hope of obtaining a payoff of $10.

In contrast, in deciding on the best strategy against an actually or potentially irrational opponent or opponents, normative game theory can provide only indirect help. Rather, what we need is an empirically supported *psychological* theory making at least probabilistic predictions about the strategies people are likely to use and, in particular, about the strategic *mistakes* they are likely to make, given the nature of the game and given their own psychological makeup. If we had such a theory, deciding on the best strategy against such an opponent would not involve any game-theoretical analysis, but would rather involve merely a solution of a simple maximization (expected-utility maximization) problem.

This is not to imply that *normative* game theory cannot provide very valuable heuristic help in developing such a *psychological* theory of actual—often quite error-prone—human behavior in game situations. To the contrary, normative theories of rational behavior can often suggest very fruitful hypotheses to empirical psychology because irrational actions can often best be interpreted as psychologically understandable *deviations* from the normative standards of rationality. For example, a psychologist trying to predict actual computation behavior will certainly benefit by knowing arithmetic, which can be regarded as the *normative* theory of correct computation, because it is easier to explain computation errors and to tell exactly what "has gone wrong" in any given case if you know the correct answers and know the correct arithmetic procedures yielding these answers.

In the same way, experience shows that it is much easier to explain people's—often mistaken—moves in experimental games if you know normative game theory than if you do not. But this does not change the fact that *normative* game theory and *psychological*—explanatory and predictive—theories of actual game-playing behavior are very different intellectual enterprises, using very different methodologies as a matter of logical necessity.

Apart from objecting to game-theoretical solution concepts, Kadane and Larkey also object to the use of probabilistic models for the mathematical representation of games with incomplete information (as I proposed in Harsanyi, [2]). But their discussion shows that their objections are based on a complete failure to understand the purpose of these models. They seem to think that the "basic probability distribution" used in these models is meant to be a factor "that influences the players' priors." In actual fact, the prior distribution is merely a mathematical artifact introduced into the model so that the *subjective* probability distributions used by the players in deciding their strategies can be replaced by *objective* (conditional) probability distributions in order to obtain a game model admitting of analysis by the usual analytic methods of game theory. The logical justification for replacing the original game containing *subjective* probability distributions with a probabilistic game model involving only *objective* probability distributions is the well-known fact that any Bayesian decision-maker (or player) will always act exactly the same way, regardless of whether he interprets the numerical probabilities he assigns to various events as *objective* probabilities corresponding to long-run frequencies or as *subjective* probabilities expressing merely his own personal beliefs.

One important advantage of this approach to games with incomplete information is that it yields probabilistic models very well suited to a study of such important game-theoretical problems as how each player can optimally *withhold* information from, or *convey* information credibly to, some of the other players in accordance with his own strategic interests in the game.

Of course, the final test for the usefulness of this approach is whether it actually *works*. Judging from the rather extensive and very successful research on incomplete-information games in the last thirteen years, I think it is reasonable to conclude that it passes this test with flying colors. Kadane and Larkey themselves list a number of papers in this area, most of which analyze infinitely many times repeated two-person zero-sum games with incomplete information, or discuss bargaining under incomplete information. In the last few years, however, this approach has been very successfully extended also to auctions and to other forms of competitive bidding. (For a survey of most of this work, see Engelbrecht-Wiggans, [1].) Not only have these various lines of research produced some very important new conceptual insights, as well as some very ingenious mathematics, but most of this work has been wholly dependent on a use of probabilistic models in analyzing incomplete-information games and could not have been done by any alternative method.

What could Kadane and Larkey propose as an alternative to this approach to games with incomplete information? All they could suggest is the very uninformative statement that in games of this kind, just as in all other games, every player should always do whatever seems best to him.

In conclusion, Kadane and Larkey have offered no real argument against the use of normative solution concepts and of the other analytical tools of game theory—except for the irrelevant observation that this practice is contrary to their own much too narrow and dogmatic interpretation of Bayesian decision theory. Historically, game theory has always identified its fundamental intellectual problem as that of finding a precise formal definition for the intuitive notion of rational behavior as applied to game situations. Trying to solve this problem, game theorists have developed many important new concepts, as well as a good deal of first-rate mathematics. Kadane and Larkey have not proposed any viable alternative to this approach. All they have proposed is to trivialize game theory by rejecting this basic intellectual problem and to replace it by the uninformative statement that every player should maximize his expected utility in terms of his subjective probabilities without giving him the slightest hint of how to choose these subjective probabilities in a rational manner.[1]

[1]The author wishes to thank the National Science Foundation for supporting this work by grant SES77-06394 to the Center for Research in Management, University of California, Berkeley.

References

1. Engelbrecht-Wiggans, R., "Auctions and Bidding: A Survey," *Management Sci.*, Vol. 26 (1980), pp. 119–142.
2. Harsanyi, J. C., "Games with Incomplete Information Played by 'Bayesian' Players," Parts I–III, *Management Sci.*, Vol. 14 (1967–1968), pp. 159–182, 320–334, and 486–502.

3. JAYNES, E. T., "Prior Probabilities," *IEEE Trans. Systems Sci. Cybernetics*, Vol. SSC-4 (1968), pp. 227–241.
4. ———, "The Well-posed Problem," in *Foundations of Statistical Inference* (edited by V. P. Godambe and D. A. Spott), 1972, pp. 342–354.
5. KADANE, J. B. AND LARKEY, P. D., "Subjective Probability and the Theory of Games," *Management Sci.*, this issue.

REPLY

MANAGEMENT SCIENCE
Vol. 28, No. 2, February 1982
Printed in U.S.A.

REPLY TO PROFESSOR HARSANYI

JOSEPH B. KADANE AND PATRICK D. LARKEY

Our differences with Professor Harsanyi are not as profound as might appear. His principal source of discomfort with our paper seems to be the indeterminancy that results from our inability to tell you what your opponent is likely to do. Our suggestion is that this is an empirical matter, and that we need studies of how different sorts of people play different sorts of games. Professor Harsanyi's position, as we understand it, is that you should assume that your opponent is "rational" and then decide what "rationality" implies for his behavior in the particular game in question, and act accordingly.

Thus we agree with Professor Harsanyi that "in deciding on the best strategy against an actually or potentially irrational opponent or opponents, normative game theory can provide only indirect help. Rather, what we need is an empirically supported *psychological* theory making at least probabilistic predictions about the strategies people are likely to use, . . . given the nature of the game and given their own psychological makeup. If we had such a theory, deciding on their best strategy against such an opponent . . . would . . . involve . . . a solution of a simple maximization . . . problem." We would add only that the empirical data cited in our paper supports the conclusion that opponents tend to be "actually or potentially irrational," and hence we attach urgency to further psychological research on actual behavior of people making decisions in game situations.

References

1. HARSANYI, J. C., "Subjective Probability and the Theory of Games: Comments on Kadane and Larkey's Paper," *Management Sci.*, this issue.
2. KADANE, J. B. AND LARKEY, P. D., "Subjective Probability and the Theory of Games," *Management Sci.*, this issue.

REJOINDER

MANAGEMENT SCIENCE
Vol. 28, No. 2, February 1982
Printed in U.S.A.

REJOINDER TO PROFESSORS KADANE AND LARKEY

JOHN C. HARSANYI

Frankly, I do not think it would serve any useful purpose to minimize the importance of our disagreement because it is about the very foundations of game

0025-1909/82/2802/0124$01.25

Published in Management Science on February 01, 1982 as DOI: 10.1287/mnsc.28.2.120.
This article has not been copyedited or formatted. The final version may differ from this version.