Econornetrica, Vol. 24, No. 2. (Apr., 1956), pp. 158-171.

THE DETERMINATION OF SUBJECTIVE CHARACTERISTIC FUNCTIONS IN GAMES WITH MISPERCEIVED PAYOFF FUNCTIONS[1]

BY R. DUNCAN LUCE AND ERNEST W. ADAMS

A basic assumption of the theory of games is that each player correctly perceives the payoff functions of the other players. This assumption seems highly unrealistic, and it is dropped in this paper and replaced by the assumption that each player has a perception of the payoff function of each of the other players; these perceptions may be incorrect. It is shown that this leads to an analogue of characteristic function theory in which each player has a subjective characteristic function which he believes represents the strength of the several coalitions. It is proposed that the coalitions be treated directly as outcomes and that individual preferences among them be ascertained. If the von Neumann utility axioms plus one other are met, then it can be argued that a simple transformation of the resulting utility function is a player's subjective characteristic function. The beginnings of an equilibrium theory are outlined for this more general model; this theory reduces to a known one when there are no misperceptions. The value of this generalized equilibrium theory is severely limited by two strong assumptions which are made.

1. INTRODUCTION

IN THE THEORY of games as formulated by von Neumann and Morgenstern [**13**] an *n-person game in normal form* consists of:

G1. a set I_n of n elements,
G2. n non-empty finite sets S_i, $i \in I_n$, and
G3. n real-valued functions M_i, $i \in I_n$, defined over the product space $S_1 \times S_2 \times \cdots \times S_n$.

The elements of I_n, which are called the *players* of the game, may always be taken to be the first n integers. The sets S_i are called the spaces of *pure strategies* and each element of S_i is called a pure strategy of player i; these sets are interpreted to represent the domains of possible choice of action to each of the players. The number $M_i(s_1, s_2, \cdots, s_n)$, where $s_i \in S_i$, is interpreted to be the payoff to player i in his units of utility when the n players have chosen the strategies $\| s_1, s_2, \cdots, s_n \|$. Each player is to choose one element from his strategy space in the absence of knowledge of the choices of the other players, and payments will be made to them according to the functions M_i.

[1] The first three sections of this paper are a considerable revision of work carried out at the Behavioral Models Project, Bureau of Applied Social Research, Columbia University, which is supported by funds extended by the Office of Naval Research. Reproduction of this article in whole or in part is permitted for any purpose of the U. S. Government. The paper may be identified as publication A189 of the Bureau of Applied Social Research. We are indebted to the referees for many valuable criticisms which resulted in a considerable reorganization of the material as well as the inclusion of the brief summary of n-person theory in Section 1.

On the basis of this structure and two principles which will be given below, one attempts to construct theories of behavior which characterize, in terms of payments to the players and their alliances with one another, the equilibrium states of the game. The two guiding principles which characterize individual behavior are, first, one of rational behavior:

P1. each player attempts to maximize his expected utility return in the sense that of two alternatives he will always choose the one with the larger expected utility;

and, second, one of information:

P2. each player is assumed to have full knowledge of the structure described by G1, G2, and G3, but no knowledge of the strategy choices of the other players.

For $n \geqslant 2$ one important possibility in a game is the formation of coalitions, and von Neumann and Morgenstern centered their theory about this phenomenon. They observed that for any coalition S, i.e., subset of I_n, the worst situation it can face when both P1 and P2 are assumed operative is for $-S$ (the complement of S with respect to I_n) also to form a coalition and for the game to reduce to a two-person game between S and $-S$. Using the unique minimax value from 2-person zero-sum theory, which we shall assume is known to the reader, they associate to each S a value $v(S)$. It can be shown that such a real-valued set function v, which is called the *characteristic function* of the game, satisfies two conditions:

C1. $v(\phi) = 0$, where ϕ denotes the empty set, and
C2. if R and S are disjoint subsets of I_n,

$$v\,(R \cup S) \geqslant v(R) + v(S).$$

Furthermore, it is impossible in general to deduce any other conditions a characteristic function must meet, for it can be shown that given any v satisfying C1 and C2 there exists an n-person game in normal form which has v as its characteristic function.

While a characteristic function can be derived from any n-person game, it is generally agreed that to interpret the theories constructed in terms of such functions an added stipulation must be made about the realization of the game. Von Neumann and Morgenstern [**13,** p. 604] phrased it as follows: the utility in terms of which the players are paid must be "substitutable and unrestrictedly transferable." There has been a tendency to abbreviate this to "transferable," which unfortunately tends to be misleading. We may state this restriction as:

R1. the utility in which a player is paid acts like money (even if it is not) in that it is infinitely divisible, freely transferable from one player to another with no restriction other than availability, and it is conserved in the sense that when a side payment is effected the total utility gained by the recipients exactly equals the amount lost by the player making the side payment.

In later publications other authors have introduced still other restrictions which exist in some, but not all, realizations of games; these can be said to be sociological assumptions in the model, whereas G1–G3 are economic and P1 and P2 psychological. We shall cite two, of which the first is:

R2. the players are not allowed, or are unable, to make any side payments.

The next restriction results from the following intuitive considerations: If we suppose the players are arranged into a system of coalitions given by some partition τ, there sometimes appear to be restrictions, which may be due to communication difficulties, economic costs, laws, etc., as to what additions or expulsions of members each of the given coalitions of τ may consider. These restrictions will be supposed to exist independent of whether or not the changes are in some sense profitable. The class of restrictions so far considered in the literature fall into the following scheme:

R3. For each possible partition τ there is assumed given a list of admissible coalition changes, such a list being denoted by $\psi(\tau)$. It is assumed that every coalition in τ is in $\psi(\tau)$. It is clear that ψ is a function with domain the partitions of I_n and range the sets of subsets of I_n. The special symbol E will be used for the function which maps every τ into the set of all subsets of I_n, i.e., the function which represents no restriction at all.

If we assume that restriction R1 is met, then it is reasonable to add up all the payments received by player i, including all side payments he receives and subtracting all payments he makes, to yield a final summary payment x_i. On the basis of principle P1, it is argued that any n-tuple of payments $\| x_1, x_2, \cdots, x_n \|$ which arises in an equilibrium state should satisfy two conditions. First, a rational player should never accept a final payment less than the amount he is certain he can obtain if all the other players form a coalition opposing him, i.e.,

I1. $x_i \geqslant v(\{i\})$, for all $i \in I_n$.

Second, it is held that if the sum of all the payments is less than the total amount that the coalition of all players can assure themselves, then each of the players could make a profit at no cost to the others, and this, it is said, is a violation of P1. So it is demanded that

I2. $\sum_{i=1}^{n} x_i = v(I_n)$.

Any n-tuple satisfying I1 and I2 is called an *imputation* of the game with characteristic function v.

It is by no means clear that the second condition can be said to follow from principle P1, for the principle is one of individual behavior and it appears that the argument leading to I2 depends upon interpersonal agreement. Shapley

[12] has discussed the effect on the von Neumann-Morgenstern theory of solutions of weakening I2 to an inequality, and in that case the differences do not appear to be too serious.

In terms of these notions we can tabulate the various equilibrium theories of n-person games. In the following table X will denote an imputation, τ a partition of the players into non-overlapping coalitions, and $\| s_i \|$ an n-tuple of strategies. In all cases principles P1 and P2 are assumed to hold:

Name of Theory	*Assumed Conditions*	*Objects in Equilibrium*	*References*
Solutions	C1, C2, R1	Sets of X's	[13]
Reasonable outcomes	C1, C2, R1	Sets of X's	[9]
ψ-stability	C1, C2, R1, R3	Pairs (X, τ)	[4, 6]
Equilibrium points	G1, G2, G3, $\tau = [\{1\}, \{2\}, \cdots, \{n\}]$	$\| s_i \|$	[10]

The only major theory omitted from this table is Shapley's notion of value [11] for games in characteristic function form, which is not an equilibrium theory but rather an a priori evaluation of the worth of the game to each player.

It should be noted that none of these theories assume the condition R2, i.e., in all of them it is assumed that the players may make side payments; however, in the theory of equilibrium points the assumption of non-cooperation, $\tau = [\{1\}, \{2\}, \cdots, \{n\}]$, implies R2. We see, therefore, that a major omission in n-person theory to date is a discussion of cooperative games in which side payments are prohibited or impossible. Such a development seems to entail major conceptual advances.

In the present paper we shall be concerned only with situations which can be represented by characteristic functions and with a generalization of such situations. In these theories it is argued that a change of the scale of measurement of utility, such as a change from dollars to cents, should not alter the strategic considerations of rational players. Also, a fixed payment to or from a player which is quite independent of the outcome of the game or of the coalition arrangements of the players should not affect the strategies chosen in the game, for these payments could be effected before the play of the game is ever begun. These observations lead one to call two characteristic functions v and v' over I_n *S-equivalent* if there exists a positive constant c and constants a_i such that

$$v(s) = cv'(S) + \sum_{i \in S} a_i .$$

The three equilibrium theories cited above which are based on characteristic functions are all invariant under S-equivalence.

Now if one supposes, as might be hoped, that the framework of n-person theory as given above is a suitable description of conflict-of-interest among n agents and if the argument that S-equivalent games are subject to the same strategic considerations is convincing, then, quite independently of any equilibrium theory we may champion, the observed behavior of the players in two S-equivalent games, with constants c and a_i relating them, should correspond. By this we mean that if an imputation $\| x_i \|$ arises in the one game, then $\| cx_i$

$+ a_i \|$ should arise in the other. Kalisch, *et al.* [**3**] performed an experiment which included two cases of two 4-person games which were described to the subjects by the money payments to each possible coalition. In these money payments, the pairs of games were S-equivalent. The resulting money imputations were averaged over eight groups and it was found that these averages did not correspond, in the sense defined above, for the S-equivalent games. As we have discussed elsewhere [**7**], it is not at all clear that one should expect S-equivalent games to lead to the same average imputations, for the dynamics of coalition formation—which need not be the same for S-equivalent games—may influence, if not determine, the probability distribution over equilibrium imputations. Nonetheless, a more subtle analysis of the data still suggests that the subjects did not accord all S-equivalent games the same strategic considerations.

Evidently, either the experiment was poorly executed and the results are meaningless or there is something wrong with n-person theory as a descriptive system. Since we do not have sufficient reasons to argue that the experiment was inadequate, we may tentatively investigate the second alternative, at least until the experiment is replicated and another result is obtained. The question we must consider, therefore, is which of the several assumptions that have been made should be altered. Usually, even when there are no data to raise doubts, the immediate, and often immoderate, attack is directed toward P1, the assumption that the players are rational in the sense of maximizing utility. This assumption is, however, the least easy to sacrifice for it is the only principle of behavior incorporated into the theory, and there is no ready substitute for it. To save the rationality assumption, one can argue that for some cases the fact that outcomes are measured in a utility which reflects individual preference patterns makes P1 tautological; however, in characteristic function theory this argument is particularly weak since to interpret the theory we are forced to suppose R1 holds, which in turn imposes quite strong requirements on the utility functions. Indeed, many will argue that R1 practically necessitates the supposition that the outcomes are in money and that utility is linear with money. The salient point is that until an adequate and mathematically usable behavioral principle is offered to replace P1, we are either obliged to abandon such theorizing as exemplified by n-person game theory or we must examine other potential sources of difficulty while retaining P1.

There are at least two other questionable assumptions which we may consider. First, one may directly question the argument that S-equivalent games should be subject to the same strategic considerations. If we suppose that the payments (or deprivations) a_i are effected independently of the outcome of the game, the net effect is to alter the total wealth of the players prior to playing the game. There seems to be some evidence that behavior in risky situations varies as the total wealth varies, a person generally being more conservative as his wealth decreases and more willing to accept greater risks as his fortune increases. If we assume that we are dealing with money, such behavior can be accounted for by a nonlinear utility function of a type described by Markowitz [**8**], but it is almost certain that such an assumption will run afoul of restriction R1. This may simply mean that the study of games in which side payments can occur is, in general, too complex to be encompassed by the characteristic function formulation and that a much more elaborate theory is needed; however, there does remain one further assumption which intuitively seems questionable. This is the principle P2 that each of the players has full knowledge of the structure of the normalized game. This means, among other things, that each player not only knows his own utility function over outcomes, which is plausible if "know" is not treated too literally, but also that he is fully informed of the payoff functions to each of the other players. It is doubtful that this is often the case even in terms of the objective payments of a game, and in terms of utility—which is a subjective concept—one can legitimately question whether it is ever so.

We propose, therefore, to begin to examine some of the effects of dropping P2; this leads us to the study of what may be described as games with misperceived payoff functions, or, as we shall say, m-games. Our principle concern is to develop a method to obtain laboratory estimates of the characteristic functions of such games. Following that we shall suggest the outline of an equilibrium theory for m-games which, however, is blemished by two strong assumptions which we are forced to make.

2. m-GAMES

By an *n-person m-game in normalized form* we mean a system consisting of:

G1. a set I_n of n elements,
G2. n non-empty finite sets S_i, $i \in I_n$, and
G3′. n^2 real-valued functions M_i^j, i, $j \in I_n$, defined over the product space $S_1 \times S_2 \times \cdots \times S_n$.

The interpretation of G1 and G2 is as before. The function M_i^j is to be interpreted as player i's perception of the payoff function player j is attempting to maximize. If for all $i, j \in I_n$, $M_i^j = M_j^j$, then every player correctly perceives every other player's payoff function and the m-game is simply a game in the sense of von Neumann and Morgenstern.

With this as our interpretation we may retain principle P1, but we shall have to replace P2 by:

P2′. Player i is assumed to know only the structure described by G1, G2, and the functions M_i^j, $j \in I_n$.

It is clear that there are other possible forms of misperception which we might have assumed. In some situations we may expect the players to be imperfectly informed as to the actual strategy spaces of the other players; for example, this is one way to interpret the role of secrecy about technical developments both in the military and in industry. We do not, however, propose to investigate this or any form of misperception other than that introduced through G3′ and P2′.

Observe that the n functions M_i^i form a game which we may term the *objective* *game* under consideration, and it has associated with it an (objective) characteristic function v. Similarly, for each $i \in I_n$, the n functions M_i^j also form a game —the game player i perceives to exist. Associated with this game there is also a characteristic function, v_i, which we shall call the *subjective characteristic function* of player i. It is clear that if an m-game is a game (in the sense of von Neumann and Morgenstern), then $v_i = v$ for all i.

From P2′ it follows that each player i can only be aware of his own characteristic function v_i, which he supposes is the characteristic function of the game, and so he may be assumed to base his behavior on it. It is, therefore, conceivable that players may react subjectively so differently to two given S-equivalent objective games that the resulting outcomes will not correspond nor will they be directly predictable from the given characteristic functions. Thus, the generalization of a game which we have offered may be sufficiently general to encompass the experimental phenomena mentioned in the first section. But, as we pointed out, there were a number of other ways we might have explained away the unpleasant data, and so it is hardly sufficient to offer an alternative which may work. Rather, we must establish that there are experimental methods which can detect any misperceptions which exist so that an empirical check can be made as to whether these account for the observed data.

The detection of misperceptions is, in a sense, no conceptual problem; all one need do is ascertain for each subject not just his own preference pattern and utility function over outcomes, but also his estimate of the preference patterns for each of the other players. From these data we can ascertain whether we have an m-game or simply a game. In either case we can solve all the necessary 2-person games to determine v and the v_i, and then, in terms of some equilibrium theory, we can discuss the observed behavior.

In principle all this is possible, but in practice there are almost insurmountable obstacles. It is hard enough to obtain usable reports from a subject about his own preferences to suggest that it will be next to impossible to obtain his views as to the preferences of others. Assuming, however, that these data have been obtained, the problem of translating them into utility functions requires either a finitistic axiom system such as presented by Davidson and Suppes [**1**] or certain approximations if the von Neumann and Morgenstern construction is used [**13**]; in either case the task is laborious. Now assuming the utility functions are known, the characteristic functions must be calculated, and this involves in each case the solution of 2^n two-person games. If the strategy spaces are at all rich in possibilities, these calculations are enormously complicated. All this is true for ordinary game theory and it is multiplied by a factor of n for m-games. In fact, so far as we know, it has never been carried out for any ordinary game which is sufficiently rich in strategies to be used in an experiment. It is true, on the other hand, for some ordinary games, such as the simple ones, that it is possible to arrive at the characteristic function directly without passing through the normal form of the game. There is an example of this beginning on page 564 of [**13**]. It is extremely doubtful that such an inspection technique can ever be used for m-games for the very reason that each player has a subjective characteristic function. Thus, we feel compelled to conclude that unless some method can be devised to bypass or to simplify these practical difficulties, we are left with a generalization which for all practical purposes is completely empty. The purpose of the next section is to offer what we feel may be such a method.

3. A PROCEDURE FOR DETERMINING SUBJECTIVE CHARACTERISTIC FUNCTIONS

The idea behind the procedure we shall propose is very simple: treat the coalitions of players as if they were outcomes and find the preferences of each player among probability mixtures of coalitions. If certain axioms are met, then a utility function exists which, when restricted to pure coalitions, is a characteristic function. For players who are rational, but with misperceptions, we shall present a result which argues strongly for taking the function so determined to be a player's subjective characteristic function.

Let A be any set of alternatives. A may be extended to what is called a mixture space of alternatives by considering all probability mixtures of elements from A: If α is a real number, $0 \leqslant \alpha \leqslant 1$, then if a and b are in A, the alternative "a with probability α and b with probability $1 - \alpha$" is an element of M which is denoted by $a\alpha b$. We shall suppose M is closed under this operation, i.e., if a and $b \in M$, then $a\alpha b \in M$. Hausner [**2**] has given an abstract characterization of a mixture space which is worth repeating: A set M is a *mixture space* if

M1. if $a, b \in M$ and α is real, $0 \leqslant \alpha \leqslant 1$, then $a\alpha b \in M$,
M2. $a\alpha b = b(1 - \alpha)a$,
M3. $a\alpha(b\beta c) = \left(a \dfrac{\alpha}{\alpha + \beta - \alpha\beta} b\right)(\alpha + \beta - \alpha\beta)c$, if $\alpha + \beta - \alpha\beta \neq 0$,
M4. $a\alpha a = a$,
M5. if $a\alpha c = b\alpha c$ for some α, $0 < \alpha \leqslant 1$, then $a = b$.

Now, let us suppose that a preference relation $\gtrsim$ is given over the mixture space. The following axioms are equivalent to those given by von Neumann and Morgenstern [**13**]:

W1. the relation $\gtrsim$ is a weak ordering of M,
W2. if $a \gtrsim b$, then $a\alpha c \gtrsim b\alpha c$ for $0 < \alpha < 1$,
W3. if $a\alpha c \gtrsim b\alpha c$ for some α such that $0 < \alpha < 1$, then $a \gtrsim b$,
W4. if $a > b > c$, then there exists an α such that $a\alpha c \sim b$.

When these axioms hold, it is known that there exists a non-empty family $U(\gtrsim)$ of real-valued functions over M, called *utility functions*, such that for each $u \in U(\gtrsim)$, and $a, b \in M$, and $0 \leqslant \alpha \leqslant 1$, the following conditions are met:

U1. $a \gtrsim b$ if and only if $u(a) \geqslant u(b)$,
U2. $u(a\alpha b) = \alpha u(a) + (1 - \alpha)u(b)$,
U3. $u' \in U(\gtrsim)$ if and only if u' is a linear transformation of u.

With this material as background, let us suppose that we have a situation which for one reason or another we think might be an m-game. Under some assumptions to be stated, let us suppose that from each of the players we obtain a report as to his preferences between each possible pair of risk situations involving coalitions. That is to say, we shall take A to consist of a set of possible coalitions and the mixture space M will consist of elements of the form "coalition R with probability α and coalition S with probability $1 - \alpha$." The assumptions under which the player is asked to operate when giving his reports are two:

(i) if he chooses the coalition R, then he can expect to receive the average value of payments to players in R, where the game is played between R and $-R$; the empty set ϕ is taken to mean non-participation;

(ii) if he chooses a mixture $R\alpha S \in M$, then with probability α he will receive the payment expected according to assumption (i) from a choice of R and with probability $1 - \alpha$ he will receive the payment expected according to assumption (i) from a choice of S.

Let $\gtrsim$ denote the resulting preference relation.

If one examines the axioms W1–W4 it does not seem unreasonable to suppose that for rational players $\gtrsim$ should meet these axioms. We acknowledge that it is unreasonable to expect people to satisfy the axioms perfectly, but one may hope that in some cases they will be approximately consistent, or in other words, that our model of a player's subjective evaluation of coalition strength is approximately correct. If this is the case, then there exists the class $U(\gtrsim)$ of utility functions satisfying U1–U3.

Further consideration of the fact that we are dealing with coalitions, not arbitrary outcomes, suggests that a further axiom should be met. On the grounds that the coalition $R \cup S$, where R and S are disjoint, can do everything R and S can do separately, and possibly more, it seems plausible that $R \cup S$ should be preferred to a mixture of R and S where the probability weighting of R is proportional to its size. We are therefore led to impose:

W5. if $R, S \in A$ and $R \cap S = \phi$, then $R \cup S \gtrsim R(|R|/|R \cup S|)S$, where $|R|$ denotes the number of elements in R.

If $u \in U(\gtrsim)$, it follows immediately from W5 that

$$u(R \cup S) \geqslant \frac{|R|}{|R \cup S|} u(R) + \frac{|S|}{|R \cup S|} u(S).$$

Let $\gtrsim$ satisfy axioms W1–W5 and let $u \in U(\gtrsim)$. We define $C(u)$ to be the class of set functions

$$v(R) = c|R|\,[u(R) - u(\phi)] + \sum_{i \in R} a_i,$$

where c is any positive constant and the a_i's are constants. It is a matter of routine verification to show the following four theorems:

1. If $v \in C(u)$, then v is a characteristic function.
2. If $v \in C(u)$, then $v' \in C(u)$ if and only if v' is S-equivalent to v.
3. If $u, u' \in U(\gtrsim)$, then $C(u) = C(u')$.

4. If v is a given characteristic function which is extended to the mixture space of coalitions by the definition $v(R\alpha S) = \alpha v(R) + (1 - \alpha)v(S)$ and if v is used rationally in answering questions (i) and (ii) in the sense that $R \gtrsim S$ if and only if $\frac{v(R)}{|R|} \geqslant \frac{v(S)}{|S|}$, then $\gtrsim$ satisfies axioms W1–W5 and $v \in C(u)$, where $u \in U(\gtrsim)$.

These results say, in effect, two things: first, if the preference relation over the mixture space of coalitions meets the given axioms, then the class C of functions is an S-equivalence class of characteristic functions, and, second, if a player has a subjective characteristic function v and if he bases his preferences among coalitions rationally on v, then the preference relation will meet the axioms and each of the resulting characteristic functions will be S-equivalent to v.

Thus, we feel that we may conclude that if a game is given and if the players are rational, then, without ever determining the normal form of the game, the above procedure will lead to the characteristic function of the game. On the other hand, if a situation is given which is thought to be an m-game, then the procedure should lead to the relevant subjective characteristic functions. It is of course, in the latter context where applications of the method may be expected.

The question can be raised as to how one should know when he is dealing with an m-game and so when to use the given procedure. There is not, we think, any simple rule to determine this other than intuition or the actual determination of the normal form, which we wish to avoid. At the present juncture we do not envisage any attempt to use such methods for other than laboratory experiments of the type studied by Kalisch *et al.* Should such an experiment prove successful (and because of the improved laboratory and theoretical methods being developed in utility theory the chances are gradually increasing that it might), one may then be emboldened to attempt some sort of limited field study on either a restricted economic or sociological conflict-of-interest situation. The latter program would clearly be some years in the future and would very much depend on the creation of an acceptable equilibrium theory for m-games. In the next section we suggest the outlines of one such theory which is applicable to some m-games.

4. AN EQUILIBRIUM THEORY

The purpose of this section is to sketch an equilibrium theory for m-games in characteristic function form which is a generalization of ψ-stability theory for ordinary games. We must admit right at the start that the generalization is not wholly satisfactory because we are forced to make two assumptions which surely will be met only under certain special circumstances. The first assumption amounts to a stipulation that the utility functions over coalitions are interpersonally comparable in a certain way, namely: all the players agree as to the most the situation can offer and they also agree about the strength of each individual when he is opposed by a coalition of all the other players. If the individual characteristic functions are v_i, then formally we suppose that there exist $n + 1$ constants C_i, $i \in I_n$, and C, where $C > \sum_{i=1}^{n} C_i$, such that

$$v_i(\{j\}) = C_j \quad \text{and} \quad v_i(I_n) = C.$$

Such an assumption may be expected to be met when, from the nature of the situation, the total possible payment to all the players is quite unambiguous and it is equally clear that an individual player who is opposed by a coalition of all the other players will receive a definite amount. It does not appear easy to cite other classes of m-games where the condition will be met. Incidently, we may as well suppose that these constants are $C_j = 0$ and $C = 1$, for if not the function

$$v_i'(S) = \frac{v_i(S) - \sum_{j \in S} C_j}{C - \sum_{j \in I_n} C_j}$$

is the characteristic function S-equivalent to v_i which also satisfies $v_i'(\{j\}) = 0$ and $v_i'(I_n) = 1$; it is said to be in 0, 1 normalization.

Next we must consider under what circumstances we may usefully speak of an imputation of an m-game. First, because the individual characteristic functions are all in 0, 1 normalization, there is no ambiguity in the definition of an imputation. The important question is whether we can give it any meaning; this is in doubt because we have assumed that misperceptions exist and it seems reasonable to expect that these may be misperceptions of the value of the various outcomes to the players. Thus, while each of the players may know the actual payments received, we can expect that each will have his own evaluation of the utility of these payments to the other players, and so instead of a single imputation of utilities there will be n of them. We do not have the slightest idea how one should deal with such a profusion of imputations, and so we are led to search for a condition which will collapse them into a single one. We believe that there are such m-games and that they can be dealt with experimentally. Let us suppose the actual payments to the players are in money, that they all agree upon one another's evaluation of money, and that there is free communication about the payments agreed upon; then the n imputations should all be identical. However, these assumptions appear to be so strong that they may wipe out the whole generalization, for can there still be misperceptions concerning the relative strengths of coalitions? We believe that this can still occur when the strategy spaces available to the players are sufficiently complex so that the consequences of the various choices cannot be worked out in any detail. Under these conditions people still behave, make choices, and have preferences, but since they cannot be objectively grounded they may vary considerably from person to person.

In summary, then, we shall suppose that the following two restrictions are met in any application of the definition to be given later:

R4. There are no misperceptions concerning the payments received by the individual players;

R5. There may be misperceptions concerning the relative strength of the various coalitions.

We shall also suppose that R1 (that payments are in a money-like commodity which is freely transferable and is conserved) and that R3 (that for each possible τ there are given restrictions, $\psi(\tau)$, on the coalition changes) are both met.

The equilibrium state of such an m-game will be supposed to be given by a pair (X, τ), where X is an imputation and τ a partition of the players into coalitions. If (X, τ) is in equilibrium and if S is any admissible change, i.e., $S \in \psi(\tau)$, then there must be a reason why S does not form and disrupt the equilibrium. The only motive for a change to occur, according to P1, is profit, and so we must suppose that for at least one player $i \in S$ the change does not look as if it would be profitable, i.e.,

$$v_i(S) \leqslant \sum_{j \in S} x_j .$$

Since it is sufficient for one member of a potential coalition to reject it to prevent it from forming, it is clear that it is sufficient to compare $\sum_{i \in S} x_i$ with $\theta(S) = \min_{i \in S} v_i(S)$. This will be our primary condition, but, as with the ψ-stability theory for games, we shall also stipulate that a player will not participate in a non-trivial coalition unless he receives more than he could assure himself without cooperating with anyone. Thus, we are led to the following:

Definition: Given an m-game with characteristic functions v_i in 0, 1 normalization and a function ψ of the type described in R3, a pair (X, τ), where X is an imputation and τ a partition of I_n, is said to be *ψ-stable* if

(i) for every $S \in \psi(\tau)$, $\theta(S) \leqslant \sum_{i \in S} x_i$, where $\theta(S) = \min_{i \in S} v_i(S)$, if $S \neq \phi$, and $\theta(\phi) = 0$, and

(ii) $x_i = 0$ implies $\{i\} \in \tau$.

Note that this definition reduces to the one already given for ordinary games [**4, 6**] when the m-game is actually a game, for in that case $v_i = v$, and so $\theta = v$. Thus, all we have done is formally to extend a known definition to a wider class of set functions than the characteristic functions; therefore, it would be well to characterize this class of functions. Ignoring the condition of 0, 1 normalization, which is trivial to handle, any function θ, where $\theta(S) = \min_{i \in S} v_i(S)$ and where the v_i are characteristic functions, satisfies the following two conditions:

Θ1. $0 \leqslant \theta(S) \leqslant 1$,

Θ2. $\min_{i} \max_{i \in R \subset S} \theta(R) = \theta(S)$.

The first condition follows directly from properties of characteristic functions in 0, 1 normalization. To show Θ2, we observe that since n is finite there must exist some $j \in S$ such that $v_j(S) = \min_{i \in S} v_i(S) = \theta(S)$. For any $R \subset S$ such that $j \in R$, property C2 of characteristic functions implies $v_j(R) \leqslant v_j(S)$, so

$$\theta(R) = \min_{i \in R} v_i(R) \leqslant v_j(R) \leqslant v_j(S) = \theta(S).$$

Thus, $\max_{j \epsilon R \subset S} \theta(R) \leqslant \theta(S)$. But, clearly, for any i, $\max_{i \epsilon R \subset S} \theta(R) \geqslant \theta(S)$, so we conclude $\min_i \max_{i \epsilon R \subset S} \theta(R) = \theta(S)$.

Conversely, if we have any set function θ in 0, 1 normalization which is defined over the subsets of I_n and which satisfies the above two conditions, then there exist n characteristic function v_i in 0, 1 normalization defined over I_n such that $\theta(S) = \min_{i \epsilon S} v_i(S)$. To establish this assertion, define

$$v_i(S) = \begin{cases} 0, & \text{if} \quad i \notin S, \\ \max\limits_{i \epsilon R \subset S} \theta(R), & \text{if} \quad i \in S. \end{cases}$$

It follows immediately from Θ1 that the v_i are in 0, 1 normalization. We show C1 and C2:

C1. $v_i(\phi) = 0$, since $i \notin \phi$,

C2. Let R and S be disjoint. If $i \notin R \cup S$, then $v_i(R \cup S) = 0 = v_i(R) + v_i(S)$. If $i \in R \cup S$, then with no loss of generality we may suppose $i \in R$, in which case

$$v_i(R) + v_i(S) = \max_{i \epsilon T \subset R} \theta(T) + 0 \leqslant \max_{i \epsilon T \subset R \cup S} \theta(T) = v_i(R \cup S).$$

The proof is complete with the observation that condition Θ2 yields

$$\min_{i \epsilon S} v_i(S) = \min_i \max_{i \epsilon R \subset S} \theta(R) = \theta(S).$$

It follows that the class of functions satisfying these two conditions is slightly more general than the monotonic set functions, i.e., those with the property that if $R \subset S$ then $\theta(R) \leqslant \theta(S)$, which class in turn includes all characteristic functions. One may hope that in some applications θ itself may be a characteristic function. For example, if we suppose that the most pessimistic view of the strength of any coalition is always held by a member of the coalition, i.e., for every S,

$$\min_{i \epsilon I_n} v_i(S) = \min_{i \epsilon S} v_i(S) = \theta(S),$$

then θ is a characteristic function. By definition $\theta(\phi) = 0$ and if R and S are disjoint,

$$\theta(R \cup S) = \min_{i \epsilon R \cup S} v_i(R \cup S) \geqslant \min_{i \epsilon R \cup S} [v_i(R) + v_i(S)] \geqslant \min_{i \epsilon R \cup S} v_i(R) + \min_{i \epsilon R \cup S} v_i(S)$$

$$\geqslant \min_{i \epsilon I_n} v_i(R) + \min_{i \epsilon I_n} v_i(S) = \theta(R) + \theta(S).$$

When θ is a characteristic function of one or another of several general types and the function ψ is also of a particular type, results are known as to the existence and nature of ψ-stable pairs [**4**, **5**]. No investigations have yet taken place concerning functions meeting Θ1 and Θ2 which are not characteristic functions; it would be of interest to define certain restricted classes of these functions and to investigate their stability properties, in particular to discover whether phenomena can occur which are qualitatively different from those found with characteristic functions.

In closing, we may say a few words concerning the functions ψ. These functions have been introduced into the theory with some hesitation, since it is not generally clear how they may be determined. Yet one has the intuitive feeling that in applications some such limiting boundary condition is needed. It is not impossible that the theory of m-games can explain the origin of this intuition for some situations. Consider the fact that the function θ always gives the most pessimistic evaluation of the strength of a coalition by the members of the coalition. There should be a tendency—though exceptions are easily produced—for the value of θ to be less than we would expect from a knowledge of the objective game. The exceptions occur when and only when each of the members of a coalition overestimates its strength. Now, if this depressing effect is sufficiently severe, for example if it results in $\theta(S) \leqslant |S|/n$, then even if all possible coalition changes are allowed (the function E defined in R3) there are stable pairs [4]. In other words, we may say that misperceptions in a game tend to increase the stability of the game in much the same way as limiting coalition change does, and it is conceivable that this probable effect of misperceptions is the source of our intuition that potential coalition changes are limited.

Columbia University;
Wesleyan University

REFERENCES

[1] Davidson, Donald and Patrick Suppes: *Finitistic Rational Choice Structures*, Report no. 2, Stanford Value Theory Project and Department of Philosophy, Stanford University (1955), 38 pp.

[2] Hausner, Melvin: "Multidimensional Utilities," *Decision Processes* (R. M. Thrall, C. H. Coombs, and R. L. Davis, Editors), John Wiley and Sons, New York (1954), pp. 167–180.

[3] Kalisch, G., J. W. Milnor, J. Nash, and E. D. Nering: "Some Experimental n-person Games," *Decision Processes* (R. M. Thrall, C. H. Coombs, and R. L. Davis, Editors), Wiley, New York (1954), pp. 301–327.

[4] Luce, R. D.: "A Definition of Stability of n-person games," *Annals of Mathematics*, 59 (1954), pp. 357–366.

[5] ———: "k-Stability of Symmetric and of Quota Games," *Annals of Mathematics*, 62 (1955), pp. 517–527.

[6] ———: "ψ-Stability: A New Equilibrium Concept for n-person Game Theory," *Mathematical Models of Human Behavior*, Dunlap and Associates, Inc., Stamford, Conn. (1955), pp. 32–44.

[7] ———: A note on the paper "Some Experimental n-person Games," submitted for publication.

[8] Markowitz, Harry: "The Utility of Wealth," *Journal of Political Economy*, 60 (1952), pp. 151–158.

[9] Milnor, J. W.: "Reasonable Outcomes for n-person Games," Memorandum RM-916, The RAND Corporation, Santa Monica (1952).

[10] Nash, John: "Non-cooperative Games," *Annals of Mathematics*, 54 (1951), pp. 286–295.

[11] Shapley, L. S.: "A Value for n-person Games," *Contributions to the Theory of Games, II*, (H. W. Kuhn and A. W. Tucker, Editors), Annals of Mathematics Studies, no. 28, Princeton University Press, Princeton (1953), pp. 307–317.

[12] ———: "Notes on the n-person Games, III; Some Variants of the von Neumann-Morgenstern Definition of Solution," Memorandum RM-817, The RAND Corporation, Santa Monica (1952).

[13] von Neumann, John and Oskar Morgenstern: *Theory of Games and Economic Behavior*, Princeton University Press, Princeton (1947).