CARL WAGNER

THE FORMAL FOUNDATIONS OF LEHRER'S THEORY OF CONSENSUS[1]

Introduction

The challenge of devising rational methods for achieving group consensus has provided decision theorists with an important class of practical and theoretical problems. How should group deliberation be structured in order to enhance the exchange and thoughtful consideration of relevant information and maximize the possibility of agreement? Are there defensible ways of combining the opinions of individuals who disagree even after exhaustive discussion? Can we develop a comprehensive theory of group rationality and perhaps refine our understanding through the discovery of limitative metatheorems, which have so often been the mark of mature axiomatization?

Suppose that a group wishes to arrive at a consensual preferential ordering of some set of alternatives. Each member of the group votes by reporting his own pattern of preference and indifference among these alternatives, and the problem is to devise a fair rule for combining the votes to produce a group ordering. An obvious candidate is the *majority rule*: the group prefers a to b when a majority of its members do so, and if there is no majority for a over b or b over a, the group is indifferent between these alternatives. As attractive as this rule seems, it breaks down in the face of certain preference profiles. Suppose that individuals I_1, I_2, and I_3 order alternatives a, b, and c as follows:

I_1	I_2	I_3
a	c	b
b	a	c
c	b	a

Bogdan, R. J. (ed.), 'Keith Lehrer', 165–180

Since I_1 and I_2 prefer a to b, and I_1 and I_3 prefer b to c, the group must do likewise. But I_2 and I_3 prefer c to a, and this majority preference cannot be incorporated in the group ordering since it would violate the transitivity of the preference relation. This *voter's paradox*, discovered by the eighteenth-century philosopher and social scientist Condorcet, stimulated a search for methods which would produce a group ordering from any profile of individual orderings.

Borda, also in the eighteenth century, suggested a rule based on the following method of assigning a score to each alternative: An alternative a receives as its *Borda count* $B(a)$ the sum, taken over all individuals, of the number of alternatives ranked strictly below a. Then the group ranks a over b if $B(a) > B(b)$ and is indifferent between them if $B(a) = B(b)$. This method always leads to a well-defined group ranking, but the fairness of the results is sometimes questionable, as in the case of the profile

I_1	I_2	I_3
a	a	b
b	b	c
c	c	d
d	d	a

Since $B(a) = 6$ and $B(b) = 7$ the group must rank b over a even though a majority favor a over b.

Following Condorcet and Borda many other scoring methods were proposed and, indeed, employed in practice, but none proved to be entirely invulnerable to criticism. Finally, in 1951, the economist Kenneth Arrow argued in the monograph, *Social Choice and Individual Values*, that all efforts to discover a rational, democratic, universally applicable method of amalgamating individual preferences were doomed to failure. Arrow's strategy involved, first of all, viewing amalgamation rules as abstract mappings from the set of all preference profiles to the set of all preferential orderings. All such mappings, which he called *social welfare functions*, yield universally applicable rules since their domain is required by definition to include all possible profiles.

Next Arrow argued that any rational, democratic social welfare function would necessarily satisfy four conditions, which he labeled *positive association of social and individual values*, *independence of irrelevant alternatives*, *citizens' sovereignty*, and *nondictatorship*. He then proved, in what is now known as the *Arrow Impossibility Theorem*, that any social welfare function satisfying the first three of these conditions necessarily violates the fourth.[2] In short, there are no rational, democratic social welfare functions. One can attempt to mitigate the implications of this theorem by arguing that a weaker characterization of rational democracy is acceptable, or by dropping the requirement that the domain of a social welfare function consist of all possible preference profiles. But the Impossibility Theorem remains the point of reference for such maneuvers and thus continues to exercise limitative force.

Decision problems in which an individual's input is limited to a report of his preferences among alternatives comprise an important, but by no means exhaustive, subclass of the problems encountered by decision-making groups. Problems involving measurement, resource allocation, or the estimation of probabilities, for example, yield unavoidably a vector of *numerical* individual opinions. In such cases the role of the aforementioned social welfare functions is played by functions which map a numerical vector to a single (consensual) number. In practice the familiar averaging functions of elementary statistics such as the median and the mean (both arithmetic and geometric) have been employed in this capacity, justified by a few brief remarks about their 'centralizing' effect. These simple averages have the property of giving equal weight, in some sense, to all individual opinions. Now while groups of adversaries may have no recourse but to compromise with an average of this sort, it is clear that an ideal community of disinterested truth seekers should avail itself of an average weighted to reflect the expertise of its members. Thus, for example, the arithmetic mean $(a_1 + a_2 + \cdots + a_n)/n$ of a vector of opinions should be regarded as just one of an infinite number of possible averaging policies provided by the class of *weighted arithmetic means* $w_1a_1 + w_2a_2 + \cdots + w_na_n$, where $0 \leq w_i \leq 1$ and $w_1 + w_2 + \cdots + w_n = 1$. The formal core of Lehrer's theory of consensus represents nothing less than an attempt to equip decision-making groups with a method for reaching a rational consensus as to which sequence of weights to employ in arithmetic averaging. This essay is devoted to a summary and assessment of that method.

1. Weighted Averaging.

Suppose that a decision-making group with n members wishes to determine consensual values of one or more numerical variables. Suppose further that after extensive research and discussion no one wishes to convey additional information to the group or to seek further information from any other individual. This state of the decision-making process, which Lehrer calls *dialectical equilibrium*, may be summarized by a matrix $A = (a_{ij})$ with n rows and one column for each variable, a_{ij} denoting individual i's estimate of the most appropriate value of variable j. Thus exact consensus obtains precisely when the rows of A are identical. In practice, of course, a group may be satisfied to arrive at a state which only approximates that ideal. If, for example, the differences between the largest and smallest entries in each column of A are uniformly bounded by some agreeably small positive constant, a group may reasonably terminate its deliberations by simple averaging of the columns of A. But how should a group proceed if A fails even to exhibit an acceptable approximate consensus? As disinterested, non-egotistical truth seekers, the individuals involved will not resort to coercion; only the forces of rational argument may be brought to bear on others. Thus a state of nonconsensual dialectical equilibrium presents something of an impasse, for such forces appear to have been given full play.

Lehrer's response to this dilemma has been to argue for the relevance of what he calls 'social' information and to specify a formal technique for incorporating such information into the deliberative process. He has maintained that rational decision makers need not restrict their attention to, say, the physics of the problem at hand, but may if necessary admit as data their considered perceptions of the expertise of their fellow physicists. Now critics of this view, while perhaps allowing that social psychologists might wish to document the situations in which social information carries any force,[3] might flatly assert that such considerations are clearly inimical to rational deliberation. But disciplining decision making groups with a rigorist regimen involving the analysis of exclusively hard data is both unrealistic and unnecessary. Group deliberation, after all, typically involves exchange, not only of indisputable facts and inferences, but also of interpretations, intuitions, and guesses which cannot be supported by rigorous logical or statistical arguments. Groups under no compulsion to come to an immediate decision might, of course, exclude such impressionistic data from consideration, and decline to adjourn unless enough hard data emerges to settle the issue. But those unable to conceive of (or perhaps afford) further relevant factual research may be forced to take account of softer kinds of data. Consideration of such data must, according to Lehrer, include as an essential component evaluation of the individual who advocates its cogency.

Those having no objection to such evaluations may assert, however, that they will typically already have been carried out (in perhaps unsystematic ways) during discussions preceding dialectical equilibrium, and thus offer nothing that could resolve nonconsensus. In fact, Lehrer's original scheme for modifying A on the basis of social information is vulnerable to this charge. Fortunately, there is a way of regimenting the deliberative agenda so that the normative algebra which he has devised can function without redundancy. Discussions preceding dialectical equilibrium are simply to be carried on through exchange of anonymous position papers.[4] This arrangement forces individuals to evaluate the opinions of others without taking their identities into account.[5] They will thus delay final evaluation of the more impressionistic assertions of others, having at this point incomplete criteria for such an evaluation.

If at dialectical equilibrium the matrix A exhibits an unacceptable deviation from consensus, authors of the position papers are identified, and individuals turn their attention to construction of an $n \times n$ weight matrix $W = (w_{ij})$ with nonnegative entries and all row sums equal to 1. Lehrer has characterized his approach to the meaning and determination of these weights as subjectivist. Thus it will suffice for this discussion simply to regard w_{ij} as that part of a unit vote which, after complete discussion, individual i is willing to proxy to individual j on the basis of his respect for j as an evaluator of the numerical variables of A. By contributing the row $w_{i1}, w_{i2}, \ldots, w_{in}$ to W, i commits himself to revise his entry a_{ij} of A to $a_{ij}^{(1)} = w_{i1}a_{1j} + w_{i2}a_{2j} + \cdots + w_{in}a_{nj}$. If each individual so revises the entries of his row of A, the resulting matrix $A^{(1)} = (a_{ij}^{(1)})$ is simply the matrix product WA.[6]

It is of interest at this point to remark on several simple theorems about matrix multiplication which guarantee that the shift from A to WA satisfies certain intuitive conditions of minimal rationality. One feels, first of all, that $A^{(1)} = WA$ should not register greater disagreement than A itself. This is in fact the case, for regardless of the entries of W, it is easy to check that the largest and smallest numbers in the jth column of $A^{(1)}$ never lie outside their counterparts in A. In particular, if A is already a consensus matrix, then $WA = A$ for any W, so that consensus, once attained, is undisturbed by the procedure under discussion. Finally, if W is a consensus matrix, the same is true for WA, for each individual k shifts his a_{kj} to the same weighted arithmetic mean of the numbers in the jth column of A.

In view of the above observations, it is clearly desirable for a group faced with nonconsensus in A to aim for the construction of a consensual weight matrix W. Thus W, like A, ought to represent a state of dialectical equilibrium, preceded by a full discussion. The subject of discussion is now, however, the expertise of members of the decisionmaking group, as indicated by their prior achievements and consequent reputation, as well as their recent, now identified, performance as authors of the first set of position papers. It is to be hoped that the group will arrive at a matrix of weights for which WA exhibits at least an acceptable approximate consensus.[8] It seems, indeed, that the failure of such consensus will simply create a higher level impasse. I describe in the next section a method proposed by Lehrer to resolve such a dilemma.

2. Iteration.

Suppose that the shift in group opinion from A to $A^{(1)} = WA$ does not yield an acceptable approximation of consensus. Note that the entries of $A^{(1)}$, like those of A, express individual opinions (albeit revised ones) of the most appropriate values of the initial decision variables. Lehrer now asks under what circumstances individuals would be willing to accord the same weights to these revised opinions as they granted to the original opinions, and thus shift from $A^{(1)}$ to $A^{(2)} = WA^{(1)}$. He argues that this will be the case precisely when individuals respect each other's assignments of weights in W to the same degree as their assignments of values to the decision variables in A. Thus, for example, if values of physical variables are at issue in A, the move from $A^{(1)}$ to $WA^{(1)}$ is rational if and only if all individuals are seen as possessing uniform, though perhaps individually differing, skills as physicists and as judges of physicists.

An attractive informal argument for this claim may be based on viewing the shift from A to $A^{(1)}$ under such circumstances as moving individual opinions to a new plateau, but one on which opinions have the same relative strengths as their counterparts in A. Lehrer's formal argument is as follows: Ignore temporarily the matrix A and multiply W by itself. The i-jth entry $w_{i1}w_{1j} + w_{i2}w_{2j} + \cdots + w_{in}w_{nj}$ of W^2 is a rational revision of the weight w_{ij} which i granted to j as a physicist, since it is a weighted average of the opinions $w_{1j}, w_{2j}, \ldots, w_{nj}$ on this matter, with coefficients determined by the weights $w_{i1}, w_{i2}, \ldots, w_{in}$ which i has granted to individuals as judges of physicists. Hence W^2A is a revision of A based on more information than the revision WA. But by the associativity of matrix multiplication, $W^2A = (WW)A = W(WA) = WA^{(1)} = A^{(2)}$.

The above procedure may be iterated to the extent that individuals perceive each other as possessing uniform higher level judgemental skills. Thus if individuals are regarded as being as equally skilled judges of judges of physicists as they are judges of physicists, W^2 may also be regarded as a rational revision of weights granted to individuals as judges of physicists, and $W^2W = W^3$ as a rational revision of weights granted to individuals as physicists. Hence $W^3A = WA^{(2)} = A^{(3)}$ is a rational revision of A incorporating still further information. At this stage one may perhaps begin to feel mesmerised by the staccato rhythms of 'judges of judges of judges' and assert, when one's head clears, that judgments of this increasingly complex kind cannot be meaningful. But the progressive modification of A by increasing powers of W does not really require that individuals assess their colleagues separately on an infinite hierarchy of judgmental skills. So long as individuals have no reason for regarding the opinions in any $A^{(n+1)} = WA^{(n)}$ as having different relative strengths than those in $A^{(n)}$, they may rationally shift to $WA^{(n+1)} = A^{(n+2)}$. In such cases there arises an infinite sequence of (synchronic) modifications W, W^2, W^3, etc. of W and a companion sequence A, WA, W^2A, etc. of modifications of A. The conditions under which these sequences approach consensus are discussed in the next section.

3. Convergence to Consensus.

Given a weight matrix $W = (w_{ij})$ let us say, following Lehrer, that i *respects* j if $w_{ij} > 0$, and that there is a *vector of positive respect from* i *to* j if there is some sequence of individuals, beginning with i and ending with j, such that each individual in the sequence respects the person listed directly after him in that sequence.[9] A matrix W for which there is some vector of positive respect from every individual to every other individual (and, hence, from every individual to himself) is known as an *ergodic* matrix. A *regular* matrix is an ergodic matrix for which there is some fixed positive integer r such that there is a vector of positive respect of length r from every individual to every other individual, and from every individual to himself. It may easily be proved by induction that for a given ordered pair (i, j) of distinct or identical individuals, there is a vector of positive respect of length r from i to j if and only if the i-jth entry of W^{r-1} is positive. Hence W is regular if and only if some power of W has exclusively positive entries.[10].

Ergodicity and regularity are concepts from the classical theory of

Markov chains. This theory deals abstractly with square matrices having nonnegative entries and unit row sums. Since it arose with a view toward applications to probability theory its theorems are commonly phrased in terms of probabilities rather than weights. But, recast in Lehrer's terminology, the basic results of this theory yield significant theorems about group decisionmaking and consensus. The most important of these theorems states that the increasing powers of a matrix W approach a limit consensus matrix L with exclusively positive entries if and only if W is regular.[11] Hence the notion of regularity captures precisely those patterns of respect which lead to a consensus in which each individual's initial opinions are accorded some positive weight, regardless of the initial state of disagreement.[12]

How does one compute in practice the entries of the limit consensus matrix L? It may appear that one could only find an approximation to L by raising W to some sufficiently high power, but there is in fact a simple method for computing L exactly. If one denotes by $\lambda_1, \lambda_2, \ldots, \lambda_n$ the entries of the (identical) rows of L, these numbers turn out to be the unique solution of the matrix equation $(\lambda_1, \lambda_2, \ldots, \lambda_n)\ W = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ subject to $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$.[13]

The above observation leads to an elegant characterization of the regular matrices which yield ultimate consensus in the form of a simple arithmetic mean of the values of the initial decision variables. For it is easy to check that $(\lambda_1, \lambda_2, \ldots, \lambda_n) = (1/n, 1/n, \ldots, 1/n)$ satisfies the above simultaneous equations if and only if the sum of the entries in each of the columns of W is equal to 1. Thus Lehrer's method leads decision-making groups to the intuitively reasonable policy of adopting simple averaging precisely when each individual receives proxies which sum to a full unit vote.

It should be noted in conclusion that there are regular matrices with remarkably few positive entries. An interesting class of examples, due to Lehrer, consists of the $n \times n$ matrices W in which each individual i grants himself the positive weight $1 - w_i$, and individual $i + 1$ (or 1, if $i = n$) the positive weight w_i. For such matrices the consensual limit weight received by i is given by $\lambda_i = (1/w_i)/(1/w_1 + 1/w_2 + \cdots + 1/w_n)$. Thus the limiting weights are what mathematicians would call 'weighted harmonic means' of the quantities $w_1, {}_2, \ldots, w_n$. In particular, $\lambda_i/\lambda_j = w_j/w_i$, so that, for example, if i proxies away twice the weight proxied by j, he will receive in the limit half the weight which j receives.

4. Generalizations of Lehrer's Model.

Having summarized the formal core of Lehrer's theory of consensus, I wish now to sketch some possible generalizations of his model. Suppose first that in dialectical equilibrium the weight matrix W is ergodic, but not regular. There is thus a vector of positive respect from each individual to every other individual and from each individual to himself. It seems reasonable that such situations would involve a communication of respect sufficient to lead the group to consensus. Indeed, one notes that a matrix of this type fails to be regular only for lack of fulfilling a rather technical condition on vector lengths. To be sure, this deficiency in W has important mathematical consequences – the sequence W, W^2, W^3, etc. simply does not converge if W is ergodic but not regular. But are there not other rational routes to consensus in such cases?

As mentioned in Note 11, an ergodic matrix W having at least one positive diagonal element w_{ii} is regular. Hence ergodic non-regular matrices arise only if each individual proxies away his entire unit vote to others. As infrequently as this may occur in practice it is desirable nevertheless to have a method for deriving consensus in such situations. I propose the following procedure: Let the group choose an agreeably small positive constant ϵ such that $0 < \epsilon < 1$. Let each individual grant himself the weight ϵ instead of 0, and revise the weights that he grants to others so that the ratios among those weights remain the same and all weights sum to 1. Thus i will revise the weight w_{ij} which he granted to j to $(1 - \epsilon)\, w_{ij}$, and W is replaced by $U_\epsilon = \epsilon I + (1 - \epsilon)W$, where I denotes the $n \times n$ matrix in which each person grants himself weight 1. By proper choice of ϵ, U_ϵ can be made as 'close' as one wishes to W, and thus it seems reasonable that there should be some U_ϵ which the group would agree to employ in place of W. Now U_ϵ has positive entries where W does, so U_ϵ is ergodic and has positive diagonal elements and is thus regular. Hence there is a consensus matrix L_ϵ with exclusively positive entries such that powers of U_ϵ approach L_ϵ as a limit and, as usual, L_ϵ is the unique solution to the matrix equation $L_\epsilon U_\epsilon = L_\epsilon$.

It may be proved that L_ϵ is also the unique consensus matrix solution of the equation $L_\epsilon W = L_\epsilon$,[14] from which it follows, remarkably, that the choice of ϵ is irrelevant: any perturbation of W of the sort described leads to the same consensual limit matrix L, where L is the unique solution to the equation $LW = L$. Thus this simple algebraic relation stands at the core of the theory of consensus, even as extended to ergodic matrices.[15]

Like Lehrer's original model, the above extension is based on the assumption that individuals are judged at least implicitly to be uniformly talented at an entire theoretical hierarchy of deliberative tasks. But suppose one wished to have a theory of decision-making which could accommodate situations in which individuals might somehow have explicit and differing perceptions of their colleagues with respect to this hierarchy.[16] In such cases the powers W, W^2, W^3, etc., would be replaced by products W_1, W_2W_1, $(W_3W_2)W_1$, etc., of possibly different weight matrices where, say, individuals grant weights to members of the group as physicists in W_1, as judges of physicists in W_2, as judges of judges of physicists in W_3, etc. A more complete discussion of this case appears in [15]. We shall mention here two basic results concerning circumstances which guarantee that the sequence of products W_1, W_2W_1, $(W_3W_2)W_1$, etc., converges to a consensus matrix. The classical proof of convergence of powers of a single regular matrix[17] admits a straightforward extension to the following theorems:

THEOREM 1. Let d_i denote the smallest entry of W_i. If $\lim_{i\to\infty}(1 - d_1)(1 - d_2) \ldots (1 - d_i) = 0$, then $W_iW_{i-1} \ldots W_1$ converges to consensus as $i \to \infty$.

In particular, if for some $d > 0$ (no matter how small), there is an n (no matter how large) such that $i \geq n$ implies $d_i \geq d$, then convergence takes place. This is a very weak assumption of minimal respect.[18]

THEOREM 2. Let i_1, i_2, i_3, etc., be any increasing sequence of positive integers. Let $U_1 = W_{i_1}W_{i_1-1} \ldots W_1$, $U_2 = W_{i_2}W_{i_2-1} \ldots W_{i_1-1}$, etc., and let e_i denote the smallest element of U_i. If $\lim_{i\to\infty}(1 - e_1)(1 - e_2) \ldots (1 - e_i) = 0$, then $W_iW_{i-1} \ldots W_1$ converges to consensus as $i \to \infty$.

Theorem 2 specializes to Theorem 1 when $i_1 = 1$, $i_2 = 2$, etc. As above, if for some $e > 0$, there is an n such that $i \geq n$ implies $e_i \geq e$, then convergence takes place. In particular, this condition obtains when the sequence W_1, W_2, etc., converges to a regular limit matrix.[19] Unlike the simple case of powers of a single matrix, there is in general no way to compute $\lim_{i\to\infty}W_iW_{i-1} \ldots W_1$ exactly. But convergence assures that a group can in principle compute by repeated multiplication as close an approximation to the limit as desired.

5. Conclusion: Some Unresolved Problems.

Granted the acceptance of weighted arithmetic means as averaging functions and a subjectivist determination of weights, the formal mechanism underlying Lehrer's theory is beyond reproach. For on these assumptions the classical Markov chain convergence theorems are incontestably relevant to the issue of rational consensus, and yield significant insights about group decision-making.

These theorems entail, first of all, the impossibility of rational disagreement where certain patterns of respect obtain among members of a group. No such results appear in decision theoretic literature prior to Lehrer's work. Even in the context of averaging by weighted arithmetic means with subjectively determined weights, earlier approaches were stymied by the possibility of disagreement about the appropriate (first order) weights, and intimidated by the specter of an infinite regress of decision problems from facing up to a hierarchy of higher order weight matrices. Yet in retrospect Lehrer's notion of consensus as the limit of a convergent Markov chain appears, to its credit, to be no more mystifying than the notion of continuously compounded interest.[20] At the same time, conditions under which rational consensus may fail are clearly delineated. For as noted above the failure of regularity (or ergodicity, if one accepts my arguments in Section 4) vitiates the guarantee of consensus. Given this limitative theorem, Lehrer's results have much the same character as the theorems of social welfare theory which show that majority rule yields a rational consensus for certain kinds of voting patterns while failing to effect consensus in general.

It should be noted, however, that Arrow's analysis of ordinal consensus goes much further than a simple description of cases where majority rule may be employed. For it rationalizes the limited use of this rule (while also proving that *no* acceptable rule is universally applicable) by showing that it satisfies certain very general properties which, it is argued, characterize democratic rational social choice. Thus a full analogue of Arrow's results would require a general explication of rational averaging in which weighted arithmetic means occupied no a priori privileged position. Such an analysis would, at the outset, regard all functions mapping opinion vectors to scalars as candidates for rational averaging functions. One would then seek to exclude certain of these functions 'from above' by noting that they fail to satisfy one or more general features identified as intuitively rational. One might feel, for example, that any averaging function deserving the name ought always to produce a number in the interval determined by the largest and smallest values of each opinion vector. This axiom would exclude many functions from consideration. Indeed, if one supplements this axiom with the requirement that an acceptable averaging function be linear, it is easy to show that one has abstractly characterized the class of weighted arithmetic means.[21] But linearity in this context seems a pleasant incidental feature of such means, rather than an intuitively desirable basic property of averaging functions It would be interesting to know if there is an alternative characterization of weighted arithmetic means that does not directly postulate linearity.[22]

In addition to the above, there remain some interesting open questions of a more localized sort. For even within the context of averaging by weighted arithmetic means, the issue of the meaning and determination of the weights deserves further attention. Violent opponents of subjectivist estimation will of course remain unconvinced of the value of any enterprise which seeks to quantify in this mode. But more can profitably be said to those inclined to sympathy, however tentative it may be. To tell an individual that his weights should sum to 1, that they should reflect his respect for the expertise of members of his group, and that they should be chosen with a willingness to employ them in weighted averaging, is, after all, to provide minimal guidance for an important task.

Given the great variety of decision problems it is unlikely that a detailed, comprehensive treatment of weights can be achieved. Worthwhile insights may nevertheless emerge from the study of special cases. Suppose, to take an example due to Lehrer, that n individuals are given a collection of N objects, each bearing a label from the set $\{1, 2, \ldots, k\}$, and must determine the fraction p_j of objects bearing the label j for each $j = 1, 2, \ldots, k$. Suppose further that the collection is partitioned into n disjoint subsets, and that individual i examines a subset with N_i objects and reports, for each $j = 1, 2, \ldots, k$, the fraction p_{ij} of objects in that set which bear the label j. The $n \times k$ matrix $A = (p_{ij})$ will only rarely exhibit consensus. But the rational sequence of weights $w_1, w_2, \ldots, w_n$ with which to average the columns of A is immediately apparent. Individual i should receive a weight proportional to the size of the set which he examines, so that $w_i = N_i/N$. For assuming that individuals count correctly, it is easy to check that $p_j = w_1p_{1j} + w_2p_{2j} + \cdots + w_np_{nj}$ correctly reports the fraction of objects in the entire set which bear the label j. Artificial as this simple combinatorial exercise may be, it illustrates in pure form the type of decision problem in which deliberative responsibility is partitioned in such as way that committee reports are accorded full credibility but weighted to reflect the scope of their unique concerns. An elaboration of this notion involving judgment based on weighted criteria appears in [15].

At the other end of the spectrum one encounters cases like the following, which is borrowed from classical statistical decision theory: Individuals are attempting with unbiased devices of differing accuracy to measure a quantity μ. Their actual estimates $a_1, a_2, \ldots, a_n$ of this number are regarded as realizations of a sequence of independent random variables $X_1, X_2, \ldots, X_n$ with $E((X_i - \mu)^2) = \sigma_i^2$. If, as is often the case, the variance is taken as a measure of the expected disutility of adopting the estimate produced by a random variable, the group should adopt as their estimate of μ the number $w_1a_1, + \cdots + w_na_n$, where the weights w_i are chosen to minimize $E((w_1X_1 + \cdots + w_nX_n - \mu)^2) = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \cdots + w_n^2\sigma_n^2$. A little partial differentiation now yields the formula $w_i = (1/\sigma_i^2)/(1/\sigma_1^2 + 1/\sigma_2^2 + \cdots + 1/\sigma_n^2)$. In decision-making contexts like the one under discussion one might initially have been inclined to endorse the estimate of the individual with smallest variance as the rational group estimate. Yet such a policy is demonstrably inferior (with respect to variance minimization) to the use of the above weighted average. One might roughly characterize this result as an indication of the wisdom of collective deliberation on a single matter, even where some individuals are more expert at this matter than others.

The foregoing examples illustrate rational approaches to the choice of weights in several highly simplified contexts involving little or no subjectivist estimation. Yet they point to two important considerations in the evaluation of individual expertise: To the extent that an individual offers a unique point of view on a problem, what is the scope of his field of vision? To the extent that he is engaged in communal deliberation, how shall his contribution to the avoidance of error be weighted? To take these difficult questions seriously, as Lehrer urges, is to view individuals as indicators of truth quite as deserving of analysis, evaluation, and improvement as the technological instruments which they often invent to assist their thinking about the world.

University of Tennessee

Notes

[1] This essay was prepared while the writer was a Fellow at the Center for Advanced Study in the Behavioral Sciences. I am grateful for financial support provided by the National Science Foundation (BNS 76–22943 A 02), the Andrew W. Mellon Foundation, and the University of Tennessee.

[2] The original proof of this theorem appears in Arrow [1]. A particularly clear exposition, including a survey of recent developments, appears in Roberts [13].

[3] A number of years ago French [3] and Harary [4] initiated just such a study of group power structures, based on the theory of homogeneous Markov chains. Lehrer's simplest normative model is based on this same mathematical structure, though rationalized from a wholly different perspective.

[4] This scheme, endorsed by Lehrer in [12], is described in detail in [15]. A related method, the 'Delphi technique', appears in Helmer [5].

[5] This will not be the case if only two individuals are involved. Even in larger groups, individuals may reveal themselves to others through a previously known expository style or philosophical orientation. In such cases individuals need to exercise a certain deliberative discipline to supplement that imposed by the format of anonymous position papers.

[6] The essentials of matrix algebra may be found in Kemeny and Snell [6, Section 1.11].

[7] However, consensus may fail in both A and W, yet obtain in WA. See [15, Section 4].

[8] As one would hope, a sufficiently close approximation to consensus in W can communicate an adequate approximate consensus to WA. See [15, Note 2] for a precise formulation.

[9] An individual may be listed more than once in such a sequence. If, for example, i respects himself, then $[i, i]$ is a vector of positive respect (of length 2) from i to himself. Of course, even if i does not respect himself, there may be a vector of positive respect from i to himself.

[10] This is the usual practical criterion for regularity. In fact, if W is $n \times n$, one need compute at most the powers $W, W^2, \ldots, W^{n^2-2n+2}$, for regularity always reveals itself, if at all, through the appearance of a matrix with exclusively positive entries somewhere in this finite list. (See Seneta [14] for a proof.) Thus there is what logicians call an *effective* procedure for determining regularity. Regularity also clearly follows from ergodicity, given the existence of at least one individual who respects himself. But there are regular matrices in which no one respects himself. Indeed, if $W = (w_{ij})$ is any $n \times n$ matrix with $w_{11} = w_{22} = \cdots = w_{nn} = 0$ and all other entries positive, then W^2 has exclusively positive entries.

[11] See Kemeny and Snell [6, Section 4.1] for a fuller discussion.

[12] For it is easy to check that if increasing powers of W converge to a consensus matrix with exclusively positive entries, then for every matrix A with n rows, WA, W^2A, etc. converges to a consensus matrix in which each individual's opinions in A are given some positive weight. Conversely, if WA, W^2A, etc, converges to such a matrix for every A, then W must be regular since convergence holds in particular for the $n \times n$ matrix $A = (a_{ij})$, where $a_{ij} = 1$ if $i = j$ and $a_{ij} = 0$ if $i \neq j$. On the other hand, a fixed A may exhibit agreement to the extent that WA, W^2A, etc. converges to consensus even if W is not regular. Indeed, if A is a consensus matrix then WA, W^2A, etc. converges (to A) for every W, by the remarks in the next to last paragraph of Section 1 of this essay. Finally, it should be noted that there are non-regular matrices, increasing powers of which converge to consensus matrices in which one or more columns consist entirely of zeroes. The initial opinions of individuals corresponding to these columns are, *with their consent*, accorded no weight. One requires a

broader notion than regularity if one wishes to capture this sort of consensus. See [2, Ch. 8] for an account of convergence of this more general variety.
[13] Kemeny and Snell [6, Section 4.1].
[14] Ibid., Section 5.1.
[15] Further justification for employing L to average A follows from the fact that when W is ergodic, $(W + W^2 + \cdots + W^n)/n$ converges to L as $n \to \infty$ [6, Section 5.1]. Of course, if W is regular the sequence W, W^2, etc. also converges to L.
[16] It should be noted that Lehrer has always been aware that the simplicity of employing powers of a single weight matrix is gained at the expense of narrowing applicability of his model. Indeed, it was Lehrer who asked me to look into convergence to consensus in a more general setting.
[17] See Note 11.
[18] We may have $\lim_{i\to\infty}(1 - d_1)(1 - d_2)\cdots(1 - d_i) = 0$ even if $\lim_{i\to\infty} d_i = 0$ as, for example, when $d_i = 1/i$. More exactly, it may be shown that $\lim_{i\to\infty}(1 - d_1)(1 - d_2)\cdots(1 - d_i) = 0$ just when $d_1 + d_2 + \cdots$ converges to no finite limit. Thus, while d_i may converge to zero, the convergence cannot occur too rapidly.
[19] See Wagner [15], pp. 344–5, for a fuller account.
[20] Recall that one dollar invested for a year at interest rate r, compounded n times, has the value $(1 + r/n)^n$ at the end of that year. Perhaps contrary to one's initial intuitions, this quantity does not increase without bound as n increases, but, as shown in most calculus texts, converges to e^r, where e is the base of the system of natural logarithms. Hence e^r may be regarded as the value of one dollar invested for a year at interest rate r, compounded continuously. As in Lehrer's model, a simple convergence theorem enables one to make sense of an initially perplexing notion.
[21] For if we denote an opinion vector by $(x_1, x_2, \ldots, x_n)$ and the averaging function by f, linearity implies that $f(x_1, x_2, \ldots, x_n) = w_1x_1 + w_2x_2 + \cdots + w_nx_n$ for some fixed sequence $w_1, w_2, \ldots, w_n$ of real numbers. For the vector $(x_1, x_2, \ldots, x_n)$, where $x_i = 1$ and $x_j = 0$ if $j \neq i$, the first axiom implies that $0 \leq f(x_1, x_2, \ldots, x_n) = w_i \leq 1$. Similarly, if $(x_1, x_2, \ldots, x_n)$ is such that $x_i = 1$ for each i, then

$$f(x_1, x_2, \ldots, x_n) = w_1 + w_2 + \cdots + w_n = 1.$$

[22] A very attractive characterization of this sort has recently (May 1979) come to light. Weighted arithmetic means

$$f(x_1, \ldots, x_n) = w_1x_1 + \cdots + w_nx_n$$

enjoy what might be called the *allocation property*: For each $n \times k$ matrix (a_{ij}) with all row sums equal to some fixed s, one has

$$f(a_{11}, \ldots, a_{n1}) + f(a_{12}, \ldots, a_{n2}) + \cdots + f(a_{1k}, \ldots, a_{nk}) = s.$$

This property is particularly desirable in the case of decision problems in which the values of the decision variables are required to have a fixed sum, for without such a property, *ad hoc* normalizations would be required after the application of the averaging function to the columns of (a_{ij}). I conjectured, but was unable to prove, that the functions $f: R^n \to R$ satisfying the allocation property and the inequality $\min\{x_i\} \leq f(x_1, \ldots, x_n) \leq \max\{x_i\}$ were precisely the weighted arithmetic means. Professor J. Aczél has kindly furnished a proof of this conjecture, which in fact assumes the allocation property only for $k = 2$ and $k = 3$.

References

[1] Arrow, Kenneth, *Social Choice and Individual Values*, Cowles Commission Monograph 12 (John Wiley and Sons, New York, 1951; second edition, 1963).
[2] Berman, A. and Plemmons, R., *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979).
[3] French, John R. P., Jr., 'A Formal Theory of Social Power', *Psychological Review* **63** (1956), 181–194.
[4] Harary, Frank, 'A Criterion for Unanimity in French's Theory of Social Power', in *Studies in Social Power*, D. Cartwright (ed.) (Institute for Social Research, Ann Arbor, Michigan, 1959), pp. 168–182.
[5] Helmer, Olaf, 'The Systematic Use of Expert Judgment in Operations Research', P-2795, The RAND Corporation, 1963.
[6] Kemeny, John G. and Snell, J. Laurie, *Finite Markov Chains* (Springer Verlag, New York, 1976).
[7] Lehrer, Keith, 'Social Consensus and Rational Agnoiology', *Synthese* **31** (1975), 141–160.
[8] Lehrer, Keith, 'When Rational Disagreement is Impossible', *Noûs* **10** (1976), 327–332.
[9] Lehrer, Keith, 'Induction, Consensus and Catastrophe', in *Local Induction*, R. J. Bogdan (ed.) (Reidel, Dordrecht-Holland, 1976), pp. 115–143.
[10] Lehrer, Keith, 'Social Information', *The Monist* (1977), 473–487.
[11] Lehrer, Keith, 'Consensus and Comparison: A Theory of Social Rationality', in Hooker, C. A. *et al.* (eds.), *Foundations and Applications of Decision Theory* (Reidel, Dordrecht-Holland, 1978), pp. 283–309.
[12] Lehrer, Keith, 'A Model of Rational Consensus in Science', to appear.
[13] Roberts, Fred S., *Discrete Mathematical Models* (Englewood Cliffs, N.J., Prentice-Hall, 1976).
[14] Seneta, Eugene, *Nonnegative Matrices* (John Wiley and Sons, New York, 1973).
[15] Wagner, Carl, 'Consensus through Respect: A Model of Rational Group Decision-making', *Philosophical Studies* **34** (1978), 335–349.