Destroying the Consensus

Author(s): Barry Loewer and Robert Laddaga
Source: *Synthese*, Vol. 62, No. 1, Consensus (Jan., 1985), pp. 79-95
Published by: Springer
Stable URL: https://www.jstor.org/stable/20116085
Accessed: 12-06-2019 16:10 UTC

BARRY LOEWER AND ROBERT LADDAGA

DESTROYING THE CONSENSUS*

Should the U.S. government proceed with the development of the MX missile? This is a complex question. Answering it involves considering many factors including assessments of alternative defense systems, national priorities and so on. One of the factors is whether or not the MX system will survive a Soviet first strike. There is a range of expert opinion on this question. There are also differences of opinion concerning who the experts are and the extent to which their opinions are to be relied on. Given this diversity of views what is a policy maker, in this case the president, to do? Is there some rational way to generate a consensus out of this wealter of opinions?

In their book *Rational Consensus in Science and Society*[1] Keith Lehrer and Carl Wagner claim that under certain circumstances there does exist a uniquely rational way of combining the opinions of 'experts' into a consensus. They argue that even though the experts disagree they may be rationally committed to a consensus. As they construe it, the consensus should embody all the information contained in the individual views. The experts, recognizing this are, under certain conditions, rationally committed to changing their views to the consensus. The president might be able to employ Lehrer and Wagner's method to extract a consensus from his advisors which he could use in his own deliberations.

Lehrer and Wagner develop their account of consensus within a Bayesian framework. They imagine a group of individuals each with a subjective probability distribution over some language. According to them consensus, if it exists, is also represented by a probability distribution. The consensus distribution is a weighted average of the individual distributions. The determination of the weights is crucial in their account and will be discussed later. The authors see their account of consensus as contributing to the solution of a number of related problems.

(1) It is intended as an analysis of the concept of consensus as in, for example, "There is a consensus among geologists that the earth is at least 4 billion years old."

Synthese **62** (1985) 79–95. 0039–7857/85/0621–0079 $01.70

(2) It attempts to explain the rational force of appeals to consensus. They suggest that appeals to 'intuition' of the sort commonly made in philosophy are often appeals to consensus.

(3) The account is offered as a solution to group decision problems in which members of the group differ in their beliefs but must select a single probability assignment before choosing a course of action.

(4) It provides an answer to an important question which has not received much attention: How ought a person alter his beliefs when he learns the beliefs held by others?

An adequate account of consensus might also contribute to the 'meta-analysis problem' which has lately been the subject of much discussion in statistical psychology.[2] Given a large number of studies in some area, say concerning the effectiveness of psychotherapy, what are the best ways of combining the studies? This problem is especially relevant in the social sciences where one frequently finds an enormous number of studies of related questions of variable reliability and often pointing to differing conclusions. A solution to the problem of how the beliefs of experts should be combined may be applicable to the problem of extracting consensus from many studies. In any case, it is clear that an account of consensus may contribute to the solutions to a number of important philosophical and practical problems.

The structure of our paper is as follows: In section 1 we explain the Lehrer and Wagner method for attaining consensus. In section 2 we show that their arguments for their method are faulty and sketch a more Bayesian account of learning from others. Furthermore, although their account is developed within a Bayesian framework we argue that in a number of respects it is quite unBayesian. There are circumstances in which their approach is incompatible with Bayesian ideas. In section 3 we argue that their account is inadequate as an analysis of 'consensus'. It focuses on certain kinds of agreements in belief at the expense of others. We show that, because of this, their method can lead to incoherent results. In section 4 we discuss the role of consensus in science from a Bayesian perspective.

1.

The best way to explain Lehrer and Wagner's method for acheiving consensus is by example. Suppose that three 'experts' Arabella, Barbarella, and Cinderella are convened to give advice on whether or not a tax cut will raise interest rates. Each begins the meeting with a degree of belief in the statements that a tax cut will raise interest rates. Typically their opinions will differ. The difference may be due, in part, to their having different information. As a first step they exchange information and the reasoning that leads to their views. Suppose that after exchanging information $P_A(S) = 0.3$, $P_B(S) = 0.6$ and $P_C(S) = 0.8$. There is agreement and disagreement in the group. All three think that the probability that a tax cut will raise interest rates is at least 0.3. Barbarella and Cinderella think that S is more likely than not; a view not shared by Arabella. Is there any way to represent whatever consensus exists in the group? Lehrer and Wagner assume that consensus should be represented by a single probability assignment. Aside from observing that if there is such a consensus probability it must be between 0.3 and 0.8 it seems that little reason can be given for favoring one value as the consensus over another. However, Lehrer and Wagner suggest that there may be more information present in the group, so far neglected, which will lead to consensus. This additional information is the extent to which each member of the group values or respects the opinions of members of the group including himself.

> What is needed is some kind of average but an average that takes into consideration the weight or respect that members of the group would assign to each other. Moreover, the average must be arrived at in such a way that all the members of the group are committed to such a probability assignment as a rational consensus. Suppose a given member of the group asks himself how much weight he would give to the opinions of each member of the group. To make the question more precise, let him ask how he would divide a unit vote among members of the group as potential selectors of an assignment for the group. (p. 20)

There are two points to notice. First, Lehrer and Wagner assume, without argument, that the consensus, if it exists, must be a *weighted average* of the individual probabilities. Second, the weights are determined by the pattern of *respect* that members of the group have for each other. We will examine both these points later. For now we will only observe that the authors are not very explicit concerning how they intend 'weights of respect' to be interpreted. Since the respect weights must sum to one irrespective of the size of the group they measure *relative* respect (as opposed to some absolute notion). A value of zero, however, is absolute. According to Lehrer and Wagner "... we suppose that an assignment of zero means that a person is worthless as a guide to selection" (p. 20). Suppose that Arabella's, Barbarella's and Cinderella's respect weights are as represented in the following matrix:

	A	B	C
A	0.8	0.1	0.1
B	0.2	0.7	0.1
C	0.4	0.1	0.5

Arabella begins by assigning a probability of 0.3 to S. When she learns that Barbarella assigns probability 0.6 and Cinderella 0.8 this may cause her to think that her assignment is in error. Perhaps they possess some information that she does not (this information did not come out during the exchange) or perhaps she allowed her heart to influence her probability assignment thus biasing it. In any case,

> It is reasonable for a person to accept the implication of the weights he assigns and attempt to improve upon his initial probability assignment by taking a weighted average, using the weights he assigns, of the probability assignments of members of the group. (p. 22)

So Arabella can *improve* on her probability assignment 0.3 by changing it to $P^1_A(S) = 0.8 \times 0.3 + 0.1 \times 0.6 + 0.1 \times 0.8 = 0.38$. Presumably, this is an improvement since she has corrected her degree of belief by taking the others' degrees of belief into account to the extent that she respects their opinions. Barbarella's and Cinderella's improved probability assignments are $P^1_B(S) = 0.56$ and $P^1_C(S) = 0.58$. The new degrees of belief are closer together than the old on any reasonable measure of closeness (e.g., the sample variance of the assignments). While the three have not yet attained consensus they are closer to it. Lehrer and Wagner argue that the Arabella *et al.* can further improve upon these probability assignments by following the same amalgamation procedure. At stage 2 disagreement will be further reduced. If the process is continued then each of the individual probability assignments converge towards a single assignment P^*. This is the rational consensus assignment.

The Lehrer–Wagner method for attaining consensus is general. Given a group of n individuals, an $n \times n$ matrix specifying the respect weights each individual assigns to himself and others (these must sum to 1), and n probability assignments to a set of exclusive and exhaustive statements we can obtain a unique consensus probability assignment as long as the respect matrix satisfies a condition Lehrer and Wagner call 'communication of respect'. This condition holds if at least one member of the group assigns positive respect to himself and there is a chain of positive respect from each member of the group to every other member. There is a chain of positive respect from person I_1 to person I_n just in case we can find a sequence of members of the group $I_2 \ldots I_{(n-1)}$ such that I_k assigns positive respect to $I_{(k+1)}$ for each k. The consensual assignment can be obtained by first obtaining the vector that results from taking the nth power of the original matrix as n goes to infinity and then multiplying the probability vector by this vector.

2.

Lehrer and Wagner claim that the consensus probability that results from employing their method "constitutes the best summary of the total information of the group" (p. 19). Furthermore, they claim that each member of the group is rationally committed to the consensus assignment. If Arabella were to fail to revise her degrees of belief to P^* she would be irrational in that she would be failing to take into consideration all the information available to her. These are interesting claims. We will begin our critique by showing that the argument which is intended to show that members of the group ought rationally to adopt the consensus view is defective.

Lehrer and Wagner's argument for the claim that it is rationally required for a member of a group (in which each member assigns positive respect to every other member) to change his probability assignment to the consensus assignment goes as follows (p. 23):

> (1) Every member of the group assigns positive respect to every other member of the group. (2) It is rational to aggregate information by taking the weighted average (the weights being the respect values) of individual probability assignments. (3) The respect values remain the same after the procedure of aggregation is carried out. (4) So it is rational to iterate the procedure. (5) In the limit repeated iterations of the individual assignments converge to a single assignment, the consensus assignment. (6) So it is rational for each member of the group to adopt the consensus assignment.

The argument depends on the crucial notion of 'respect'. Why is it rational to aggregate information by taking the average weighted by respect values? Lehrer and Wagner seem to offer two arguments for this procedure. One is that it would be inconsistent to assign nonzero respect values to other people's views and not change one's own view to the weighted average. For example, they write: "Since, in fact, he does not assign them a weight of zero, his refusal to average is inconsistent with the weights he actually assigns" (p. 23). So it is part of what they mean by 'respect' that one must, on pain of inconsistency, update one's views by averaging. The second argument is that their procedure follows from some apparently plausible axioms governing the choice of a group probability assignment. We deal with the arguments in turn.

The central notion of respect is left frustratingly obscure in Lehrer and Wagner's discussion. When we say that we respect someone's opinion we usually mean that it should be given a hearing or that we intend to take it seriously in forming our own opinion. Generally, though not inevitably, when we respect someone's opinion we will alter our view to take his view into account. (Respecting someone's opinion is different from respecting his right to have or voice an opinion.) But there certainly seems to be nothing inconsistent in saying that we respect Arabella's opinion and yet not alter our views to the consensual probability assignment. After seriously considering her opinion we might think we perceive an error in her reasoning and so not let her view influence ours. Or we might take her view into account but not in the way counseled by weighted averaging.

We think that Lehrer and Wagner's argument for their method trades on an ambiguity. They sometimes use 'respect' as a technical concept whose very meaning implies that it is inconsistent not to average. At other times they seem to intend some nontechnical notion. The difference in interpretations is crucial. On the technical interpretation step 2 of the argument is trivial. However, on that interpretation one might very well question why anyone does or should assign anything other than zero respect to other people. On the other hand, if 'respecting someone's views' means giving them serious consideration then it is clear that we frequently do and should assign positive respect to others. However, it is far from obvious that the correct way of taking other people's views into account is the Lehrer–Wagner procedure.

We are not certain why Lehrer and Wagner think that their procedure is the only rational way of taking into account the views of others whom we respect. They may think that it is the correct procedure because they can show that every other procedure is irrational. Perhaps this is what they think their derivation of the procedure from axioms is supposed to show. We do not wish to quarrel with their axioms (see p. 109) at this point but we will observe that their approach involves some questionable assumptions.

First, they assume that consensus is to be represented uniquely and by a single probability assignment. They also assume that it is rational for each member of the group to adopt the consensus assignment. It is only given these assumptions that their result that weighted averaging is the only way to achieve consensus while satisfying their axioms follows. Neither assumption seems to us to be especially plausible. We can see no reason why there might not be more than one probability assignment which can serve as a 'consensus assignment'. More importantly, it may be that whatever consensus exists in a group is better represented by a set of probability assignments than by a single assignment. Isaac Levi makes such a suggestion in his contribution to this volume and although we do not agree that we should limit ourselves to convex sets of assignments as he does we do think that his suggestion is quite natural.

Even if we grant that consensus should be represented as a single probability assignment it is not at all obvious why members of the group are rationally committed to altering their probability assignments to the consensual assignment. Perhaps they think that this conclusion follows from their claim that refusing to take another's views into account by averaging where the average is weighted by respect values is equivalent to assigning zero respect and that is equivalent to not taking the other's view into account at all; that is seeing her as no better than random device. So they seem to be saying that either you assimilate the information that A has a certain view by their method of weighted averaging or you are condemned to irrationality for neglecting this important information. We will show that this conclusion cannot be correct by sketching an alternative model of respect and the assimilation of other people's opinions. On our alternative account a person may consider the fact that someone has certain beliefs to be valuable information and may alter her views in light of this without following the Lehrer–Wagner procedure.

Suppose that Arabella assigns a probability of 0.7 to the hypothesis H that a certain hurricane will strike the coast. Barbarella is another meteorologist whose views Arabella respects; she is ready to alter her opinion in the light of Barbarella's beliefs. We suppose that she asks Barbarella her opinion. To keep things simple, we will suppose that Barbarella can answer in only three ways: 'it will strike the coast', W, 'it will not strike the coast', N, and 'it is as likely to strike the coast as not', E. (These correspond to Barbarella's assigning degrees of belief of 1, 0, and 0.5 to H.) Respect for Barbarella is reflected in her assigning the following likelihoods: $P(Y/H) = 0.9$, $P(Y/-H) = 0.1$, $P(N/H) = 0.9$, $P(N/-H) = 0.1$, $P(E/H) = 0.5$, $P(E/-H) = 0.5$. This assignment reflects the fact that Arabella thinks that Barbarella's opinion is indicative of the truth. Suppose that Barbarella answers that the hurricane will strike the coast. Then Arabella can take this into account as evidence by using Bayes' theorem to calculate that $P(H/Y) = 0.954$. Arabella exhibits her 'respect' for Barbarella's opinion by considering it evidence to be conditionalized on. This model can be generalized to allow for conditionalization on any degree of belief that Barbarella may have. All that is required is that $P_A(P_BH = x/H)$ and $P_A(P_BH = x/-H)$ be well defined. If analogous functions are well defined for Barbarella then each can conditionalize on the other's degrees of belief. As long as $P_i(P_jH = x/H) > P_i(P_jH = x/-H)$ and $P_jH > 1/2$ Arabella's and Barbarella's posterior beliefs will be closer together than their prior beliefs, though of course, they typically will not be identical.

On the Bayesian model the 'respect' that Arabella has for Barbarella's opinion, that is the extent to which learning Barbarella's beliefs will affect her own views is given by the likelihood functions, $P(P_BH = x/H)$ and $P(P_BH = x/-H)$. These are, respectively, the probability that Barbarella assigns a degree of belief of x to H if H is true and if H is false. In a particular case, we can say she positively respects Barbarella's opinion if upon learning it her own degree of belief moves closer to it. This need not be the case. Learning that someone endorses H may be a reason for lowering our degree of belief in it. Zero respect for someone's opinion is represented by the situation in which learning her degree of belief effects no change.

Notice how this way of thinking about respect differs from Lehrer and Wagner's. On their account, as long as I consider every member of a group to be better than random devices for selecting opinions I must assign them respect weights which sum to 1. So if there are three members of the group who I consider highly reliable but equally so I must assign respect weights of 1/3 and the same if I consider them not very reliable but equally so. If the group is enlarged by adding someone whose views I respect then the respect weights I assign to other members of the group must decrease (to insure that these sum to 1).

There are two further ways in which the Bayesian model for assimilating other's views differs from Lehrer and Wagner's approach. First, there is no reason to iterate the Bayesian approach. To do so would be to count the information that A has a certain probability assignment twice and this would surely be irrational. Second, while Arabella's and Barbarella's views may be closer together after each conditionalizes on the other's views, their updated degrees of belief will typically still differ. Consensus of the sort sought by Lehrer and Wagner, that is a single probability assignment is not achieved. This shows, we think, that when the members of a group take one another's views into account in forming their opinions the result need not be a single agreed upon probability assignment.

We must register a misgiving concerning the Bayesian model of learning from the views of others. It requires that the likelihood assignments $P(P_B(H) = x/H)$ and $P(P_B(H) = y/-H)$ exist. But this seems quite an extreme idealization. The idealization is especially severe since these are the assignments after the members of the group have exchanged the information on which they based their assignments. It would be very nice to have a more realistic account of how the beliefs of others should be assimilated. But the nonrealism of the model does not affect our argument against Lehrer and Wagner. Our main reason for sketching the Bayesian model was to show that individuals may respect one another, in that they are prepared to alter their views in the light of learning the others views, without being committed to Lehrer and Wagner's account. If this is correct then the choice offered by Lehrer and Wagner, either assign zero respect and consider a person's opinion as no better than random or assign positive respect and alter your views by the method of weighted averaging, is a false one. We now want to show that their approach is quite 'unBayesian' in certain other respects as well.

It is clear from our previous discussion that changing one's degrees of belief by the method of consensus will typically yield results which differ from conditionalization. Generally the Lehrer–Wagner method grants too much weight to the evidence of other's opinions. Here is a simple example which illustrates this point. Arabella and Barbarella agree that an urn contains either 60% red balls and 40% white balls, hypothesis R, or 40% red balls and 60% white, W. However $PA(R) = 0.99$ and $P_B(R) = 0.01$. Suppose also that each of the two assigns a respect value of 0.9 to herself and 0.1 to the other. In this case the Lehrer–Wagner consensus assignment is $P^*(R) = 0.5$. From Arabella's perspective this is the change that results from evidence that has a likelihood ratio of 99 : 1 in favor of W. Such evidence would be obtained, for example, by selecting 11 white balls in succession. But surely this is too much weight for Arabella to grant to Barbarella's opinion, especially since she assigns her such a low respect value. Notice further that we obtain the same consensus assignment irrespective of how reliable or accurate Arabella and Barbarella consider each other as long as their views of each other are symmetrical. So whether Arabella has a great deal of respect for Barbarella (considers her a very reliable authority) or very little, the Lehrer–Wagner method can result in her counting Barbarella's opinion as very weighty evidence. The consensus assignment respesents a *compromise* assignment located 'between' the individual assignments. While compromise may be important in some circumstances, generally one would not and ought not be willing to change one's belief for the sake of compromise. Experts with different views may be willing to compromise in order to issue a joint report but they would not be expected to adopt the official position of the report in place of their own views.

There is a second way in which the Lehrer–Wagner method is incompatible with Bayesian ideas. Consider again the example discussed in the previous paragraph. Suppose that Arabella and Barbarella sample the urn and select (with replacement) three red balls, e. Suppose that they first conditionalized on this evidence and then obtained consensus. In this case $P_A(R/e) = 0.997$ and $P_B(R/e)$ so $P^*(R) = 0.626$. But if they first obtained consensus and then conditionalized they would assign $P^*(R) = 0.773$. This is typical. The assignment obtained by conditionalizing and then taking consensus generally differs from the assignment obtained by taking consensus and then conditionalizing. Now, conditionalizing on two pieces of evidence, e and e' has the property that the order in which the conditionalization is done makes no difference. Since order does make a difference when taking consensus we think that this shows that taking consensus is not equivalent to conditionalizing on evidence. We believe that taken together our arguments show that it is not only unBayesian but irrational to change one's views by the Lehrer–Wagner method.

3.

We have argued that a strictly Bayesian account of change of belief in the light of learning the beliefs of others does not conform to Lehrer and Wagner's method. Generally after the members of a group exchange views and reasons they will not adopt the same degrees of belief. In this section we will show that Lehrer and Wagner's method produces a probability assignment which may conflict with what consensus does exist and further it may not be rational for the group to select a probability assignment by their method in the context of a group decision problem.

"Consensus" means agreement. It can correctly be said that there is a consensus among biologists that Darwinian natural selection is an important cause of evolution though there is currently no consensus concerning Gould's hypothesis of speciation. This means that there is widespread agreement among biologists concerning the first matter but disagreement concerning the second. Within an area of investigation there may be certain matters concerning which there exists a consensus and certain matters for which there is no consensus. As long as the investigators respect one another (in Lehrer and Wagner's sense of 'respect') it is not possible to make this distinction since they are committed to adopting the same probability assignment. If they followed the method biologists (assuming that the communication of respect condition, etc., is satisfied) would all be committed to assigning the same probability to both the evolution hypothesis and Gould's hypothesis. Patterns of agreement and disagreement are obscured.

In addition the Lehrer–Wagner account of consensus can fail to represent certain agreements in the group. It was pointed out by Laddaga[4] that even though each member of a group has a probability assignment in which two propositions, H, K are probabilistically independent the consensus assignment determined by Lehrer and Wagner's method will not render H, K independent. Weighted averaging does not preserve independence. Here is a simple example: Arabella: $P(H \& K) = 0.12$, $P(H \& -K) = 0.08$, $P(-H \& K) = 0.48$, $P(-H \& -K) = 0.32$. Barbarella: $P(H \& K) = 0.30$, $P(H \& -K) = 0.20$, $P(-H \& K) = 0.30$, $P(-H \& -K) = 0.20$. Suppose that Arabella and Barbarella each have equal respect for themselves and for each other. Then the consensus assignment is $P(H \& K) = 0.21$, $P(H \& -K) = 0.14$, $P(-H \& K) = 0.39$, and $P(-H \& -K) = 0.26$. H, K are independent in Arabella and Barbarella's assignments but not in the consensus assignment. Although both agree that H, K are independent this bit of agreement is lost in the Lehrer–Wagner consensus probability. This suggests that their consensus assignment, whatever it uses, does not represent the agreement that exists in the group. Lehrer and Wagner respond to the failure of averaging to preserve independence with the comments "... in countless cases independence is simply not of much interest" and "In such situations the independence of certain compounds of these propositions is largely fortuitous".[5] But certainly there are situations in which independence is important and it is important to know that all the members of a group agree that two propositions are independent. For example, agreement concerning independence would typically be of great importance if the group is a team of scientists designing an experiment.

Laddaga's point about independence reveals not merely an undesirable feature but an actual inconsistency in Lehrer and Wagner's method of probability amalgamation. Suppose that we add the proposition I, which says that H, K are independent, to our example thus splitting each conjunction into two, one with I unnegated and one with it negated. Since Arabella and Barbarella assign a probability of 1 to I the consensus assignment P^* also assigns 1 to *I even though H, K are not independent according to P^**. Someone who followed Lehrer and Wagner's procedure would end up with a probability assignment in which H, K are not independent but in which they are certain that they are independent. Such an assignment is literally incoherent.

If the problem we have been discussing only involved independence perhaps one could respond, as Lehrer and Wagner do, that their method is intended to be used only in cases where independence is not important. However, this problem is endemic to their approach. There are many features of probability assignments which are not preserved by arithmetic averaging. Most theoretically grounded families of assignments are not closed under averaging. For example, Arabella and Barbarella may agree that a certain quantity is binomially distributed, however in the Lehrer–Wagner consensus the quantity is not binomially distributed. The problem arises, of course, since it is impossible for a single probability assignment to possess all and only the features that two different assignments have in common. An assignment obtained by averaging will always preserve certain features, e.g., mutual exclusivity, but it will fail to preserve others and it will have some characteristics not possessed by any of the original assignments. As Lehrer and Wagner observe, some of these features may be more important than others. But there is no reason to think that the features preserved by averaging are always the most important. Averaging will typically destroy the epistemic structure of a belief system. While Arabella and Barbarella may have belief systems in which certain propositions provide reasons for others this epistemic structure will be absent from averages of the two systems. This is one reason why it would generally be irrational to adopt an average of probability assignments (even if they belong to people whose views you respect) as your own assignment.

There is one application of the Lehrer–Wagner procedure for which, at first, it seems especially well suited. Suppose that a group of people face a common decision problem. All the members of the group must agree on an action. If anyone dissents then nothing can be done and all view that situation as bad. Each member of the group has his own probability assignment over the relevant states of nature and his own utility assignment over the possible consequences of the actions. For example, the group may be a committee which must reach a unanimous decision on some matter. In this situation it may seem appropriate for the group to obtain a single probability assignment and then choose the action which maximizes expected utility relative to that assignment. The Lehrer–Wagner method does seem to be a natural way of selecting an assignment in this case. Even if, as we have argued, it is a bad account of rational belief change and a poor analysis of consensus it does seem to be a reasonable way of coming up with a 'compromise' probability assignment, and that is what seems to be needed here. However, there is a problem with this approach. If the group members have differing utility assignments then it is possible that before they attain Lehrer–Wagner consensus with respect to the states of nature they already agree on which action is best but after they attain consensus they are thrown into disagreement. The following pair of decision problems illustrate the difficulty:

A's decision problem

	S_1	S_2
d_1	0.8	0.5
d_2	0.2	0.5
	0.6	0.4

B's decision problem

	S_1	S_2
d_1	0.35	0.6
d_2	0.65	0.4
	0.4	0.6

Of course the situation is just a special case of our previous observation that weighted averaging preserves some features of a probability assignment but not others. One feature that fails to be preserved is the ordering of expected utilities of actions computed relative to the probability assignments. Lehrer and Wagner might reply to this objection that it is still better for the group to compute the consensus probability otherwise they would be neglecting information. We have already answered this by arguing that their method does not take information into account in a rational manner. We have seen that the Lehrer–Wagner approach can actually destroy a consensus that exists in a group concerning the solution to a decision problem.

4.

Our discussion has, unfortunately, been mostly negative. We have shown that Lehrer and Wagner's model fails to solve any of the problems mentioned at the outset that a theory of consensus ought to solve. We think that they make two fundamental mistakes. The first is their claim that it is rational for an individual to change his degrees of belief by the method of weighted averaging. The second is their assumption that consensus should be represented by a single probability assignment. Given the mathematical result that under certain conditions iterated weighted averaging results in a single assignment it is not surprising that they would identify this with the consensus assignment. It is our view that the two questions: (i) 'How should an individual alter his beliefs in the light of learning the beliefs of others?' and (ii) 'How should consensus be represented?' are different and demand different answers. To answer the first we suggested the model of standard Bayesian conditionalization. We do not have a general answer to the second question. But we think that we can say something of interest concerning it by examining the role that consensus plays in science.

There is no doubt that consensus is an important concept in science. A scientific discipline, say spectroscopy theory, is in part identified in terms of the theories, problems, procedures, etc., on which there is agreement among members of the discipline. Membership in a discipline generally requires that one accept most of the views on which agreement is widespread. Furthermore, dissensus is viewed as generating problems to be solved. If there are two groups within a scientific discipline which differ in their beliefs concerning some theory then one expects the two to engage in mutual criticism and experimentation to settle their differences. We would not expect the differences to be settled by the adoption of some compromise view located between the two competing views.

The important question then is this: How can members of a group attain consensus, that is agreement, with regard to some important matter? We suggest the following procedure. First, they should exchange information (including their probability assignments) and conditionalize on this information. At this point they may discover that they are in agreement, but typically this will not be the case. Now, it is well known that if people begin with different probability assignments and experiment together conditionalizing on the evidence obtained then their posterior assignments generally will be closer together. If they continue experimenting then in the limit their posterior assignments converge. Furthermore, their posterior assignments not only converge with each other but also on the truth. That is: it is highly likely that after a number of experiments the investigators probability assignments will converge with each other and with the truth. This result requires that a number of conditions obtain; (a) each member originally assigns a positive probability to the true hypothesis, (b) each member assigns the same likelihood to the results of the experiments given the hypotheses (these must be the correct likelihoods to guarantee probabilistic convergence on the truth) and (c) the experiments must be independent and believed to be so. Incidentally, one might remark that here is a case where agreement on independence is crucial. So if there is consensus (or near consensus) with regard to the likelihood function then they will believe that continued experimentation will bring them closer to consensus concerning the hypotheses being tested.

In certain circumstances Bayesian theory can provide a more detailed account of how to achieve consensus. The Bayesian theory of experimental information enables one to calculate the expected information of an experiment. This allows one to rank possible experiments with respect to their expected information value. One can also calculate the expected increase in 'agreement' of an experiment. That is, the extent to which agreement (as measured by the variance of degrees of belief) can be expected to be decreased by an experiment. The interesting point is that even though the members of the group disagree (assigning different probabilities to H) they may agree that a particular experiment from among a collection has the greatest or a relatively great expectation of increasing consensus. So although they do not agree concerning the assignment to H they can agree about the course of investigation to pursue. On this model a group of scientists do not achieve consensus by adopting a compromise view. To do so would literally be to compromise the truth. Instead, investigators who agree on a program of research may come to agreement concerning hypotheses by carrying out that research.

Consensus and dissensus play a dynamic role in science. A certain amount of dissensus within a science is healthy. It encourages the development of different theories and the pursuit of different lines of research. But continued disagreement is undesirable since it indicates that the truth has not been found. As we have argued it would be a mistake to obtain agreement by adopting a compromise view. To do so would be to compromise the truth. On the model we sketched consensus is achieved, if at all, by experimentation and argumentation. A research tradition in a science can tolerate, may even encourage, a certain amount of disagreement on particular hypotheses and explanations. It is much more important to the cohesiveness of a research tradition that there be agreement concerning which programs of investigations are likely to prove fruitful. If there is disagreement over some issue there is reason to settle it. But successful resolution of disagreement requires agreement on a research program and on the interpretation of the results of research.

The notion of consensus plays an important role in science, politics, and philosophy. Lehrer and Wagner's work is valuable if only because it has brought the problems associated with consensus to the attention of philosophers. Our own discussion (and the other contributions to this issue) show that these problems are difficult and that, as of now, there is little consensus concerning how to to solve them. Though a consensus on 'consensus' does not exist perhaps there is sufficient agreement to lead to some progress on these issues.

NOTES

* We would like to thank Davis Baird, Bob Jannerone, and Ferdy Schoeman for useful discussions on consensus.

[1] Lehrer, Keith and Wagner, Carl, *Rational Consensus in Science and Society* (D. Reidel, Dordrecht, 1981). Parenthetical page references are to this book.

[2] Glass, Gene *et. al. Meta-analysis in Social Research*, 1981. One approach to meta-analysis employs weighted averages in a manner similar to Lehrer and Wagner's account of consensus.
[3] To apply the Lehrer–Wagner method the members of a group need not really be experts. They must, however, value one anothers views to a certain extent. The conditions under which Lehrer–Wagner consensus exists are given later in the paper.
[4] Laddaga, Robert, 'Lehrer and the Consensus Proposal', *Synthese* **36** (1977), 473–77.
[5] Lehrer and Wagner, 'Probability Amalgamation and the Independence Issue', *Synthese* **55** (1983) 339–46.
[6] This criticism of Lehrer–Wagner's method was suggested to us by Ferdy Schoeman.
[7] See Savage, Leonard, *The Foundations of Statistics* (Dover) for precise statements of these theorems.

Department of Philosophy
University at South Carolina
Columbia, SC 29208
U.S.A.