# Taking Advantage of Difference in Opinion

Richard Bradley
Department of Philosophy, Logic and Scientific Method
London School of Economics

January 26, 2007

**Abstract**

Diversity of opinion both presents problems and affords opportunities. Differences of opinion can stand in the way of reaching an agreement within a group on what decisions to take. But at the same time, the fact that the differences in question could derive from access to different information or from the exercise of different judgemental skills, means that they present individuals with the opportunity to improve their own opinions. This paper explores the implications for solutions to the former (aggregation) problem of supposing that individuals exploit these opportunities. In particular, it argues that rational individual revision of opinion implies that aggregation problems are instable in a certain sense and that solving them by exploiting the information embedded in individual opinion has profound implications for the kinds of conditions that we should impose on aggregation procedures.

## 1 Introduction

Diversity of opinion, even amongst experts, is a fact of life that from the point of view of decision making both presents problems and affords opportunities. When a group must make a decision about some matter, differences of opinion on, for instance, the likelihood or desirability of the consequences of undertaking possible courses of action, can stand in the way of reaching agreement on what to do. On the other hand, differences of opinion may reflect the fact that individuals are drawing from different sources of information or exercising different judgemental skills. And this presents individuals with an opportunity; namely to improve their own opinions, and hence their own decisions, by drawing on the information or skills embedded in the judgements of others. Let us call the problem associated with the first observation, that of deriving a single judgement on each of the relevant issues from the diverse set of individual ones, *the aggregation problem* and the problem associated with the second, of how to improve one's judgements in the face of the different judgements of others, *the revision problem.*

Much attention has been given to the aggregation problem engendered by diversity of opinion; somewhat less to the opportunities it affords and the significance of learning for aggregation. This is regrettable for (at least) two reasons. Firstly, since much of what we learn is from other people, how to deal with the diverse opinion that confronts us is a matter of central importance. Indeed it has some claim to being one of the constitutive problems of social epistemology. And secondly, the possibility of individual learning raises interesting questions for aggregation theory: Is it possible to summarise or encode within the collective judgement all the information distributed amongst the individuals making up the group? Is it possible to aggregate in a way which is robust with respect to possible improvements of individual opinion in the sense that these revisions do not change the aggregate judgements? And, if not, can we aggregate in such as way as to ensure that the aggregate of individuals' initial opinions are not totally at odds with the aggregate of the opinions that they would arrive at were they allowed to improve them?

At the heart of these questions is, I think, an issue about the compatibility of two kinds of demands: the demand that the judgement basis for a collective decision be 'appropriately' related to the opinions of the individuals making up the group and the demand that the collective judgement optimally (or adequately) reflects the information contained in individual judgements. I will argue in this paper that the two demands are difficult to reconcile and that making aggregation sensitive to the information contained in the opinions of individuals requires giving up some rather central principles of current aggregation theory.

I will proceed as follows. In the first two sections I will define the aggregation problem more carefully and then argue for two propositions: that the statement of an aggregation problem engenders opportunities for learning and hence revision problems of a particular kind and that, consequently, aggregation problems are unstable in the face of the demand that individuals' judgements should incorporate all the information available to them. In the sections that follow, I consider the possibility of responding to this instability by letting what we know about how opinions will be revised act as a constraint on the solution to the initial aggregation problem. This works best if revision tends to produce consensus amongst individuals and so in the third section I evaluate (and reject) a model of revision that suggests that it will. In the final section I consider the question of whether this method of producing aggregate judgements is consistent with some common principles of aggregation theory.

## 2 Aggregation Theory

To start, some basic assumptions and definitions. Let us take the objects of judgements or opinions to be prospects of any kind: that it rains tomorrow, that the government falls, that taxes rise, that the government will fall if taxes rise, and so on. Judgements may also be of various kinds, including categorical judgements, such as when a prospect is judged as true or false or as good or bad, comparative judgements such as those contained in a preference ranking of prospects, and quantitative judgements such as when a probability or utility is attached to a prospect. For convenience we represent a judgement on a prospect by a real number. What values it can take will depend, of course, on the kind of judgement involved: just zero or one in the case of categorical judgements, any number in the zero to one interval in the case of probability judgements, and so on. I will focus on probability judgements in this paper, but little will depend on this choice.

Let us call a set of judgements, one for each prospect in a given set of them, a *judgement set*. And call a vector of such judgements sets, one for each individual, a *state of opinion*.[1] A state of opinion can be represented by a table of the kind displayed below where each cell at the intersection of the $i$-th row and the $j$-th column displays the judgement of the $i$-th individual on the $j$-th prospect.

| | | | *Prospects* | | |
|---|---|---|---|---|---|
| *Individuals* | $X^1$ | $X^2$ | ... | ... | $X^m$ |
| $I_1$ | $x_1^1$ | $x_1^2$ | ... | ... | $x_1^m$ |
| $I_2$ | $x_2^1$ | ... | ... | ... | ... |
| ... | ... | ... | ... | $x_i^j$ | ... |
| $I_n$ | $x_n^1$ | ... | ... | ... | $x_n^m$ |
| | | | $\underbrace{\textit{aggregation}}$ | | |
| *Collective* | $x^1$ | $x^2$ | ... | | $x^m$ |

The aggregation problem associated with a particular state of opinion is simply that of deriving an acceptable aggregate or collective judgement profile from the individual ones. An aggregation procedure is a general purpose method for doing so, i.e. it is a way of deriving, from any given state of opinion, an acceptable aggregate or collective judgement on each of the relevant prospects. The interesting part of the problem lies in characterising acceptable or desirable profiles of collective judgements and in identifying the class of aggregation methods that produce them.

In its different instantiations, the problem has been extensively studied in social choice theory, statistics and elsewhere, typically by identifying the class of aggregation procedures consistent with one or more conditions on the relation between individual and collective judgements. In one rather general form the problem is at the heart of the theory of judgement aggregation developed by List, Dietrich and others: see for instance [14] and [6]. The problem of aggregation of diverse probability judgements has mainly occupied the attention of statisticians: Genest and Zidek [8] contains a very useful overview. Finally, the problem of aggregating utilities, qua values of a function representing a ranking of some kind, received its classic treatment in the work of Sen [19] and has since been studied by amongst others Gevers [9], D'Aspremont and Gevers [1] and Roberts [18]. It is impossible to do justice in a few paragraphs to the wealth of different results and applications that these investigations have produced and

[1] In social choice theory, a state of opinion is typically called a profile of judgements.

here I will simply comment on some of the more common conditions postulated in them.

Firstly, for a judgement profile to count as an *aggregate* of a set of individual ones it needs to be related to the latter in certain ways: in particular, the aggregate judgements need to be positively sensitive to the individual ones. This idea finds minimal expression in aggregation theories in 'unanimity preservation' conditions that require that unanimous individual judgements be reproduced or preserved at the aggregate level. Thus if everyone believes or accepts or prefers that $p$, the relevant unanimity preservation condition will require collective belief, acceptance or preference that $p$. Stronger expressions of the positive dependence idea may be justifiable too e.g. the requirement that any increase in the strength in individual opinion concerning $p$ should not lead to a reduction in the strength of the aggregate judgement that $p$. But only the minimal condition will play a role in our discussion.

Secondly, aggregation theories impose rationality conditions on both individual and collective judgements, typically in the form of consistency requirements such as that preferences be transitive or probabilities be additive. Consistency is a natural condition on judgement (some would say constitutive of it);furthermore it is difficult to see what sense could be made of the idea of an aggregate judgement depending on individual ones if either are inconsistent.[2] But as we shall see, rationality conditions can be very demanding.

Thirdly, independence conditions of various kinds are often invoked of which the most common, the independence of irrelevant alternatives condition, is the requirement that the collective judgement on some particular prospect depends only the individual judgements on this prospect and not on their judgements on any other (irrelevant) prospect. In terms of our tabular representation this means that the value of any collective judgement, $x^j$, should depend only on the $x_i^j$s, and not on the values contained in any other column.

Finally, it is very common to impose a 'universal domain' condition on aggregation functions to the effect that the domain of the function should contain all vectors of consistent individual judgement sets over the prospects in question. The thought here seems not to be that opinions are likely to take any possible form, but that the aggregation method itself should impose no constraints on judgements of individuals - rational individuals are in this sense sovereign.

# 3 Learning from Others

Any receipt of new information poses a revision problem for an agent; namely how they should modify their opinion in order to accommodate it. As with aggregation the interesting question is what counts as an acceptable revision and what rationality requires of us in this regard. Answers to this question are given by theories of revision: Bayesian conditioning models for probability revision are perhaps the best known - see Joyce [11] - but there are also well developed

[2] The controversy tends to centre on whether there *is* such a thing as a collective judgement, and not some much as to whether it should be rational in the event that it does.

theories of revision for categorical belief or acceptance - see Gärdenfors [7] - and nascent theories of preference revision - see Bradley [4]. But although these theories have been applied to a large variety of different types of information, models that explicitly address the kind of situation we are addressing, where an individual is presented with a diverse set of opinions, are thin on the ground and have received little attention from aggregation theorists. (There is a notable exception: the theory developed by Lehrer and Wagner [12], which we will examine more closely later on.) The lack is regrettable because I claim:

**Proposition 1** *The statement of an aggregation problem for a group engenders a revision problem for each individual in it.*

The basis of this proposition is essentially the observation that there are situations in which others' judgements, or the expressions of them, provide grounds for modifying one's own. The most obvious examples are those in which somebody has information that one does not hold. Suppose, for instance, that the question on which we hold different opinions concerns which of the three urns before us contains the red ball and that I know that you have sneaked a look in the first urn. Then I should adopt the opinion that the red ball is in the first urn iff you declare it to be (assuming that I have no grounds to question your honesty).

Cases in which we have reason to believe that others' judgements reflect information relevant to the question under consideration are common, if rarely as simple as the one described in the urn example. But someone's authority on a question may also derive from some expertise that they have or special training or method. Doctors may be able to make better judgements about someone's condition because their diagnostic abilities have been honed by experience, even though they may have no special information about that person's condition. Whenever someone is an authority on a question, whatever its source, we should regard their declared opinions as evidence, for or against, the truth of particular propositions.

Now the observation that expressions by others of their beliefs and preferences can be informative suggests that a central tenet of scientific methodology, the Principle of Total Evidence, applies in cases of diverse opinion. The principle says that my beliefs must be consistent with all the evidence available to me and, further, if I acquire new evidence, I should revise my beliefs to accommodate it. So it follows, in particular, that if I believe that someone is an authority on question A, then I should revise my beliefs in the light of his or her expressed judgement about A, just as I should revise my opinion in the face of any reliable evidence concerning A.

The implication is that, in a certain sense, the aggregation problem that was defined at the beginning of the paper may not be a stable one, given the strengthened rationality condition introduced here. For the very statement of the original aggregation problem generates information about the judgements of the various individuals which the Principle of Total Evidence requires each to take into account. And only in very special cases will this not imply that individuals are rationally obliged to revise their judgements. Hence:

**Proposition 2** *Aggregation problems are not generally stable if individual judgements must respect the Principle of Total Evidence.*

What follows for aggregation theory from the instability of the aggregation problem? This depends on how individuals can be expected to revise their opinions, a question about which we have said very little so far. The revision problem for a state of opinion is simply the question of how to revise an existing judgement profile in the light of the information presented by the statement of these opinions. A revision method in turn is a general-purpose procedure for deriving a new judgement profile from an old one on the basis of the given state of opinion. As such it will determine a transformation of one state of opinion into another by specifying for each individual $I_i$ the relationship between her initial judgement, $x_i^j$, on prospect $X^j$, and her new judgement $y_i^j$ on this prospect.

$$\begin{array}{c} \text{Opinion at } \mathrm{T}_0 \\ \textit{Prospects} \\ \begin{array}{|c||c|c|c|} \hline & X^1 & ... & X^m \\ \hline I_1 & x_1^1 & ... & x_1^m \\ \hline I_2 & x_2^1 & ... & ... \\ \hline ... & ... & ... & ... \\ \hline I_n & x_n^1 & ... & x_n^m \\ \hline \end{array} \\ \underbrace{\quad}_{} \, agg \\ \begin{array}{|c|c|c|c|} \hline C & x^1 & ... & x^m \\ \hline \end{array} \end{array} \xrightarrow[transformation]{} \begin{array}{c} \text{Opinion at } \mathrm{T}_1 \\ \textit{Prospects} \\ \begin{array}{|c|c|c|} \hline X^1 & ... & X^m \\ \hline y_1^1 & ... & y_1^m \\ \hline y_2^1 & ... & ... \\ \hline ... & ... & ... \\ \hline y_n^1 & ... & y_n^m \\ \hline \end{array} \\ \underbrace{\quad}_{} \, agg \\ \begin{array}{|c|c|c|} \hline y^1 & ... & y^m \\ \hline \end{array} \end{array} \xrightarrow[trans]{} ...$$

As we argued above, the transformation of the state of opinion induced by the individual revisions yields a new aggregation problem. It also yields new revision problems, for the manner in which individuals revise their judgements generate information about what they know about the relative reliability of the judgements of the others. Anne may know that Bob knows whether Cara is reliable on matters X. If she observes Bob to revise his judgements on X so as to bring them in line with Cara's, then Anne may surmise that Cara's judgements on X can be expected to be closely correlated to the truth. Anne should now revise her own judgements once again in the light of this. Such revisions must in practice come to a halt at some point; in principle they need not - unless a consensus is achieved. And each iteration will produce both a new aggregation problem and a new opportunity for revision.

All this leaves aggregation theory with both a problem and an opportunity. Aggregation theorists often argue that the possibility that opinions will change is in a sense not their problem, because at some point revisions induced by deliberation or enquiry must come to an end and a decision made on some collectively acceptable basis. But this retort is somewhat beside the point here. If the statement of the aggregation problem itself creates opportunities for improvement of individual opinion, and the aggregation theorist can recognise them, then there is a strong intuition that the aggregation method should be sensitive to the revisions that might be induced. To not do so would in effect be to throw away information available to individuals that could improve the collective judgement itself. In the worst case, application of aggregation procedures to the initial state of opinion without regard to the revisions it induces might produce a collective judgement that all would reject from the standpoint of their improved opinion.

The latter threat could be neutralised if the information creating the opportunities for learning could be incorporated in the aggregate judgement. One way of achieving this could be to let what we know about how opinions will be revised act as a constraint on the solution to the initial aggregation problem. More generally, we could let what we know about how individuals revise their opinions constrain our choice of aggregation principles and methods. Suppose, for instance, that our revision theory tells us that the improved judgements of individuals will agree on the value of some prospect. Then it would be natural to require that the aggregate judgement on that prospect should be just the consensual judgement everyone will eventually arrive at by revising their judgements in the light of the reported judgements of others.

This constraint on aggregation will have force only to the extent that individual revisions actually produce consensus. Sometimes it seems reasonable to expect this. Recall our urn example in which opinion differs on the question on which of three urn contains the red ball. If I know that you have looked in the first urn and you know that I have looked in the second, you do not declare the ball to be in the first urn and I do not declare it to be in the second, then both of us should conclude that it is in the third. An aggregation method that took cognisance of not just our opinions on the whereabouts of the red ball but also of our opinions about the quality of each others opinions on the whereabouts of the ball could determine this judgement 'up front', thereby exploiting all the subjectively available information contained in individual opinion. But is this example typical or exceptional?

# 4 Revision and Consensus

In this section I will evaluate a model of revision separately developed by Morris DeGroot [5] and by Keith Lehrer and Carl Wagner [12] (hereafter the DLW model) and which purports to show that rational revision must lead to a consensus.[3] Their model is applicable to the revision of both probability and utility judgements, but we will confine attention to the former. The DLW model assumes that individuals have not only opinions about some basic set of prospects, but also opinions about the quality of each other's opinions on these prospects. These latter judgements determine an 'respect' weight, taking a value between zero and one, on each person's opinions, roughly representing the probability that their opinion is the best of those on offer. Rationality requires of every

[3]The model and especially the claim of Lehrer and Wagner that specifies the uniquely rational way of aggregating probability judgements has been attacked from a number of different angles: See for instance Goodin [10] and the papers appearing in *Synthese* vol. 52 (1985). The discussion here picks up mainly on the themes addressed in Loewer and Laddaga [15].

individual that they modify their opinions in the direction of another person's declared probability judgements to the degree that they respect that person's judgements. In the DLW model this is achieved by each person adopting the respect weighted sum of the set of declared (prior) probabilities as their new (posterior) probabilities.

Consider a simple example with three individuals with prior probabilities for some particular event of 0.1, 0.2 and 0.15 respectively. Let the respect of each individual for each one of them (including themselves) be given by the $3 \times 3$ matrix displayed below, with the $i$-th row giving individual $i$'s degree of respect for the judgements of each of the three individuals (including herself). Then the posterior probability of the event for individual $i$, given the posted probabilities of all, is obtained by multiplying each individual's prior by $i$'s respect for that person and taking the sum. For instance, Anne's probability for the event in question is just $(0.4 \times 0.1) + (0.2 \times 0.2) + (0.4 \times 0.15)$. In the case displayed below this yields consensus on a probability of 0.14 for the event in question.

$$\begin{array}{cccc} & \textit{Weights} & \textit{Priors} & \textit{Posteriors} \\ \begin{array}{l}\text{Anne}\\ \text{Bob}\\ \text{Cara}\end{array} & \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.6 & 0.4 & 0 \\ 0.5 & 0.4 & 0.2 \end{pmatrix} & \begin{pmatrix} 0.1 \\ 0.2 \\ 0.15 \end{pmatrix} = & \begin{pmatrix} 0.14 \\ 0.14 \\ 0.14 \end{pmatrix} \end{array}$$

So long as the weights are strictly positive, revisions of this kind will produce a convergence of opinion, though not typically a consensus. But there is no reason for revision to stop at this point. Individuals should continue to revise their opinions in the light of those of the others until they have exhausted all the information contained both in the posted judgements and in the judgements about each other's judgemental authority reflected in the revisions induced by their posting. Opinions about the reliability of others' judgements about others' reliability may also be encoded in respect weights on their judgements. These weights may not be the same as those that they attach to others' judgements about the basic prospects: I might consider that someone is a poor judge of tomorrow's weather but an impeccable judge of people's skill at judging the weather; at 'knowing who knows'. But whether the weights vary or not makes little difference to the important conclusions; namely[4]:

1. Iterated revision of this kind will eventually produce a consensus on all issues (including that of how much respect should be accorded to each person's initial opinions) in a rather broad class of cases: roughly whenever there is some individual $i$ who respects him or herself and is such that there is a chain of strictly positive respect from each member of the group to $i$.

2. The agreed probability for each prospect will take the form of a linear or weighted average of the prior probabilities of each individual, with the weights on these probabilities being the agreed degrees of respect for these probability judgements.

[4]See Berger [3] and Wagner [20] for proofs of these claims.

The DLW model thus offers an answer to our opening question: the optimal aggregate judgement on any prospect is simply a respect-weighted average of the individual judgements. It is optimal in the sense that the consensual judgements that it preserves are the best summaries of the information contained in the group, including information that individuals hold about one another's judgemental competences (at least insofar as the information can be extracted from the judgements). It is also therefore robust with respect to rational revision. Finally it is compatible with the principles of aggregation theory discussed earlier: linear averaging of this kind respects the conditions of unanimity preservation, consistency of collective judgements (because a weighted sum of a probability measures is itself a probability measure) and a version of the independence of irrelevant alternatives condition.

But should individuals revise their judgements in the manner claimed by the DLW model? The best way of answering this is to test the implications of the model against those of generally accepted Bayesian updating principles governing rational belief revision in the face of new evidence. On the Bayesian view of things any individual $i$'s revised probability for some event $X$, $q_i(X)$, after having observed the prior probabilities of the other individuals, should equal her conditional probability for $X$, given these probability judgements, i.e.:

$$q_i(X) = p_i(X|p_1(X) = x_1, p_2(X) = x_2, ..., p_n(X) = x_n)$$

In the simplest case of just two individuals, $i$ and $j$, the dual constraints of the Bayesian conditioning and respect-weighted averaging yields:

$$p_i(X|p_j(X) = x_j) = wp_j(X) + (1 - w)p_i(X)$$

where $w$ is $i$'s respect weight on $j$'s probability judgement on $X$. The relation expressed here is potentially very useful. In the form given above it 'tells' $i$ how to form her conditional probabilities in the light of her respect for $j$'s judgements, thereby offering a solution to the hard problem of how to determine conditional probabilities given the testimony of others. On the other hand, by reorganisation we can derive the respect weights from the posterior probabilities:

$$w = \frac{p_i(X|p_j(X) = x_j) - p_i(X)}{p_j(X) - p_i(X)}$$

In case $i$'s conditional probabilities for $X$ given $j$'s expressed probability for $X$ equals $x_j$, $j$'s prior for $X$, her respect for $j$ must be at the maximum of one. And vice versa. At the other extreme, if her conditional probabilities for $X$ given $j$'s judgements just equals her prior for $X$, her respect for $j$ is zero. Furthermore by Bayes' Theorem :

$$p_i(X|p_j(X) = x_j) = \frac{p_i(p_j(X) = x_j|X)}{p_i(p_j(X) = x_j)} p_i(X)$$

So $p_i(X|p_j(X) = x_j) = p_i(X)$ just in case $p_i(p_j(X) = x_j|X) = p_i(p_j(X) = x_j)$, i.e. whenever $j$'s probability judgements are independent of the truth. So it seems that a zero respect weight on someone else's probabilities coincides with the judgement that they are probabilistically independent of the truth; a useful result.

So far so good. However a couple of problems emerge even in this simple case.

1. Suppose that $i$ and $j$ have the same beliefs about $X$ at a particular point in time, but that $j$ is subsequently able to make an additional relevant observation. If $j$ now declares his new probabilities, how should $i$ respond? Intuitively, and provided that $i$ does not doubt $j$'s powers of observation, $i$ should simply adopt $j$'s new probabilities as her own. But this is tantamount to zero weighting her own judgement, which in the light of the preceding judgement would seem to be equivalent to regarding her own judgements as probabilistically independent of the truth. But $i$ may very well regard her judgements as perfectly good, even if not as well-informed as $j$'s. So intuitively adopting someone else's probabilities does not commit one to the view that one's own judgements are independent of the truth.

2. Suppose that $p_i(X|p_j(X) = x_j) < p_i(X)$ because, for instance, $i$ regards $j$'s judgements as systematically biased in some way. In this case $i$'s respect weight for $j$ should be *negative*. But this cannot be the case in the DLW model where weights are assumed to be non-negative.

A more serious difficulty for the reconciliation of Bayesian revision and linear averaging emerges in when we consider larger groups. Suppose we have three individuals $i$, $j$ and $k$. Then Bayesian updating requires that:

$$\begin{aligned} q_i(X) &= p_i(X|p_j(X) = x_j, p_k(X) = x_k) \\ &= \frac{p_i(p_j(X) = x_j|p_k(X) = x_k, X).p_i(p_k(X) = x_k|X)}{p_i(p_j(X) = x_j|p_k(X) = x_k).p_i(p_k(X) = x_k)} p_i(X) \end{aligned}$$

The factor of interest here is the probabilistic independence or otherwise of the judgements of individuals $j$ and $k$ and its significance for $i$. In case they are independent we obtain:

$$q_i(X) = \frac{p_i(p_j(X) = x_j|X).p_i(p_k(X) = x_k|X)}{p_i(p_j(X) = x_j).p_i(p_k(X) = x_k)} p_i(X)$$

On the other hand when $k$ and $j$'s judgement on $X$ is perfectly correlated in the sense that $p_i(p_j(X) = x_j|p_k(X) = x_k, X) = p_i(p_j(X) = x_j|p_k(X) = x_k) = 1$, we obtain:

$$q_i(X) = \frac{p_i(p_k(X) = x_k|X)}{p_i(p_k(X) = x_k)} p_i(X)$$

Clearly $i$'s posterior probabilities for $X$ will generally differ in these two cases; indeed they only agree when $j$'s judgement on $X$ is independent of its truth. So we must conclude:

**Proposition 3** *On the Bayesian account it cannot be the case that someone's posterior probabilities, given the judgements of others, depends only on these judgements and the epistemic weight that they attach to them.*

The fact that the method of linear averaging ignores the interdependence of expressed judgements is a significant weakness. To see how this can lead us astray, compare a situation in which two scientists conduct separate experiments to try and settle some question with one in which they conduct a single experiment together. Suppose that in both cases the scientists report that as a result of their experiments they consider $X$ to be highly probable. In the former case, we would probably want to considerably raise our own probability for $X$ because of the convergence of expert testimony. In the latter case too we would want to raise our probability for $X$, but less so, because their joint testimony in favour of $X$ is based on same information. To revise once in the light of the testimony of the first scientist and then again in the light of that of the second would in effect be to update twice on the same evidence, akin to an individual scientist conditioning twice on the same experimental result.

Two things follow from this discussion. Firstly since it has not been established that respect weighted averaging is the uniquely rational way of revising belief in the light of differences of opinion, it cannot be claimed that a consensus is the inevitable outcome of rational individual learning. Secondly, since linear averaging does not take into account interdependencies between the judgements of the different individuals making up the group, it cannot be claimed that it produces an optimal summary of the information contained in these judgements. So whatever the merits of the DLW model it does not solve the problem we set ourselves.

# 5 The Robustness of Consensus

The inadequacy of the DLW model leaves open the question of the extent to which rational revisions will tend to produce a consensus and hence the extent to which aggregation can hope to summarise the information contained in the diverse opinions of individuals by simple adoption of any consensual judgements that it produces. In this final section, I will assume that some revision method may be found that is at least partially successful in producing consensus and show that any aggregation method that is successful in exploiting the information contained in he consensual judgements will violate some frequently invokes conditions on aggregation.

It is easy enough to establish this claim for the condition of independence of irrelevant alternatives. For it follows immediately from the claim that revision of opinion should be driven by the judgements individuals make about the epistemic authority of others, that any aggregation function that was sensitive to individual learning should determine a value for the collective judgement on some prospect that depends on what judgements individuals reach about each other's authority with respect to that prospect. Individual opinion on the quality of others' judgements is not irrelevant therefore.

This conclusion may not be of serious concern to aggregation theorist, since there are already many other reasons for questioning the independence of irrelevant alternatives condition. What is much more significant is that the unanimity preservation condition does not stand up to this line of questioning. To make our point more precise let us start with a definition of the robustness of an aggregation principle that is intended to capture the idea that features of the aggregate judgements derived from the initial state of opinion that are required by a principle should not be disallowed by the principle when applied to the state of opinion reached by rational revision of individual opinion.

**Definition 4** *An aggregation principle (or set of them) will be said to be minimally robust with respect to rational revision of opinion just in case, for any state of opinion, S, the set of profiles of aggregate judgements derived from S and consistent with the principle (or set of them) has a non-empty intersection with the set of profiles of aggregate judgements consistent with the principle (or set of them) derived from any state of opinion S' reached by rational individual revision in the light of information contained in S.*

The definition of minimal robustness leaves unspecified what rationality requires of revision of opinion. If one assumes, as I have done, that the Bayesian conditioning model provides the basic normative standard for the revision of probabilities then the rule of 'zero unanimity' - that if all individuals attach zero probability to a prospect then the aggregate probability for this prospect should be zero too - is minimally robust. The same is not true, however, of the more demanding unanimity preservation condition mentioned earlier, that requires that if all any individuals agree on the probability of some prospect then this consensual judgement should be adopted as the aggregate one. On the contrary:

**Proposition 5** *The condition of unanimity preservation for probability judgements is not (even) minimally robust with respect to rational revision of probabilities.*

The proposition can be established without making any controversial assumptions about what rationality requires (in particular, there is no need to invoke the Bayesian model). Consider the following example. Suppose that Anne has observed that $A$ is true, but thinks that $B$ is improbable, while Bob has observed that $B$ is true, but thinks that $A$ is improbable. Suppose that their degrees of belief are more precisely represented as follows:

| | $AB$ | $A\neg B$ | $\neg AB$ | $\neg A\neg B$ |
|---|---|---|---|---|
| Anne | 0.1 | 0.9 | 0 | 0 |
| Bob | 0.1 | 0 | 0.9 | 0 |
| Aggregate | 0.1 | ? | ? | 0 |

By application of the unanimity preservation condition to the initial state of opinion, the aggregate probability for $AB$ must be 0.1 and for $\neg A \neg B$ it must be 0. Now suppose that Anne knows that Bob has observed whether $B$ is true and that Bob knows that Anne has observed whether $A$ is true. Then Anne should infer from Bob's expressed beliefs that $B$ is true, and Bob should infer from what Anne says that $A$ is true. After they revise their degrees of belief, the new state of opinion is:

| | $AB$ | $A\neg B$ | $\neg AB$ | $\neg A \neg B$ |
|---|---|---|---|---|
| Anne | 1 | 0 | 0 | 0 |
| Bob | 1 | 0 | 0 | 0 |
| Aggregate | 1 | 0 | 0 | 0 |

Unanimity preservation now requires that the aggregate probability for $AB$ be 1 and for all other possibilities be 0. So unanimity preservation is not minimally robust.

The point can be made for other types of opinion too. Consider the following example for preferences. John and Jane both prefer restaurant X to restaurant Y. So unanimity preservation requires an aggregate preference for X over Y. Now suppose that John knows that Jane has no taste and forms her preferences on the basis of cost alone. Jane on the other hand knows that John has no budget constraints and forms his preferences on the basis of quality alone. John knows that Jane has been to restaurant X but not to Y, while Jane knows that John has been to Y but not to X. John concludes that X is cheap. Jane concludes that Y is expensive. If both now revise their preferences to take into account what they have learnt, the new state of opinion will be one in which both John and Jane prefer restaurant Y to restaurant X. Unanimity preservation now requires an aggregate preference for Y over X: the exact reverse of what it required initially. So the principle that unanimous preferences should be preserved is not minimally robust in the face of rational preference revision.[5]

## 6 Conclusion

In this paper I have argued that diversity of opinion presents both a problem for collective decision making and an opportunity for learning and the improvement of opinion. This raises the question of whether it is possible to resolve the aggregation problem associated with the former in a way which optimally exploits the revisions of opinion associated with the latter. The question is a pressing one for aggregation theory because an aggregation method that is insensitive to the information present in the judgements of individuals risks finding that the collective judgements that it determines are undermined by the revisions made (rationally) in response to this information.

[5] To make this argument more precise, a theory of preference revision would have to be spelled out (see, for instance [4]). But such a theory should vindicate the revisions assumed in the example (on pain of being a poor theory).

The focus of the investigation of the question has been the prospects for an informationally sensitive aggregation procedure that works by adopting as the contents of the collective judgement any unanimous opinions achieved by rational revision. Two claims were made in this regard. Firstly I argued that, contrary to the DLW model, rational revision could not be guaranteed to lead to consensus. Hence, we have no reason to think that the proposed method will completely determine collective judgements. And secondly, that to the extent that the method can be applied it will lead to values for collective judgements that are in conflict with principles of standard aggregation theory, including the independence of irrelevant alternatives and the unanimity preservation principle itself. This suggestion that the demands expressed by common aggregation principles are not completely compatible with the epistemic demands driving rational revision and that some trade-off between them must be accepted.

# References


[1] D'Aspremont. C. and Gevers, L. (1977) "Equity and the Informational Basis of Collective Choice", *Review of Economic Studies* 44: 199-209

[2] Broome, J. (1990) "Bolker-Jeffrey Expected Utility Theory and Axiomatic Utilitarianism", *Review of Economic Studies* 57: 477-502

[3] Berger, R. L. (1981) "A Necessary and Sufficient Condition for Reaching Consensus by De Groot's Method", *J. Amer. Statis. Assoc.* 76: 415-418

[4] Bradley, R (forthcoming) "The Kinematics of Belief and Desire", *Synthese*

[5] De Groot, M. H. (1974) "Reaching a Consensus", *J. Amer. Statis. Assoc.* 69: 118-121

[6] Dietrich, F. and List, C. (forthcoming) "Arrow's Theorem in Judgement Aggregation", *Social Choice and Welfare*

[7] Gärdenfors, P. (1988) *Knowledge in Flux.* Cambridge MA: MIT Press.

[8] Genest, C. and J. V. Zidek (1986) "Combining Probability Distributions: A Critique and Annotated Bibliography", *Statistical Science* 1 (1): 113-135

[9] Gevers, L. (1979) "On Interpersonal Comparability and Social Welfare Orderings", *Econometrica* 47: 75-89

[10] Goodin, R. E. (2001) "Consensus Interruptus", *The Journal of Ethics* 5: 121-131

[11] Joyce, J. (2003) "Bayes Theorem", *The Stanford Encyclopedia of Philosophy*, E. Zalta (ed.), URL = http://plato.stanford.edu/entries/bayes-theorem/#4.

[12] Lehrer, K and C. Wagner (1981) *Rational Consensus in Science and Society*, Dordrecht: Reidel

[13] Levi, I. (1985) "Pareto-unanimity and Consensus", *Journal of Philosophy* 87

[14] List, C. and P. Pettit (2002) "Aggregating Sets of Judgments: An Impossibility Result", *Economics and Philosophy* 18: 89-110

[15] Loewer, B. and Laddager, R. (1985) "Destroying the Consensus", *Synthese* 62: 3-11

[16] McConway (1981) "Marginalization and Linear Opinion Pools", *Journal of American Statistical Association* 76: 410-14

[17] Mongin, P. (1995) "The Pareto Principle and Spurious Agreement" *LSE Choice Group Working Papers* 1 (4)

[18] Roberts, K. (1980) "Interpersonal Comparability and Social Choice Theory", *Review of Economic Studies* 47: 421-439

[19] Sen, A. (1982) *Collective Choice and Social Welfare*, Edinburgh: Oliver and Boyd

[20] Wagner, C. (1978) "Consensus through Respect: A Model of Rational Group Decision-making", *Philosophical Studies* 34: 335-349

[21] Wagner, C. (1982) "Allocation, Lehrer Models, and the Consensus of Probabilities", *Theory and Decision* 14: 207-220

[22] Wagner, C. (1985) "On the Formal Properties of Weighted Averaging as a Method of Aggregation", *Synthese* 62: 97-108