**Penultimate draft of DOI:10.1007/s11098-009-9362-1**
**The original publication is available at**
**http://www.springerlink.com/content/586830751765pgx4/**

# Disagreement, Equal Weight and Commutativity

Alastair Wilson
University College, Oxford

How should we respond to cases of disagreement where two epistemic agents have the same evidence but come to different conclusions? Adam Elga has provided a Bayesian framework for addressing this question[i]. In this paper, I shall highlight two unfortunate consequences of this framework, which Elga does not anticipate.

Elga calls his proposal the 'equal weight view', and gives the following formulation:

> Equal weight view:
> Upon finding out that an advisor disagrees, your probability that you are right should equal your prior conditional probability that you would be right. Prior to what? Prior to your thinking through the disputed issue, and finding out what the advisor thinks of it. Conditional on what? On whatever you have learned about the circumstances of the disagreement.
>
> Elga (ibid.), p.490

This proposal, puzzlingly, seems to amount to nothing more than the following: 'conditionalize on new evidence, where the evidence is information about the circumstances of disagreement'. This advice corresponds exactly to standard conditionalization[ii]. By itself, it does not deserve to be called the 'equal weight view': the 'conditionalization view' might be a better name. To see why, notice that it's entirely compatible with Elga's formulation that your prior conditional credence distribution should amount to one of the following: 'in all cases of disagreement with agent A, whatever the circumstances, stick to your original credences', or 'in all cases of disagreement with agent B, whatever the circumstances, believe that B is correct'. There is no equal weighting involved in these distributions.

So why does Elga give his view the name that he does? The puzzle is resolved by noticing that the 'equal weighting' element of the view depends on our treating an advisor as an 'epistemic peer'. The notion of an epistemic peer is in fact the key element of the equal weight view. Elga defines it as follows:

> On my usage, you count your friend as an epistemic peer with respect to an about-to-be-judged claim if and only if you think that, conditional on the two of you disagreeing about the claim, the two of you are equally likely to be mistaken.
>
> Elga (ibid.), p.487 n.21

A consequence of this is that, if you treat someone as an epistemic peer, it must be the case that your prior credence in their being right, conditional on a neutral disagreement (one with no special factors that make you think either of you more or less likely to be right than usual), is 50%. Elga explicitly acknowledges this elsewhere:

> Suppose that before evaluating a claim, you think that you and your friend are equally likely to evaluate it correctly. When you find out that your friend disagrees with your verdict, how likely should you think it that you are correct? The equal weight view says: 50%.
>
> Elga (ibid.), p.488

The general strategy is therefore that in cases of neutral disagreement with an epistemic peer (with respect to a particular proposition P) we should adjust our credence in P so as to meet the peer halfway. Thus, if I judge it to be 30% likely that the world will end tomorrow, and a peer judges it to be 70% likely, and we each know, say, the contents of every newspaper yet published (and nothing more), we should each end up with 50% credence in the world ending tomorrow.

Could Elga deny that he is committed to this sort of 'averaging' view? I have two main reasons for thinking that he is indeed committed to it. Firstly, the following passage strongly suggests the view:

> ...note that one might have differing assessments of an advisor's abilities with respect to different issues. For example, one might count an advisor as a peer with respect to arithmetic, but as less than a peer with respect to disputes about euthanasia. So despite the name, the equal weight view does not in general call for simply averaging together one's probability function with that of one's advisor.
>
> Elga (ibid.), p.489

Elga's reason for claiming that the equal weight view does not in general call for averaging one's entire credence functions with that of a peer is only that one may count an advisor as a peer with respect to some issues but not others. I may rate you as a peer when it comes to predicting football results, but not when it comes to predicting election results. But when this kind of reason is absent – when two agents count themselves as peers with respect to every issue – Elga clearly recommends averaging.

Moreover, note that even if we do not count an advisor as a peer on all issues, we can have disagreements with them which relate to only those issues about which we do count them a peer. In such a case, a natural extrapolation of the equal weight view calls for averaging with the advisor that part of our credence function which relates to the issue in question.

Another reason for thinking that Elga is committed to the averaging view is that he gives us no other account of how to incorporate equal weighting into our credences. That is to say, it is unclear that the equal weight view recommends a positive strategy at all, if it does not recommend averaging your credences with an advisor (at least about the issues with respect to which you consider that advisor an epistemic peer.)

Interpreted as an (at least partial) averaging proposal, the equal weight view has plenty of initial plausibility. If I treat someone as an epistemic peer, I should respect his credences exactly as much as my own; and a natural Bayesian implementation of this respect is to meet the peer's credences halfway. But this plausibility is easily undermined, as I shall now argue.

In general epistemic agents who endorse the equal weight view will need to update their credences both by averaging out credences with epistemic peers when they learn of disagreements and by conditionalization on other evidence. The problem to which I want to draw attention concerns the interaction between evidence of disagreement and other evidence. In general the combination of conditionalization on other evidence and application of the equal weight view by epistemic peers will not satisfy the requirement of commutativity; that is, the credences agents end up with will depend on the *order* in which they receive new evidence and learn of disagreements. This has long been recognized in the literature on

Bayesian updating and disagreement; however, the examples to follow aim to show that non-commutativity is difficult to live with.

To get a clearer idea of non-commutativity, consider the analogy of a price in a shop that has been reduced in a sale. A tree initially on sale for $100 is marked 'reduced by $5', since it has finished blossoming, and 'reduced by 20%', since the January sales are in progress. In which order should the reductions be applied? If the price is first reduced by 20%, and then by $5, the resultant price is $75; if the $5 reduction comes first, the resultant price is $76. The two processes of addition and multiplication are, like conditionalization and averaging of credences, not commutative. In the case of the tree, the shop can freely choose to charge the higher price. Unfortunately, non-commutativity raises problems in the epistemic case, as the following example shows.

Two agents, who consider each other epistemic peers and have equal access to information about coin tossing dynamics and bias distributions, come to different conclusions about the bias of a coin. One thinks that a bias towards heads is 80% likely, the other thinks that it is 20% likely. They each are sure that the coin is either fair or biased in such a way that it lands heads every time. Now at $t_1$ the toin is tossed, and lands heads, and the result is shown to the pair; at $t_2$, the two confer and average out their credences in accordance with the equal weight view. The equal weight view, due to the failure of commutativity, predicts a difference in final credence for the agents between cases in which $t_1$ is earlier than $t_2$, and cases in which the order of these two times is reversed.

Consider first the limiting case where one agent begins with credence 1 that the coin is biased, and the other with credence 0. In this case, conditionalization will not alter either agent's credence at all if it is applied before the equal weight view, but will alter it if applied afterwards. The limiting case does not raise a direct problem for Elga's proposal, since Elga takes evidence to be any proposition given credence 1, and by hypothesis the agents have exactly the same evidence available. However, the same kind of problem arises for cases other than the limiting case, as I will now argue.

Assume one agent begins with credence 0.8 in a bias and the other with credence 0.2. If the agents compare first, their credences in a bias are averaged to 0.5 according to the equal weight view. If a heads result is then observed, conditionalization excludes half of the possible cases of fairness, so a bias will now seem twice as likely to them as fairness, and the agents' credences in the bias increase as follows:

$$0.5 \text{ goes to } \frac{0.5}{(0.5+0.5(1-0.5)} = 0.666\ldots$$

If the agents flip first, and conditionalize on a result of heads, then their credences in the bias change as follows:

$$0.8 \text{ goes to } \frac{0.8}{(0.8+0.5(1-0.8)} = 0.888\ldots \qquad 0.2 \text{ goes to } \frac{0.2}{(0.2+0.5(1-0.2)} = 0.333\ldots$$

Applying the equal weight view and taking the average of these, the agents' final credences that the coin is biased towards heads will be 0.6111111. For other initial credence distributions that average to 0.5, the following results are obtained:

0.6 → 0.75, 0.4 → 0.5701, average is 0.660

0.9 → 0.9473, 0.1→ 0.1818, average is 0.565

0.99 → 0.9945, 0.01 → 0.0202, average is 0.508

If the agents compare views first, their final credences always end up as 0.666..., whatever their initial divergence. If the agents flip first, their final credences will vary between .5 and .666… depending on how far their initial credences diverged. Hence the end result when comparing first will always be different from the end result when flipping first, unless the agents start off with the same credence; in this case, there is no trouble, as there is no disagreement in the first place. The greater the disagreement, the greater the problem the failure of commutativity generates for the equal weight view.

This failure of commutativity seems troubling, but perhaps we could learn to live with the dependency of our credences on the order in which we receive evidence. Unfortunately, there is a further problem to worry about. When $t_1 = t_2$ (that is, when one simultaneously learns of the disagreement and obtains pertinent evidence) the value predicted by the equal weight view for the final credences is undefined, since the order in which the two processes should be applied is unconstrained. An agent will not be uniquely guided by the equal weight view if he cannot distinguish any time difference between learning the result of the coin toss and learning his peer's original opinion about the bias.

We cannot avoid this second problem by simply denying that it is possible to discover a disagreement and learn of the result of a coin flip at *exactly* the same time. In certain cases, learning of a disagreement is itself a source of evidence, so that conditionalization and the averaging of credences have to operate simultaneously. The problem can best be illuminated with another example. Two agents, who have the same evidence and take each other to be epistemic peers, are wondering whether or not they disagree about more than one thing at time t[iii]. At t, one agent thinks the probability of more-than-single-disagreement is 0.1; the other thinks the probability is 0.9. They voice their opinions on this matter at time t+1 and each decides to average his credence with his peer, in accordance with the equal weight view. But at t+1 they have also found through comparing notes that they disagreed about at least one thing (the probability in question) at t; so both of their credences in more-than-single-disagreement at t should go up via conditionalization. Now the question is – should conditionalization on this new evidence occur before or after the application of the equal weight view? Whether it happens before or after, the pair will end up agreeing on the probability of more-than-single-disagreement. However, the actual credence that the pair of them ends up with will be different depending on the order of application. Given that the two processes in fact occur simultaneously, it looks like there can be no good reason to prefer one order over another[iv].

Note that the problem generalizes to cases of disagreement with agents other than those we count as epistemic peers. If I have any respect whatsoever for an agent's epistemic abilities, and if we share all the relevant evidence, Elga's model suggests that in cases of disagreement I should adjust my credences towards his to some degree or other. But *any* adjustment of this sort is enough to provide a commutativity problem with conditionalization. The only special cases of the 'equal weight view' that avoid this problem are the cases in which you ignore an agent completely, or in which you take their credences to be always correct.

I agree with the intuition behind Elga's proposal, that partial deference is the right approach to take in actual cases of disagreement. We should certainly take others' views into account, and not dogmatically prioritize our own credences. But unless we can live with both the non-commutativity and the indeterminacy that I have alluded to, this deference needs to be modelled in a different way from Elga's proposal[v][vi].

[i] Adam Elga, 'Reflection and Disagreement', *Nous* 41:3 pp.478-502 (2007).

[ii] Elga does specify a coarse-graining constraint on circumstances of disagreement, but this won't be relevant to my discussion.

[iii] In the relevant sense, two agents disagree about more than one thing if there are two families of propositions, logically independent from each other, over which the two agents have different credence distributions. Obviously we can't say that disagreeing about more than one thing is disagreeing about more than one proposition, since if two agents have different credences in p they will *ipso facto* have different credences in not-p.

[iv] One possible response would be to take the average of the credences resulting from applying the processes in different orders. Another would be to insist that one of the two processes should always be applied first. Without independent motivation, such proposals seem unacceptably *ad hoc*.

[v] I have more general concerns about Elga's proposal, since it applies Bayesian methods, which assume logical infallibility, to scenarios more naturally thought of as involving imperfect reasoning. However, I will not pursue these concerns here.

[vi] Thanks to an audience at Oxford, and particularly to Frank Arntzenius and John Hawthorne.