Epistemic Compromise between Difference Splitting and Scoring Rules:

A Response to Moss

Armin Schulz

Department of Philosophy

University of Wisconsin-Madison

5185 Helen C. White Hall

600 N Park St.

Madison, WI 53706

Epistemic Compromise between Difference Splitting and Scoring Rules: A Response to Moss

Sarah Moss advocates adopting a novel approach towards the determination of the perfect epistemic comprise – one based on scoring rules (SR's). She makes her case lucidly, presents interesting and innovative arguments, and is to be commended for advocating the use of formal tools to push the investigation of traditional epistemological problems forward.

However, upon closer scrutiny, it turns out to be less than fully clear what exactly her contribution to the literature on epistemic compromise amounts to. In particular, this comment attempts to show that, on the one hand, Sarah fails to give entirely convincing reasons for *abandoning* the traditional approach towards epistemic compromise, but that on the other, she does manage to place that approach *on a firmer theoretical foundation*. While falling short of the aims she has set herself, her argument thus nonetheless deserves to be taken very seriously.

This comment is structured as follows. Section I lays out the key aspects of Sarah's paper. Section II critically discusses the reasons she gives for adopting the SR approach. Section III concludes.

I. Moss on Scoring Rules and Epistemic Compromise[1]

The overarching concern of Sarah's argument is the epistemology of disagreement: what credence ought two epistemic peers (two relevantly similar epistemic agents with the same evidence) adopt when their individual credences diverge? Sarah notes that this question can be broken into two parts: firstly, what *is* the perfect epistemic compromise? Secondly, what role

[1] Note that, since the aim in this section is only to provide a foundation for the discussion to follow, not all the issues of interest in Sarah's paper will be considered. Note also that, throughout this paper, page numbers without identifying reference point to Sarah's manuscript, available at http://socrates.berkeley.edu/~fitelson/few/moss.pdf.

ought that compromise play in the agents' further epistemic actions? In laying out her argument, Sarah concentrates on the second question, and this comment will follow her in that.

When it comes to answering that question, the traditional way to proceed is by appealing to 'splitting the difference': if my credence in some proposition p is x, and yours is y, the perfect compromise credence is (x+y)/2. Sarah disagrees with this answer, and sets out to provide an alternative; in particular, she sets herself two explicit aims (p. 2): (1) To show that using SR's is a *coherent* way to determine the perfect epistemic compromise. (2) To show that we ought to *adopt* scoring rules to do so. Her argument towards these aims proceeds as follows.

Sarah begins by defining the expected epistemic value of credence x in a proposition p that is actually believed to degree □ as follows:

$$EV(x) = \square\, f_1(x) + (1-\square)\, f_0(x).$$

Here, $f_0(x)$ is the epistemic value of credence x in p if p is *false*, and $f_1(x)$ is the epistemic value of credence x in p if p is *true*. The two functions $f_0(x)$ and $f_1(x)$ together make up the agent's *scoring rule*. Sarah also suggests that that this approach can be easily extended to an *algebra* P of propositions by summing the EV's of all the *atomic propositions* of P.

She then lays out an important distinction among SR's: that between *credence-eliciting* and *non-credence-eliciting* ones. The former are such that the credence that maximises EV(x) is always the credence that the agent actually has; the latter, by contrast, may be maximised at values other than the actual credence of the agent. What is important about this distinction here is that non-credence eliciting SR's are plausibly seen to be *irrational*: they can lead to the updating

of beliefs even in the absence of good reasons (e.g. without there being novel evidence).[2] This, together with the fact that the entailed belief changes can be *large* and *cyclical*,[3] mean that non-credence eliciting SR's hold only a very dubious epistemological status.

Using these tools, Sarah argues that the perfect epistemic compromise is determined by *maximising the average of the expected epistemic values* of the agents in question. She puts forward three justifications to support this argument (pp. 11-13): firstly, doing so maximises the joint epistemic values of the two argents, and thus gives them something that they intrinsically epistemically care about. Secondly, it fits well to what we do in cases of *joint action*, where Sarah takes it to be 'intuitively reasonable' (p. 12) that we maximise our average expected utilities. Thirdly, it has the benefit of leading to outcomes that are not *dominated* in epistemic value (in what follows, this will be referred to as the 'dominance principle').

Overall, therefore, Sarah suggests and seeks to justify an alternative route towards the determination of the perfect epistemic compromise – one based on SR's. However, as the next section tries to show, more work is needed in order to make her argument fully compelling.

II. Questions about the Scoring Rule Approach towards Epistemic Compromise

What needs to be asked about Sarah's argument is to what extent it has managed to fulfil the two aims she has set herself – after all, these aims define her inquiry, and can thus be used as a standard with which its success is assessed. In this, most of the attention is on aim (2), but some remarks about (1) are useful, too. Consider them in turn.

[2] It is not entirely obvious that Sarah fully endorses this statement (see e.g. p. 5), but be that as it may: it is sufficient that there are good reasons (acknowledged by Sarah) that suggest that having these kinds of SR's is irrational.
[3] For example, Sarah's 'Brie cubed' SR maximises epistemic value at $x = 0$ or $x = 1/2$ if the actual credence in p is 1; however, *starting* with a credence of 1/2, epistemic value is maximised for $x = 1$.

1. The Coherence of the Scoring Rule Approach[4]

Abstractly, Sarah's SR approach towards epistemic compromise takes individual judgements of epistemic value and aggregates them to construct a collective suchlike judgement. The trouble with this is that there are numerous results suggesting that this kind of judgment aggregation is in fact inconsistent with certain reasonable-sounding assumptions (see e.g. Seidenfeld et al., 1989; List & Pettit, 2002, and Arrow, 1963). In particular, Seidenfeld et al. (1989) have shown that *no* linear aggregation rule – Sarah's included – is consistent with a *unanimity principle*: if both agents value credence x in p more than credence y, the perfect compromise should reflect this.

This is troubling, as the satisfaction of such a principle, at least on the face of it, seems to be a minimal requirement for any acceptable account of epistemic compromise. The existence of Seidenfeld et al's (1989) impossibility result therefore appears to pose severe problems for Sarah's approach: if appealing to SR's is internally incoherent, then it cannot be a credible alternative to splitting the difference. However, there are two reasons to think that the impact of these problems will be limited in the present context.

Firstly, the standard way of responding to impossibility theorems of this form is by making clear why one or more of the assumptions in the background of the theorem do not apply to the case at hand (see e.g. Seidenfeld et al., 1989, List & Pettit, 2002). Furthermore, there is no reason to think that Sarah's approach is any worse off than similar ones in other contexts: for example, it may well be possible to justify failures of unanimity by noting that *epistemic compromises* are not subject to the same kinds of rationality constraints as *agents* are. In this way, the Seidenfeld et al. (1989) result and other impossibility theorems like it need not pose an insurmountable hurdle for Sarah's account.

[4] The discussion in this section was prompted by and has profited from some useful correspondence with Teddy Seidenfeld.

Secondly, it seems that this debate is anyway in many ways tangential to the main thrust of Sarah's argument. To see this, note that her argument retains interest even if formulated conditionally: that is, it is still philosophically important to find out whether it is true that, *if* it is possible to aggregate individual epistemic judgments in the way suggested by Sarah, *then* we ought to use SR's to determine the perfect epistemic compromise. For this reason, the existence of the present sort of worry need not interfere with an evaluation of the rest of her argument (see also Sober, 1981).

Note that none of this is meant to downplay the importance of the difficulties raised by the impossibility results. The point is just that they should not be seen as making a proper engagement with Sarah's discussion worthless.[5]

2. The Justification of the Scoring Rule Approach towards Epistemic Compromise

When it comes to Sarah's justification for the switch to the SR account of epistemic compromise, there are two cases that need to be distinguished: on the one had, the two agents may have the same SR, one the other, they have different ones.[6] Consider them in turn.

a. Same Scoring Rule

If the two agents have the same SR f, that rule can be either credence-eliciting or not. Again, consider the two cases in turn.

[5] Another interesting issue surrounding the coherence of her approach concerns the assumptions about a 'universal logical atomism' she needs to make. Remember that the approach takes the expected epistemic value of an algebra of propositions to be the sum of the expected epistemic values of the atomic propositions spanning it. However, it is not clear that agents who individuate the world in fundamentally different ways, and who thus have fundamentally different atomic propositions at the base of their algebra of believed propositions, can be easily compared. For example, it is not clear that it is meaningful to determine the perfect epistemic compromise between an agent with two atomic propositions and one with two thousand atomic propositions *in the way Sarah suggests*. However, for the same reasons as the ones cited in the text, I shall not discuss this worry further here.

[6] This case-distinction, as well as the one following it, is also implicit in Sarah's own presentation (see e.g. pp. 7-9).

a_1. f is credence-eliciting

In this case, splitting the difference and the SR approach yield the same compromise credences (pp. 7-8). Since all of the major reasons Sarah cites for adopting the SR approach are 'outcome-focused' – i.e. since they solely concern properties of the resulting compromise (e.g. the fact that it maximises epistemic value, or that it satisfies the dominance principle) – this means that she cannot provide grounds for favouring that approach over splitting the difference.

Moreover, if reasons other than outcome-focused ones are considered, the situation looks to actually get worse for the SR approach: for example, as noted by Sarah herself (see e.g. p. 6), the *practical applicability* of her approach is considerably lower than that of difference-splitting, while its *computational complexity* is greater. Overall, this thus means that, at the very least, this case does not favour the SR approach over splitting the difference.

a_2. f is non-credence-eliciting

As noted earlier, non-credence-eliciting scoring rules are commonly taken to be *irrational*. If adopting Sarah's treatment were to rest on this case, therefore, it would take the following form: *if two epistemically irrational agents were to disagree, then the rational way to determine the prefect epistemic compromise between them is by means of SR's*.

If not outright incoherent, this is at least highly peculiar: it seems clear that one should not rest an epistemological theory on situations that are judged to be epistemically defective. Just like one would not want to rest a defence of an account of rational belief revision that is meant to apply *generally* on what inferences are rational for *pathologically deluded* people, one should not rest an argument about what makes for a perfect epistemic compromise on what is rational for people with irrational epistemic values.

b. Different Scoring Rules[7]

If the two agents have different SR's, then in order for Sarah's approach to get off the ground, their SR's need to be *unit comparable*: they must not differ by more than an additive constant (see e.g. p. 18; Seidenfeld et al., 1989). However, this kind of unit comparability is, in fact, unlikely to be given.

To see this, begin by noting that Sarah justifies the unit comparability of epistemic values by appeal to the (purported) fact that it is 'intuitively reasonable' that we maximise average expected *utilities* when making joint decisions (p. 12). This argument is quite puzzling, though, since expected utilities are in fact notorious for *lacking* a justification for being unit comparable. Furthermore, there is every reason to think that the situation with respect to epistemic values is no better than that of the more standard 'practical utilities'. This can be brought out easily by briefly considering the three common ways of arguing for unit comparability.[8]

Firstly, one could claim that we commonly act 'as if' that comparability were possible, and that this is all there is to this question (this may be what Sarah has in mind, too – see e.g. p. 12). That is, one might reason that, since we commonly behave in ways that suggest that epistemic values are unit comparable, these values in fact *are* unit comparable. Put this way, though, it is easy to see that this argument is fallacious: facts about how we act do not entail facts about how *justified* we are in acting in this way. The point at issue here is an *evaluative* one: we are interested in whether we are *justified* in taking ourselves to be unit comparable, not in whether we *in fact take ourselves to be so* (see also Goldman, 1995).

[7] Technically, one would also have to distinguish credence-eliciting and non-credence eliciting SR's here; however, since the discussion in the latter case mirrors that where the two agents have the same SR, it can be suppressed here.
[8] For more on interpersonal comparisons, see e.g. Elster & Roemer (1991), Goldman (1995) and Hausman (1995). I also thank Mauro Rossi for many useful discussions about this issue.

Secondly, one could take 'epistemic values' to be objectively existing mental states with their own proprietary units, and assume that the epistemic agents in question are psychologically similar. The problem with this is that we have absolutely no reason to think that this is true: all the evidence that bears on this question is indeterminate between unit comparability and the lack thereof (see also Goldman, 1995). This route is thus either question-begging or without the required empirical support.

Thirdly, one may think that unit comparability is part of the (tacit) theory with which we assign epistemic values and credences to another agent (see also Goldman, 1995). The problem with this suggestion is that we are not just interested in the *content* of the theory with which we epistemically interpret each other (assuming there is one such), but in whether that theory is *true*. The situation here is unlike that in Chomskian linguistics, where a theory of this kind *constitutes* the ability to make grammaticality judgements (say): in the present case, the issue is not accounting for our *intuitions* about epistemic values, but finding out about the epistemic values *themselves*. Alas, we have absolutely no reason to think that our tacit theory about these epistemic values is true: as just noted, the evidence concerning this matter is completely indeterminate.

Thus, there seem to be no good reasons to think that SR's are unit comparable. However, if that is so, then there is no justification for using SR's to determine the perfect epistemic compromise when the two agents have different SR's.

Overall, therefore, Sarah's justification for switching to SR's over splitting the difference does not seem to rest on entirely solid argumentative grounds: the SR approach either does not improve on difference-splitting, or does not rest on epistemically acceptable considerations.

However, there may be another way of reading Sarah's argument that lets her avoid these worries.

3. Scoring Rules as a way of Grounding Difference Splitting

According to this re-interpretation of her argument, she ought to be seen, not as putting forward an alternative account of what the perfect compromise consists in, but as gesturing at the fact that SR's are well placed to *support* the existing approach. SR's are at the heart of a well-established research programme in statistics and philosophy, so that locating the literature on epistemic disagreement within the confines of that research programme could significantly strengthen its foundations.

In particular, the above argument makes clear that Sarah's approach is strongest when assuming that the two agents use the same, credence-eliciting SR. In that case, there are no problems with irrationality or interpersonal comparability to worry about, and the appeal to SR's brings out various beneficial features of the resulting compromise. Because of this, it may be best to take the theory to be resting on this case, and, through it, to *further support* the existing, difference-splitting approach to epistemic disagreement.

On this reading, Sarah's arguments bring out that this traditional approach rests on a solid theoretical foundation: it shows that it quite plausibly gets the basic epistemological facts right. By appealing to SR's, splitting the difference can be shown – at least in certain cases – to be based on a well-founded theoretical account of how epistemic values are being determined. In turn, this makes clear that it is more than an intuitively compelling way of finding epistemic compromises: it has profound theoretical support, too. Furthermore, the appeal to SR's also shows that splitting the difference yields compromises that maximise the epistemic value of the

two agents, that it cannot be dominated, and that it has many other beneficial properties.[9] On top of all of this, appealing to SR's allows for the exploration of many further features of these compromises (as, for example, issues surrounding imprecise credences – see e.g. pp. 14-16).

It needs to be noted that this defence of the traditional approach is limited: it rests on the two agents having the same credence-eliciting SR. However, this limitation should not be overstated: it may very well be that instances where the two agents share credence-eliciting SR's are widespread, and at any rate, even such a restricted defence can be very interesting.

On this reading, therefore, Sarah's argument provides valuable epistemological insights by using the formal apparatus of SR's to illuminate the reasons *why* the traditional account is the right approach to determining the perfect epistemic compromise. Of course, interpreted in this way, Sarah can no longer be seen to propose a new account of epistemic disagreement; however, despite this, she can still be taken to make a useful contribution to the literature surrounding that issue.

III. Conclusion

I have tried to argue that Sarah has not fully justified the switch towards a SR's account of epistemic compromise: this account either does not yield outcomes that improve on splitting the difference, or rests on questionable assumptions. I have then suggested that a different way of looking at Sarah's results may be more plausible: on this view, Sarah generalises the existing literature and connects it to a different field – which, despite leaving the existing epistemological literature largely intact, is a very useful outcome in its own right.

[9] Not all of these benefits may be able to be maintained after worries about the coherence of the approach have been addressed. However, some certainly might.

Bibliography

Arrow, Kenneth (1963). *Social Choice and Individual Values*. New York: John Wiley.

Elster, Jon and Roemer, John (eds.) (1991). *Interpersonal Comparisons of Well-Being*. Cambridge: Cambridge University Press.

Goldman, Alvin (1995). 'Simulation and Interpersonal Utility'. *Ethics* 105: 709-726.

Hausman, Daniel (1995). 'The Impossibility of Interpersonal Utility Comparisons'. *Mind* 104: 473-490.

List, Christian, and Pettit, Philip (2002). 'Aggregating Sets of Judgments: An Impossibility Result'. *Economics and Philosophy* 18: 89-110.

Moss, Sarah (2008). 'Scoring Rules and Epistemic Compromise'. Available at http://socrates.berkeley.edu/~fitelson/few/moss.pdf.

Seidenfeld, Teddy; Kadane, Joseph; and Schervish, Mark (1989). 'On the Shared Preferences of Two Bayesian Decision Makers'. *The Journal of Philosophy* 86: 225-244.

Sober, Elliott (1981). 'The Evolution of Rationality'. *Synthese* 46: 95-120