Studies in History and Philosophy of Science 41 (2010) 148–157

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Simulating peer disagreements

Igor Douven

Institute of Philosophy, University of Leuven, Kardinaal Mercierplein 2-bus 3200, B-3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 17 January 2009
Received in revised form 23 November 2009

Keywords:
Peer disagreement
Computer simulations
Opinion dynamics
Hegselmann–Krause model
Social epistemology

ABSTRACT

It has been claimed that epistemic peers, upon discovering that they disagree on some issue, should give up their opposing views and 'split the difference'. The present paper challenges this claim by showing, with the help of computer simulations, that what the rational response to the discovery of peer disagreement is—whether it is sticking to one's belief or splitting the difference—depends on factors that are contingent and highly context-sensitive.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Science*

1. Introduction

How should we react to the discovery that we disagree on some issue with an epistemic peer, that is, a person having the same evidence and judgmental skills as we do? Can we rationally continue to believe as we do, or should we abandon our belief? Or, when it makes sense, should we adopt a belief that is a mixture between the other person's belief and our current one?

It has been argued that, upon such a discovery, peers should give up their opposing views and 'split the difference'. On this view, to stick to one's opinion in the face of peer disagreement betokens irrationality on one's part. The significance of this claim—call it the 'Irrationality Claim' (IC)—is hard to overrate. For instance, if true, IC would seem to undermine the familiar and widely cherished view that science is a paradigmatically rational enterprise and that scientists are (typically) highly rational agents. After all, longstanding disagreements, even among practitioners generally regarded to be each other's peers, are commonplace in science. In fact, such disagreements may even be more pervasive in philosophy. It is scarcely surprising, therefore, that the tenability or otherwise of IC is presently being hotly debated by philosophers.

Elsewhere, I have argued that the extant arguments for IC rest on feeble grounds (Douven, 2009). Here I want to challenge IC more directly. In fact, this paper aims to undermine a presupposition that seems to be shared not only by all *proponents* of IC but also by many of its *opponents*, to wit, that how we ought to respond to peer disagreement can be established in an a priori manner. Computer simulations will be used to show that what the rational response in such cases is, may well depend on factors that can only be empirically discovered. The simulations will also show that these factors may be highly context-dependent: in some contexts sticking to one's belief in the face of disagreement with a peer may be irrational, in others this may be precisely the rational thing to do. This contradicts IC, which asserts, after all, that sticking to one's belief in the face of peer disagreement is *never*, and thus in *no* context, rational.

I begin, in Section 2, by describing the model that will be used for our simulations; then, in Section 3, I present the results of the simulations; and in Section 4, I bring these to bear on IC. In Section 5, I address two objections that some might have to my approach to the disagreement debate.

2. The Hegselmann–Krause model

Most examples discussed in the literature about disagreement concern situations in which the disagreement between the peers consists in their holding contradictory beliefs, beliefs that cannot both be true and cannot both be false. However, some authors also

E-mail address: igor.douven@hiw.kuleuven.be

0039-3681/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsa.2010.03.010

consider examples in which the disagreeing parties hold (not merely contradictory but) contrary beliefs, beliefs that cannot both be true but that might both be false.[1] Indeed, disagreement among epistemic peers seems to be much more often of this variety than the examples discussed in the literature suggest. Think of scientists disagreeing about the age of the Earth, or the mass of some elementary particle, or the strength of some physical force, or the average housing prices in the United Kingdom in 2015. In such cases, it is typically not true that one scientist holds that the value (age, mass, ...) equals x while the other merely thinks it does not equal x; the other typically thinks it equals y rather than x. Further, as Christensen (2007) rightly emphasizes, peers may disagree in that they assign different probabilities to a given proposition. Patently, in these cases peers can split the difference in the straightforward sense that they can adopt as a new belief or probability one that is a mixture between the beliefs or probabilities that each of them held prior to the discovery of the disagreement. For instance, if you believe that some value we are both interested in equals x and I believe that it equals y, then we split the difference in this sense when each of us adopts as a new belief that the value equals $(x+y)/2$. Similarly, if one scientist's probability for a hypothesis H equals .6, and another scientist's probability for H equals .4, then they split the difference by adopting .5 as their new probability for H.

In this and the next section, I focus on this 'averaging' way of splitting the difference, which some have recommended as the rational response to the discovery that a peer holds a belief contrary to one's own, or that she assigns a different probability to a given proposition.[2] In Section 4, I will also consider other sorts of difference-splitting that have been proposed in the literature on peer disagreement and relate them to the above one.

To appraise the recommendation to split the difference with a disagreeing peer, we need not quite start from scratch. Researchers from diverse disciplines, including mathematics, physics, and computer science, have studied this difference-splitting way of forming new beliefs, or adapting probabilities, in artificial societies, that is, communities of artificial agents, simulated in computers. They have concentrated on questions concerning the circumstances under which communities of initially disagreeing agents converge, fully or partly, and those under which they tend to polarize, supposing agents are willing to split the difference with at least some of their fellow agents; normative questions have remained largely unaddressed so far. Nevertheless, the various frameworks which they have developed for investigating the said descriptive questions can also be used to investigate normative ones, such as, most notably, under what circumstances it would be rational to split the difference with one's epistemic peers.

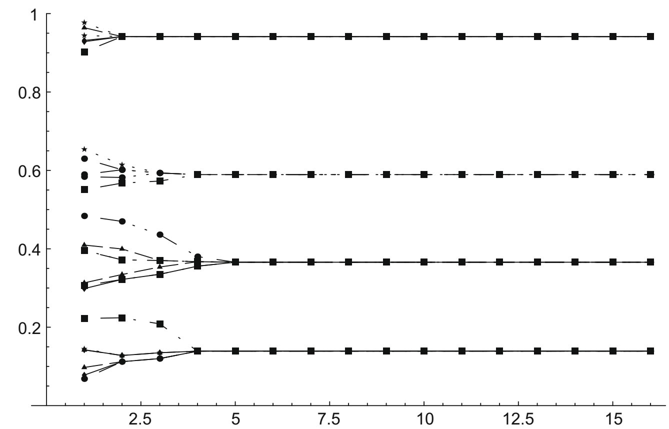

Fig. 1. Repeated difference-splitting.

The framework developed in a number of papers by Hegselmann and Krause is particularly suited for this purpose.[3] In these papers, they study by means of computer simulations the opinion dynamics of communities of agents who are individually trying to determine the value of a certain parameter, where the agents know that the true value lies in the half-open interval $\langle 0,1]$. The agents are willing to split the difference only with agents whose opinions are within a distance of ε from their own, for some given $\varepsilon \in [0,1]$.[4] For instance, if twenty-five agents, each with different initial opinions, update their opinions by averaging over all opinions that are within a distance of .1 from their own, then their opinions may, given repeated simultaneous updating, evolve as shown in Fig. 1.[5] While Hegselmann and Krause do not explicitly consider the possibility of interpreting this process as the sequential updating of probabilities, such an interpretation would certainly make sense. For instance, we can think of the agents as trying to determine the chance of some given event, and imagine that in the updating process each agent is trying to set his or her subjective probability that this event will occur equal to the event's chance. In the following, I will often just speak of the opinions of agents, and how these opinions change over time, but all the simulations to be presented can also be interpreted in probabilistic terms.

For the most part, Hegselmann and Krause study a more complicated and also more interesting way of updating than the

[1] See Christensen (2007), Kelly (Forthcoming), and Goldman (Forthcoming).

[2] See, for example, Lehrer (1976, 1980), Wagner (1978), and Lehrer and Wagner (1981). These authors consider groups of agents possessing the same relevant evidence who are to decide about the value of a variable with range $\langle 0,1\rangle$—like for instance the probability to be assigned to a given proposition—and who assign real-valued weights between 0 and 1, with 0 and 1 included, to each other (including themselves) such that the weights any given agent assigns to the members of the group (including herself) sum to unity. Informally, the weight one agent assigns to another is meant to reflect the former's judgment about the latter's expertise on whatever the relevant issue is. Lehrer and Wagner propose that the agents update their opinions/probabilities by taking weighted averages of the opinions/probabilities of all the agents in the group. They show that, under fairly weak conditions, iteration of this procedure will lead to (what they call) a consensual opinion/probability, to which, Lehrer and Wagner claim, the agents are then rationally committed. In these terms, peer disagreements are naturally conceived as cases in which at a given time two or more agents assign the same weight to the other or others as they assign to themselves. Lehrer and Wagner's proposal then effectively amounts to a recommendation to split the difference with one's peers in the averaging sense. (Technically speaking, in the case the agents assign equal weights to each other, the weight matrix is identical to any of its powers, so that the consensual opinion/probability can be calculated simply by multiplying the vector of the agents' opinions/probabilities by the weight matrix. This yields as output a vector of equal opinions/probabilities that are the arithmetic average of the initial opinions/probabilities. See Lehrer & Wagner (1981), Ch. 7.4.) At least with respect to cases of peers' assigning differing probabilities, Christensen (2007) and Elga (2007) also think that this is the right way to split the difference.

[3] See Hegselmann & Krause (2002, 2005, 2006); see also Dittmer (2001) and Fortunato (2004). Similar models were developed independently by Deffuant et al. (2000) (see also Weisbuch et al., 2002) and Ramirez-Cano & Pitt (2006). See Lorenz (2007) for a useful overview of the main technical results in this area.

[4] In the terminology of Hegselmann and Krause, ε generates a *symmetric confidence interval*. It is called 'symmetric' because it is immaterial whether an agent assigns a value to the parameter at issue that is higher than the one oneself assigns or whether this value is lower: as long as the agent assigns a value that is within ε of one's own assignment, she is among those one is willing to split the difference with. In some studies, Hegselmann and Krause also consider agents whose confidence intervals are asymmetric. Also, a society of agents is said to be *homogenous* iff the confidence interval is the same for all agents. Hegselmann & Krause (2002, 2006) contain some interesting results about non-homogenous societies. All simulations to be presented in this paper assume both symmetry and homogeneity.

[5] Most authors concerned with peer disagreement limit their attention to cases which involve only two disagreeing peers (Elga, 2007, p. 484, is a notable exception). I take it, however, that this is merely for reasons of simplicity. It would at least be odd to hold that two disagreeing peers should split the difference, but that three or more disagreeing peers should do something else, especially in view of the fact that the arguments that have been presented in support of the claim that disagreeing peers should split the difference do not at all depend on the precise number of such peers. It thus is justified to consider IC in contexts in which agents may have more than one disagreeing peer.

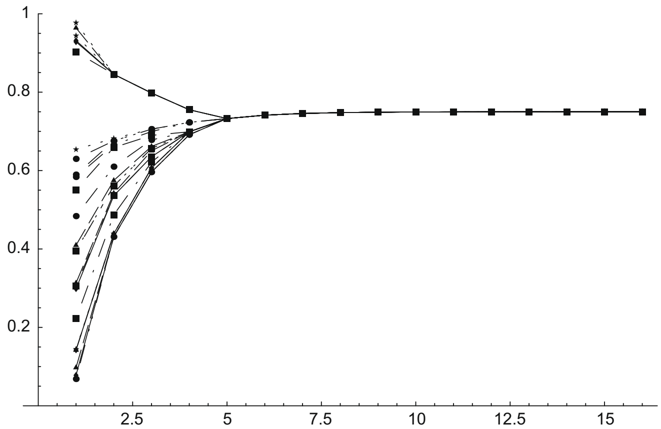

Fig. 2. Repeated difference-splitting with data gathering.

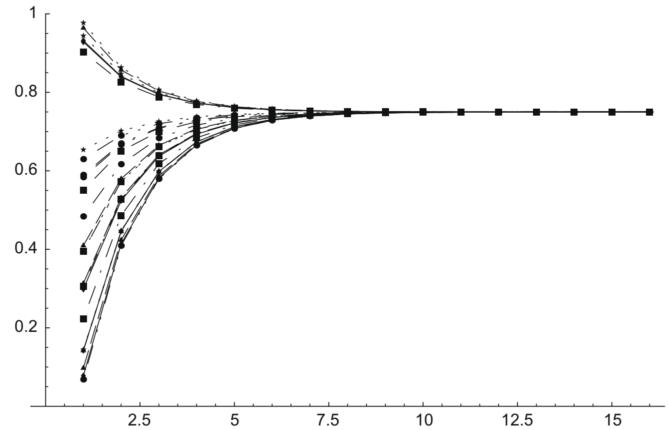

Fig. 3. Data gathering, but no difference-splitting.

just-described one. They assume that the agents not only change their opinions by taking into account other agents' opinions insofar as these are close enough to their own, but also gather information about the world directly, say, by doing experiments. More exactly, these agents are supposed to update their opinions by taking a weighted average of, on the one hand, the (straight) average of the opinions that are close enough to their own and, on the other hand, the truth value of the parameter they are trying to determine. Let $\tau \in \langle 0, 1]$ be the true value of the parameter, $x_i(u)$ the opinion of agent x_i after the u-th update, and $\alpha \in [0, 1]$ the weighting factor used. Further define $X_i(u) := \{j : |x_i(u) - x_j(u)| \leqslant \varepsilon\}$, and let $|X_i(u)|$ be the cardinality of $X_i(u)$.[6] Then the opinion of agent x_i after the $(u+1)$-st update is given by this equation:

$$x_i(u+1) = \alpha \frac{1}{|X_i(u)|} \sum_{j \in X_i(u)} x_j(u) + (1-\alpha)\tau. \tag{1}$$

If we set $\tau = .75$, $\alpha = .5$, and $\varepsilon = .1$, then if the agents considered in Fig. 1 updated their opinions by (1), these would evolve over time as shown in Fig. 2. Eq. (1) will serve as our basic model for appraising difference-splitting as a response to the discovery of peer disagreement.

To forestall misunderstanding about this model, it is worth underscoring that the model explicitly does *not* suppose that the agents themselves know the true value of the parameter. The idea is that an agent gets information, for instance by performing experiments, which points in the direction of that value, and which, together with the opinions of her peers, determines her new opinion in a way captured by Eq. (1). The algorithm or algorithms the agents themselves use to accommodate the data are left unspecified by Hegselmann and Krause.[7]

3. Simulating disagreements

Naturally, we can stipulate that the people whose relevant opinions (or relevant probabilities) are close enough (within ε) to one's own at a given point happen to be precisely those who one regards as one's epistemic peers at that point.[8] In that case, the graphs in Figs. 1 and 2 give us an indication of how the opinions of agents who respond to disagreement with their peers in a certain uniform manner may evolve over time when they do not, respectively when they do, also gather data about the world directly. As to the latter kind of situation, one may reasonably ask why people who gather information about the world which in some way points them in the direction of the truth should take into account the opinions of others at all. Why not go purely by the data they gather? In terms closer to our present debate, why should they attach any significance to disagreement with epistemic peers rather than accommodating their opinion solely on the basis of the data they gather? The answer to be argued for in this paper is: it is not a priori that they should attach any significance to the opinions of their peers, nor, however, is it a priori that they should *not* do so. What it is best to do, whether they should listen to any of the other agents, or even let others heavily influence their updates, or rather ignore them, or at least pay little attention to what they have to say and (mostly) go by the evidence, may largely depend on contingent and highly variable factors. This, I will claim, follows from the results of the simulations to be presented in this section.

First compare the development of the opinions of a community of twenty-five agents as represented by Fig. 2—so, with $\alpha = .5$, $\varepsilon = .1$, and $\tau = .75$—with that represented by Fig. 3. The latter shows how the opinions of these agents would evolve if no agent took into account the opinion of any of the others and each agent updated in such a way that her opinion immediately after the update is a straight average of her opinion immediately before the update and the true value of τ, that is, if agent x_i's opinion after update $u+1$ is given by (1) with $\alpha = .5$, $\varepsilon = 0$, and $\tau = .75$. Already comparing these figures suggests that taking into account the opinions of others makes hardly any difference to how the opinions of the individual agents develop. Simulations with 100 societies, each consisting of twenty-five agents, all with different initial opinions, confirm that the average distance from the truth of the agents' opinions after five, ten, twenty-five, and fifty updates differs at most negligibly when the agents take into account the opinions of others from when they do not do so (further updates were not considered, given that after ten updates the average distance from the truth was already approximately 0).[9] This result turned out to be fairly robust for different values of ε.

In the above simulations, it was supposed that the data the agents receive by investigating the world at each step point pre-

[6] Note that it is always the case that $i \in X_i(u)$ and also that, supposing symmetry and homogeneity (cf. note 4), $i \in X_j(u)$ iff $j \in X_i(u)$. Clearly, though, these conditions do not guarantee that if $i \in X_j(u)$ and $j \in X_k(u)$, then $i \in X_k(u)$.

[7] See Hegselmann & Krause (2006), Sect. 1, for more on how the data gathering element is to be interpreted.

[8] Some might object to this stipulation on the grounds that we do not then generally have that if i regards j to be her peer and j regards k to be her peer, then i will also regard k to be her peer (see n. 6). As will be seen later on, it is questionable whether this is problematic given a proper understanding of epistemic peerhood. More importantly, it will also be seen that we can do without the stipulation if we want.

[9] For all agents x_i and updates u, the distance of x_i's opinion from the truth after u is simply taken to equal $|\tau - x_i(u)|$.

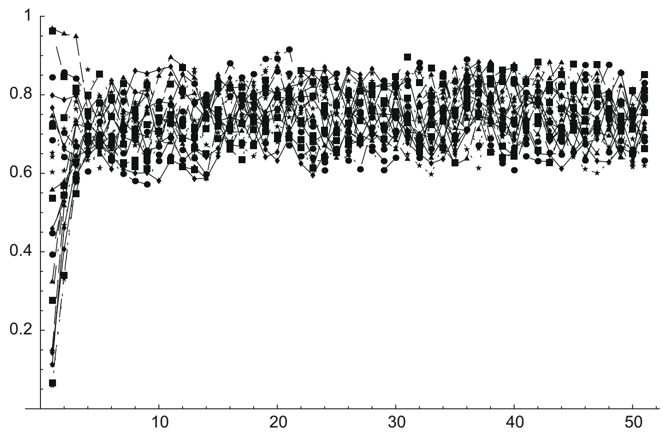

Fig. 4. Noisy data, no difference-splitting.

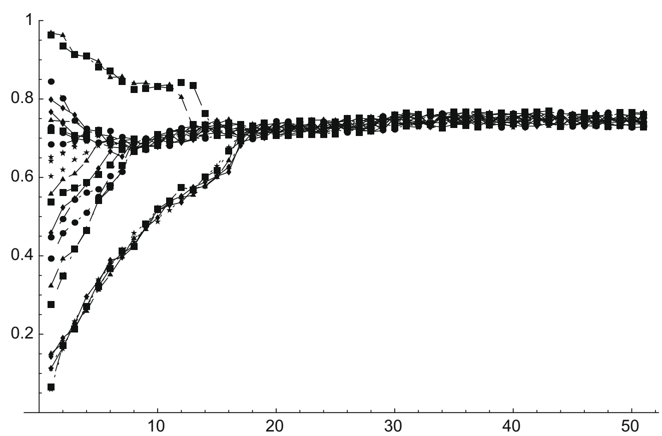

Fig. 5. Noisy data, repeated difference-splitting.

cisely in the direction of the truth. It must be admitted, however, that this supposition is not very realistic. In real life, we have to deal with measurement errors and other facts that may make our data 'noisy'. To accommodate this, I performed simulations in which the data about τ the agents receive by directly investigating the world may be 'off' a bit; that is, the data received at each update do not necessarily point in the direction of precisely τ but possibly only to some value close to it. More exactly, the opinion of agent x_i after update $u+1$ is given not by (1) but by

$$x_i(u+1) = \alpha \frac{1}{|X_i(u)|} \sum_{j \in X_i(u)} x_j(u) + (1-\alpha)(\tau + \text{rnd}(\zeta)), \tag{2}$$

where rnd(ζ) is a function returning a unique uniformly distributed random real number in the interval $[-\zeta, +\zeta]$, with $\zeta \in [0,1]$, each time the function is invoked (so the value of rnd(ζ) may be different for each agent and also from one update to the next). Fig. 4 shows the development of the opinions of twenty-five agents who update in this way, with $\alpha = .5$, $\varepsilon = 0$ (so, no difference-splitting), $\tau = .75$,

Table 1
Average distances from the truth, rounded to three decimal places; left: $\alpha = .5$, $\varepsilon = 0$, $\tau = .75$, $\zeta = .2$; right: $\alpha = .9$, $\varepsilon = .1$, $\tau = .75$, $\zeta = .2$

Number of updates	No difference-splitting	Difference-splitting
5	.058	.202
10	.055	.117
25	.055	.021
50	.055	.01

and $\zeta = .2$. Fig. 5 shows the development of the opinions of twenty-five agents with the same initial opinions as the agents considered in Fig. 4 but who, by contrast, do take into account at each update opinions that do not differ too much from their own. In fact, they attach great weight to those opinions; to be precise, their updates are given by (2) with $\alpha = .9$ and $\varepsilon = .1$ ($\tau = .75$ and $zeta = .2$, as above).

Table 1 summarizes the results of 100 runs with different societies, where the numbers are the averages of the average distance from the truth of the opinions of the twenty-five agents in those societies after five, ten, twenty-five, and fifty updates, both when it is assumed that the agents do not talk to each other ($\alpha = .5$, $\varepsilon = 0$) and when it is assumed that they do talk to other agents with 'close enough' opinions and give great weight to these other opinions ($\alpha = .9$, $\varepsilon = .1$); no results about further updates are stated because the averages remain stable after fifty updates.[10] These data confirm what one might already have guessed on the basis of Figs. 4 and 5, namely that, first, on average the members of the no-difference-splitting societies get within a moderate distance of the truth relatively quickly, but for them that is about as good as it gets in terms of truth closeness; and second, on average, convergence towards the truth occurs at a slower pace for the members of the difference-splitting societies, but in the somewhat longer run they are on average much closer to the truth than the members of the no-difference-splitting societies.[11]

As a first comment on these results, let me mention that, mathematically, they are not hard to make sense of. To see why the difference-splitting societies end up being closer to the truth than the no-difference-splitting societies, recall that the noise on the data the agents receive is randomly distributed in the interval $[.55, .95]$. Consequently, the opinions of the agents in the no-difference-splitting societies, who go purely by the data, also are, from a certain (actually very early) point in time on, randomly distributed in this interval around the truth. Averaging over those of the thus distributed opinions that are within a distance of ε of any given one of those opinions would then normally result in a value that is at least as close to the truth as the given opinion itself is (because averaging over a set of values randomly spread over a certain interval will normally yield a value close to the midpoint of that interval). As a moment's reflection will reveal, however, for the opinions in the intervals $[.55, .65\rangle$ and $\langle .85, .95]$, this averaging would normally result in a value that is *closer* to the truth. So, were they to average over the opinions within ε-distance of their own, the agents in the no-difference-splitting societies would, on average, end up being closer to the truth than they do now. But averaging in the said way is precisely what the agents in the

[10] The characterization of the case in which $\alpha = .5$ and $\varepsilon = 0$ as one in which agents do not talk to each other is to understood with some care. In a way, an agent with $\varepsilon = 0$ could be said to talk to *all* other agents in the model should it occur that all other agents have exactly the same opinion that she has. However, it is easy to see that, where n agents share exactly the same opinion, the 'social component' of each of these agents' new opinion is then ($n\times$ the agent's previous opinion)/n, which equals the agent's previous opinion and is therefore equivalent to having no social exchange at all—which is what is meant by 'not talking to others'. (Thanks to an anonymous referee for spotting this possible cause of misunderstanding.)

[11] The members of the no-difference-splitting societies could actually have done better in regard to approximating the truth if in updating they had given more weight to their own opinion just before the update. More precisely, 100 runs with different societies, each consisting of twenty-five agents who update by (2) with $\alpha = .9$, $\varepsilon = 0$, $\tau = .75$, and $\zeta = .2$ gives the following averages of the average distance from the truth after five, ten, twenty-five, and fifty updates, respectively: .207, .125, .033, .021 (again the average remains stable after fifty updates). Note, though, that if they had also taken into account at each update the opinions differing from their own by no more than .1, they would on average have ended up being twice as close to the truth as they ended up now.

Table 2
Average least squares penalties, rounded to four decimal places; left: $\alpha = .5$, $\varepsilon = 0$, $\tau = .75$, $\zeta = .2$; right: $\alpha = .9$, $\varepsilon = .1$, $\tau = .75$, $\zeta = .2$

Number of updates	No difference-splitting	Difference-splitting
5	.0098	.1169
10	.0089	.0394
25	.0088	.0011
50	.0088	.0004

Table 3
Average distances from the truth (average least squares scores); left: $\alpha = .5$, $\varepsilon = 0$, $\tau = .75$, $\zeta = .2$; right: $\alpha = .9$, $\varepsilon = 1$, $\tau = .75$, $\zeta = .2$

Number of updates	No difference-splitting	Difference-splitting
5	.058 (.0098)	.162 (.0558)
10	.055 (.0089)	.096 (.0196)
25	.055 (.0088)	.02 (.0011)
50	.055 (.0088)	.01 (.0003)

difference-splitting societies do! As to why difference-splitting societies converge more slowly, note that the agents in these societies start out by giving great weight (.9) to something in the proximity of an initial guess that is no better than random, and little weight (.1) to a value that is comparatively close to the truth. By contrast, the agents in the no-difference-splitting societies start out by giving less weight (.5 as opposed to .9) to an initial guess that is no better than random, and more weight (.5 as opposed to .1) to a value that is comparatively close to the truth.[12]

It must further be remarked that while average distance from the truth is a plausible measure for assessing, say, estimates of the mass of an elementary particle, it is not a good measure for assessing subjective probabilities or degrees of belief. Average distance from the truth is what statisticians call an 'improper scoring rule', where, roughly, a *scoring rule* is a rule for penalizing inaccuracies in people's probabilities, and an *improper* scoring rule is a scoring rule that might give people an incentive to misrepresent their actual probabilities; put more exactly, a scoring rule is proper iff the agent minimizes her expected penalty by being up front about her probabilities (see, for instance, Rosenkrantz, 1981, Sect. 2.2, for details). So, for the case that we interpret the simulations as representing agents' evolving probabilities—which, as intimated, is a possible interpretation—I reran the simulations with the least squares metric, which *is* a proper scoring rule. Let $p = (p_1, \ldots, p_n)$ be a vector of objective probabilities summing to 1 and $p' = (p'_1, \ldots, p'_n)$ the vector of a given agent's corresponding subjective probabilities; for instance, p_c might be the chance of drawing a ball of color c from a given urn containing balls of n different colors and p'_c the agent's subjective probability that a ball of color c will be drawn. Then according to the least squares metric this agent's penalty is

$$S(p', p) = \sum_{c=1}^{n} (p'_c - p_c)^2.$$

So, in the cases considered in the simulations the least squares metric penalizes agent x_i after update u by

$$\begin{aligned} S((x_i(u), 1 - x_i(u)), (\tau, 1 - \tau)) &= (x_i(u) - \tau)^2 + (1 - x_i(u) - (1 - \tau))^2 \\ &= (x_i(u) - \tau)^2 + (\tau - x_i(u))^2 \\ &= 2(x_i(u) - \tau)^2. \end{aligned}$$

100 runs with no-difference-splitting societies and 100 runs with difference-splitting societies gave the averages of the average least squares penalties of the members of those societies after five, ten, twenty-five, and fifty updates that are presented in Table 2. Qualitatively, we have the same result here as above. After five updates, the members of the difference-splitting societies on average receive a penalty that is about twelve times as big as the average penalty the members of the no-difference-splitting societies receive. But after fifty updates the situation is very different; then the members of the no-difference-splitting societies receive on average a penalty that is twenty-two times as big as the average penalty the members of the difference-splitting societies receive.

Another point worth emphasizing is that we can do without the assumption that one's peers at a given point in time are precisely those people whose opinions (or probabilities) are close to one's own. For although all extant studies of the Hegselmann–Krause model consider only relatively small values of ε,[13] there is no principled reason against taking, for instance, $\varepsilon = 1$, so that throughout a simulation all members of the society talk to each other. Such a simulation could be plausibly interpreted as representing the development of opinions of a single group of peers who continue to regard each other as peers throughout the simulated process in which they repeatedly perform new experiments and talk to each other. The remarkable thing is that if we set $\varepsilon = 1$ indeed, then we still find that the members of the no-difference-splitting societies on average converge to the truth more rapidly than the members of the difference-splitting societies (or that the former on average receive a smaller least squares penalty than the latter), but that eventually the members of the difference-splitting societies on average converge to the truth much closer (or receive a much smaller penalty) than the members of the no-difference-splitting societies. The right column of Table 3 represents the results of 100 runs with difference-splitting societies, where again $\alpha = .9$, $\tau = .75$, and $\zeta = .2$, but this time $\varepsilon = 1$. To facilitate comparison with no-difference-splitting societies, the left column restates the results about the no-difference-splitting societies already reported in Tables 1 and 2.

Nor is it essential that the agents update their opinions by simultaneously taking into account the opinions (or probabilities) of their peers as well as information they receive from, say, doing an experiment. Consider for instance Eq. (3), which is an extension of the Hegselmann–Krause model that allows us to simulate societies whose members are free at any given time to update their opinions either on the basis of experimental information or by averaging their peers' opinions (so at time u an agent might do an experiment and update her opinion on the information she receives from that, at $u + 1$ she might do another experiment and update on the new information she receives, at $u + 2$ she might update her opinion by averaging her peers' opinions at that time, and so on):

$$x_i(u+1) = \begin{cases} \alpha x_i(u) + (1-\alpha)(\tau + \mathrm{rnd}(\zeta)) & \text{if } g(i,u) \text{ is heads,} \\ \frac{1}{|X_i(u)|} \sum\limits_{j \in X_i(u)} x_j(u) & \text{if } g(i,u) \text{ is tails.} \end{cases} \tag{3}$$

Here, g is a function that returns the outcome of the flip of a fair coin.[14] If we let g depend on both i and u, as the notation sug-

[12] I owe this observation to Jake Chandler.

[13] This is unsurprising, given that these studies do not take into consideration the possibility of noisy data. If the data are non-noisy, then setting $\varepsilon = 1$ will result in all agents' having exactly the same opinion from the second update onward, which of course makes the case rather uninteresting when one's goal is to investigate 'the cognitive interaction between truth seeking individuals' (Hegselmann & Krause (2006), p. 4).

[14] It would seem interesting to experiment with taking biased coins, and even with taking differently biased coins for different agents or groups of agents (consider, for instance, that scientists with a lot of confidence in their experimental skills may have a greater tendency to do experiments, as opposed to discussing matters with their colleagues, than those have who are not so confident in the same regard), but this must remain for future research.

Table 4
Individual vs. collective updating; $\alpha = .9$, $\varepsilon = 1$, $\tau = .75$, $\zeta = .2$

Number of updates	Individual updating	Collective updating
5	.204 (.0961)	.2 (.0864)
10	.155 (.0556)	.151 (.0501)
25	.071 (.0121)	.07 (.011)
50	.021 (.0012)	.021 (.0011)
75	.01 (.0004)	.01 (.0003)
100	.009 (.0003)	.009 (.0003)

gests, then at any given time the agents update their opinions in a way that is independent of what the other agents do at the same time; indeed, independent of what any other agent does at any time (as well as of what she herself does at other times; call this 'individual updating'). But we can also make g a function of u only, in which case at each time either all agents receive experimental information and update on that—if the coin lands heads—or all agents update by averaging over the opinions of their peers, if the coin lands tails (call this 'collective updating'). The left column of Table 4 gives averages over 100 runs of average distances from the truth and, in brackets, average least squares penalties for 'individual updating' societies, and the right column gives the corresponding averages over 100 runs with 'collective updating' societies. As is readily seen, there are but minor differences between individual and collective updating, both in terms of truth closeness and in terms of least squares penalties. More importantly for our concerns, a comparison of these results with those presented in Tables 1 and 2 reveals that it takes on average longer to get close to the truth (respectively, to minimize least squares penalties) for members of difference-splitting societies—who now update by means of (3)—than for members of no-difference-splitting societies, but that given enough time the former end up being on average closer to the truth (receiving smaller penalties) than the latter.

There might still be a worry about the potential relevance of the above simulations to the peer disagreement debate. For it might be said that when the data are noisy, the agents are unlikely to receive *exactly* the same evidence, and that therefore they will stop qualifying as peers once the simulation starts running (supposing they qualified as peers at the beginning). As various authors have noted, however, a conception of epistemic peerhood that requires epistemic peers to have *exactly* the same evidence is hard to uphold, or at least would make the debate of marginal significance at best.[15] After all, on that conception we would rarely, if ever, be among epistemic peers. Nor would it seem to fit our pretheoretical conception of an epistemic peer: surely two scientists who have independently performed the same experiment can qualify as epistemic peers even if the results they got are only the same up to some small measurement errors.[16,17]

4. How to respond to peer disagreement?

It seems at least intuitively clear that, in the kind of situation assumed in the simulations with precise, 'non-noisy' data, rationality does not mandate that we account for the opinions of others, given that doing so makes no discernible difference. Rationality may even go *against* splitting the difference in those situations, inasmuch as averaging over the opinions of others may be expected to consume resources that might be spent more wisely. But what should one do in the—more realistic—kind of situation assumed in the simulations with noisy data, where it does appear to make a discernible difference whether or not one takes into account one's peers' opinions? The answer will depend on the answer to the question of which is a better epistemic strategy: one that brings us relatively fast within a moderate distance of the truth but in general no further, or one that eventually brings us *very* close to the truth, even though it may take a while before it brings us even *moderately* close to the truth.

Following Foley (1993), Ch. 1, I propose that we understand questions of epistemic rationality (or justification) in terms of conduciveness to our epistemic goal. While other authors besides Foley have argued for a conceptual connection between rationality and our epistemic goal,[18] these authors were typically concerned with the rationality of beliefs only. By contrast, Foley (ibid., p. 4) makes it clear that judgments concerning the rationality of other things, such as, most notably, strategies and methods, are best thought of as judgments of how effective those things are in helping us to achieve our goal or goals. Indeed, I find it immediately compelling to think that we judge epistemic strategies to be rational if they help us in achieving our epistemic goal. That is what we want them to do, after all, and it is the only thing we want them to do.

So now the above question becomes which of the strategies considered—whether difference-splitting or no-difference-splitting—is most conducive to the realization of our epistemic goal. In trying to answer this question, we face the problem that none of the statements of our epistemic goal to be found in the literature refer to the concept of approximate truth or truthlikeness.[19] This concept appears relevant to the question, given that with noisy data we may even in the long run be unable to do any better than to have approximately true beliefs about some propositions. Nevertheless, I do think we can bring the common formulations of our epistemic goal to bear on the issue of peer disagreement.

According to the most frequently encountered formulation, our epistemic goal is to believe the truth and nothing but the truth.[20] It would seem that, if this is our epistemic goal indeed, then both difference-splitting and no-difference-splitting could qualify as rational responses to peer disagreement. For one could argue that fast convergence to some value reasonably close to the truth and slower convergence to some value very close to the truth are, albeit in different ways, both conducive to the said goal in the kind of situations

[15] See, for example, Douven (2009), Goldman (Forthcoming), and Sosa (Forthcoming). But see also Elga (2007), who admits, though, that the notion of epistemic peer he uses is 'nonstandard' (p. 499).

[16] There might be more to the worry if we wanted to interpret the individual-updating simulations considered one paragraph back in terms of peer disagreement. In principle, it could happen that at the end of such a simulation some agents have only updated on experimental data while others have only updated by averaging the opinions of the other members of their society. But of course the results of the collective-updating simulations suffice perfectly to show that we can make the same point about convergence to the truth whether we assume that all agents update each time by both taking into account new experimental information and averaging the opinions of the others, or whether we assume, perhaps more realistically, that they do only one of these at a time. And the worry that could arise in connection with the individual-updating simulations cannot arise in connection with the collective-updating simulations, given that in the latter all agents update in the same way at each time step.

[17] If Christensen (2007), p. 212, is right that the arguments for splitting the difference with peers entail more generally that disagreeing parties having only approximately the same evidence should move their beliefs closer to one another (even if they should not quite split the difference), then what is to follow is still bad news for the advocates of IC even if it is granted that agents who do not share exactly the same evidence fail to qualify as epistemic peers. For simulations in which every agent gave more weight to her own opinion in updating than to those of others (such that she only moved her belief toward those, without precisely splitting the difference), but that were otherwise exactly like the ones presented above, have yielded again conclusions about accuracy and speed of convergence that were qualitatively the same as the conclusions obtained in the earlier simulations.

[18] See Douven (2008) and the references given therein.

[19] See Oddie (2008) for a useful overview of the explications of this concept that have been proposed in the literature.

[20] See, among others, Rescher (1973), p. 21, Lehrer (1974), p. 202, BonJour (1985), p. 8, and Foley (1992), p. 183.

we simulated. For my dialectical purposes, this would already be enough. After all, it would show that sticking with one's belief in the face of peer disagreement may be rational; that is, it would show IC to be false.

Alston may be right, however, that the standard conception of our epistemic goal is not quite correct, and that we should rather conceive of our epistemic goal as 'maximizing true beliefs and minimizing false beliefs *about matters of interest and importance*' (Alston, 2005, p. 32; my italics). In that case, it is no longer true that both difference-splitting and no-difference-splitting can be said to be rational epistemic strategies. Nor, however, could the one generally be said to be rational and the other irrational. For it is not hard to make up some story in which being within a moderate distance of the truth relatively quickly is of the highest importance—it makes the difference between life and death, say—but getting within an even closer distance has no additional benefits; at the same time, it is no harder to make up a story in which it is of the highest importance to get very close to the truth—even if this should take a while—because now the difference between being very close to the truth and being not *quite* that close to the truth is the difference between life and death.

Hence, if 'rational response' is understood in terms of conduciveness to our epistemic goal, and if Alston is right about what our epistemic goal is, then our simulations show that what the rational response to peer disagreement is depends on factors that can be found out about, if at all, only empirically—such as whether the data we receive are noisy—and on purposes that may vary from one occasion to another. For instance, if it is a matter of life and death to have, for a decision we are about to make, a rather exact estimate of the value of some parameter, and we have reason to believe that the data we get about this parameter contain random noise—as we often have—then, in view of the simulations, it may be irrational to discount the opinions of our peers. If, on the other hand, we have reason to believe that the data contain little or no noise, or if it is more important to be within a moderately close distance from the truth relatively quickly than to be, in the somewhat longer run, very close to the truth, then, as we also saw, it may well be rational to go purely by the data.

A first comment on this is that the conclusion about how we should respond to peer disagreement may be even more complex than the preceding paragraph suggests. To see why, notice that the simulations presented in the previous section all simulate societies whose agents stick to one and the same policy of responding to peer disagreement throughout the simulation. But suppose the members of what starts out as a no-difference-splitting society come to suspect that the data they are receiving are noisy. Then, perhaps because they are aware of the outcomes of simulation studies similar to the ones we did, they might want to switch to the policy of splitting the difference. That this might yield the best of both worlds is already suggested by Fig. 6, which represents the beliefs of a society of agents who receive noisy data ($\zeta = .2$) and, after the fifth update, all switch from ignoring the opinions of their peers ($\alpha = .5$, $\varepsilon = 0$) to taking them into account, and attaching great weight to them ($\alpha = .9$, $\varepsilon = .1$), in updating their own opinions: as is clearly shown, from about step 10 onward all agents have an opinion very close to $.75 (= \tau)$. This suggestion is confirmed by the data from 100 runs with similar societies, which yielded .058, .013, .011, .01, and .01 as the averages of the average distance from the truth after five, ten, twenty-five, and fifty updates, respectively. Comparison of these numbers with those given in Table 1 shows that the policy-switching societies do better, both on the count of accuracy and on that of speed of convergence, than the (uniformly) no-difference-splitting societies, and that while the (uniformly) difference-splitting societies *eventually* attain the same level of accuracy as the policy-switching societies, it takes them considerably longer to attain that level. Further simulations

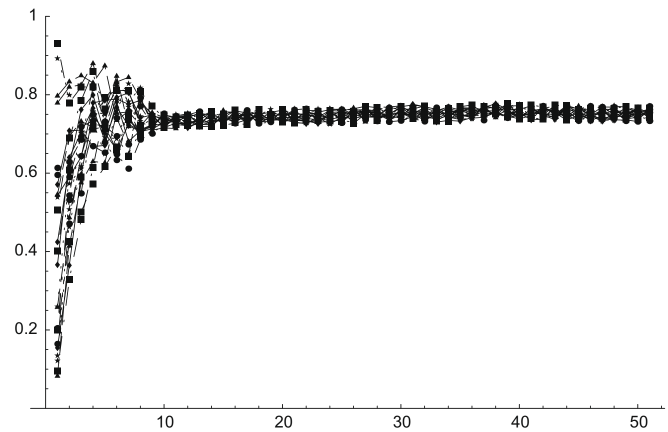

Fig. 6. Noisy data, policy-switching.

showed that if only some agents switch policies at some point in the simulated updating process, then they will also on average do better in the relevant respects than the members of the uniformly updating societies—*how much* better depending on the number of agents that switch policies. Note, however, that this does nothing to undermine my critique of IC. Quite the opposite, in fact: if, at least under the circumstances simulated here, it is rational first to follow the policy of sticking to one's opinion and then, after some updates, to switch to the policy of splitting the difference with one's peers, then, a fortiori, it can be rational to stick to one's opinion in the face of peer disagreement, contrary to what IC pronounces.

A further comment concerns the fact that I have only considered a particular way of difference-splitting and that, as intimated earlier, other ways have been proposed by IC loyalists. For instance, when Feldman (2007) asserts that difference-splitting is the rational response to peer disagreement, he means that in the face of peer disagreement one should suspend judgment on the contested claim (recall that he is only concerned with cases in which peers hold contradictory beliefs). And Christensen's (2007, p. 193) discussion of his 'restaurant case', in which two peers at a restaurant come up with contrary answers when trying to determine the equal shares of the bill that each of the dining companions should pay, makes it clear that, in his view, peers holding contrary beliefs should, upon discovering this, split the difference in the sense that they lower their confidence in the correctness of their own belief and raise the confidence in the correctness of the peer's belief. As the simulations consider difference-splitting in neither of these senses, they might be said to be not telling at all against Feldman's and Christensen's positions. In particular, it might be said that, for all I have shown, in the relevant kind of situations, difference-splitting in the way of Feldman or Christensen is always superior, from an epistemic perspective, to no-difference-splitting.

Concerning Feldman, I should begin by noting that even if he is right that difference-splitting is the rational response in the kind of cases he considers—peers holding contradictory beliefs—this is by no means enough to establish IC, as not all cases of peer disagreement are of that kind. Of course, one could consider suspension of judgment also as a possible response to the discovery of a peer's holding a contrary belief (even if Feldman himself does not do so). The main reason why I did not consider this type of response, and indeed why it may be extremely hard to model this within our approach, is that Feldman does not say how, once one has suspended belief, new evidence is to impinge on one's belief state; nor do extant confirmation theories suggest a clear answer to this question. At the same time, it seems pretheoretically obvious that in at least many situations of the type that was assumed in our simulations one will, by suspending belief in response to peer disagreement, do worse in terms of truth approximation than by either of the types of responses—splitting the difference with one's peers in the averaging sense and, respectively, ignoring their opinions—that were studied in the simulations. For instance, if a certain bacterium is .85 μm in diameter, then in an intuitive sense (which is admittedly hard to make precise) one seems to be closer to the truth if one believes the bacterium to have a diameter of .86 μm than if one suspends belief on the issue of the bacterium's diameter. Already for this reason I would be surprised if there were some argument to the effect that Feldman's proposal is generally superior to either of the said types of response. In any event, I am happy to leave it as a challenge to Feldman, or anyone sharing his view, to provide such an argument.

As to Christensen, it is not straightforward either to relate our simulations to his sense of difference-splitting for peers holding contrary beliefs, as there is no indication in Christensen's paper of how, according to him, the peers in the restaurant case, or peers in relevantly similar situations, are to proceed once they have lowered their confidence in the correctness of their own answer. On the other hand, it would seem natural to assume that the peers in the restaurant case will want to do the requisite calculation again, this time perhaps using pencil and paper. Likewise, if the beliefs at issue in a disagreement are of an empirical nature, and are partly based on experimental information, the peers might well want to repeat an experiment they performed earlier, or perform a new one, or try to get 'worldly' information in some other manner, and then form a new, possibly different opinion on this basis, presumably with renewed confidence. It also seems reasonable to think that if the new opinion appears to diverge again from the new opinion of their peer (or the new opinions of their peers), then, on Christensen's account, the peers should lower their confidences in their own new opinions. They might then want to do the math a third time or, in case of empirical beliefs, do still another experiment, and so on. Notice that if what I am assuming here is correct, then the earlier simulations with no-difference-splitting societies ('no difference-splitting' in our sense, that is) can still be interpreted as representing the development of the peers' doxastic development insofar as their categorically held beliefs are concerned, even if the lowering and raising of the confidences with which they hold these beliefs at the various stages of development are not being represented.[21] And that may be enough for the above considerations about accuracy and speed of convergence to the truth to apply, from which it would then follow that Christensen's proposal of how to respond to the kind of peer disagreement at stake here may or may not be the rational way to proceed, depending on contingent circumstances. Of course, this is all conditional on the supposition that I have identified the right way for the peers to proceed after the discovery of the disagreement. Christensen may deny this, or he may for some other reason think that the said considerations are irrelevant to his proposal. If so, then here too I am happy to leave it as a challenge to Christensen to be more forthcoming about what his view amounts to.

As a final comment, I note that the formulations of our epistemic goal considered so far concern our categorically held beliefs only. If we want to interpret the credal states of the agents modelled in our simulations in terms of subjective probabilities or degrees of belief—as, I suggested, we can do—then instead of evaluating our epistemic strategies in light of the said epistemic goal we could appeal to something like the following 'truth-possessional criterion of system superiority' presented in Goldman (Forthcoming, p. 6 of typescript):

TPCSS Epistemic system E is better than epistemic system E^* iff conformity to E would produce (in the long run) a higher total amount of degrees of truth-possession than conformity to E^* would produce,

where again a scoring rule is used to measure degree of truth-possession.[22] Evaluated in light of this principle, and assuming the credal states of our agents to be representable by probabilities indeed, the simulations seem to show that difference-splitting is the better epistemic strategy (supposing we want to commit ourselves to one uniform strategy for dealing with cases of peer disagreement). However, while Goldman expresses some sympathy for TPCSS, he expects it to be open to criticism. A complaint one might have about it is that speed of convergence is no factor in the principle, even though it is unlikely that we would prefer an epistemic system to another if, in the short run, the former led to a higher total amount of degrees of truth-possession than the latter, whereas the latter led to a higher such amount in the long run, where, as Keynes famously reminded us, 'we are all dead'. In view of this, a version of the principle sensitive to contextual (pragmatic) factors would seem more plausible. But on such a principle, we should again not expect an unequivocal answer to the question of whether difference-splitting or no-difference-splitting is the better epistemic strategy (or, in terms of epistemic systems, whether an epistemic system comprising the strategy of splitting the difference with one's peers is, all else being equal, better than one comprising the strategy of not splitting the difference with one's peers).

5. Anticipated objections

This section addresses two objections that some might want to level against the above approach to peer disagreement.

First, I earlier said that the simulations *show* that it may depend on our contextually determined practical purposes, and on such factors as whether the data we receive are noisy, what the rational response to peer disagreement is. But it might be objected that I should have chosen a more cautious wording instead, like perhaps that the simulations *suggest* this. They are just *simulations*, after all.

Here it is good to keep in mind what purpose the above simulations are meant to serve. When computer simulations are being used to make predictions about global warming effects or stock market behavior, or even to propose environmental or economic policies, then, before we take such predictions seriously, or decide to implement such policies, we better make sure that the model underlying the simulations fits reality 'well enough'. For the kinds of models at issue in such simulations, there is typically some non-negligible level of doubt about their fit with reality, which warrants some caution with regard to the interpretation of the results of simulations that make use of them and also, concomitantly, with regard to any recommendations policymakers base on those results. By contrast, the simulations presented in this paper are not meant to model reality to any degree of precision; the situations modelled in the simulations may never have occurred in the real world, and maybe never will occur. The point is that these situations certainly *could* occur; a group of scientists, all of whom are trying to determine the value of a certain parameter, *could* epistemically behave in

[21] It may be worth noting here that on most accounts of the relationship between categorical belief and subjective probability, it is perfectly possible to maintain a categorical belief while lowering the probability one assigns to it, at least so long as this probability does not fall below a certain threshold value.

[22] Or we could postulate as our epistemic goal that we should minimize inaccuracy in our degrees of belief, where inaccuracy would again be measured by means of a scoring rule; see, for example, Joyce (1998) for a proposal in this vein. But to this proposal the same remarks would apply that are now to be made about Goldman's proposal.

the ways the artificial agents in the simulations were programmed to behave. And describing situations which could occur and in which it might be rational or irrational to stick with one's belief in the face of peer disagreement, depending on one's practical purposes, for instance, is enough to claim that there is no uniform answer to the question of how we ought to respond to peer disagreement. A fortiori, it is enough to refute the general claim IC.[23]

Second, some might have worries about the intrusion of pragmatic elements into the determination of what the rational response is to a case of peer disagreement. After all, 'rational' is here supposed to be understood as 'epistemically rational', and—it might be asked—why should it matter to whether a given response counts as epistemically rational what one's *practical* concerns at the time of the response are?

In reply to this, let me note that whatever one thinks of the connection between epistemic rationality and practical concerns, it is not an issue that need specifically interest us in the current context. If we understand epistemic rationality in terms of conduciveness to our epistemic goal, as per Foley's proposal, and we accept Alston's statement of that goal, then, because the latter explicitly brings pragmatic matters into the picture, the connection follows as a matter of course. One reason why this need not interest us is that, as intimated, for my dialectical purposes, the standard conception of our epistemic goal, which does not refer to what is of interest or importance to us, suffices entirely; that already allows me to conclude from the above simulations that it can be rational to stick with one's belief after the discovery that a peer disagrees. Another reason why it seems unnecessary to go into the issue is that several authors have recently argued for 'pragmatic encroachment' (to use Weatherson's 2005 term) on grounds that are wholly independent of the foregoing, and that their arguments have found some acclaim in the analytical community. So, the issue may not be so controversial anyway.

6. Conclusion

In this paper, we have studied by means of computer simulations different policies of responding to peer disagreement, and we have evaluated these policies with regard to their ability to track the truth, which in turn we related to the concept of epistemic rationality. The results of our studies are clearly good news for the opponents of IC, who hold the, after all relatively weak, thesis that it is not necessarily irrational to continue believing as one does in the face of disagreement with one's epistemic peer or peers. Of course, in a sense these results are also bad news for all—proponents and opponents of IC alike—who have thought that there are some general and illuminating things to be said about what we ought or ought not to do in cases of peer disagreement, or even that this question can be settled, merely by reflecting on the concepts of disagreement and epistemic peerhood, and some related ones perhaps. For surely, the main lesson to be learned from the simulations we have looked at is that what it is best to do in cases of disagreement with peers may depend on circumstances the obtaining of which is neither general nor a priori.

Acknowledgements

My greatest debt is to an anonymous referee for very helpful comments on a previous version of this paper. Thanks also to Jake Chandler, Rafael De Clercq, Carlo Martini, Alexander Riegler, Jantine Schröder, and Christopher von Bülow for useful comments and discussions. Previous versions of this paper were presented at a workshop on formal methods in philosophy, organized by Stephan Hartmann and Hannes Leitgeb and held in August 2008 in Cracow as part of ECAP 6, and at a workshop on formal modelling in social epistemology, organized by Stephan Hartmann, Carlo Martini, and Jan Sprenger, and held in October 2008 in Tilburg. I am grateful to the organizers for the invitations and to the audiences on those occasions for useful questions and remarks.

References

Alston, W. (2005). *Beyond "justification"*. Ithaca, NY: Cornell University Press.
BonJour, L. (1985). *The structure of empirical knowledge*. Cambridge, MA: Harvard University Press.
Christensen, D. (2007). Epistemology of disagreement: The good news. *Philosophical Review, 116*, 187–217.
Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems, 3*, 87–98.
Dittmer, J. C. (2001). Consensus formation under bounded confidence. *Nonlinear Analysis, 7*, 4615–4621.
Douven, I. (2008). The lottery paradox and our epistemic goal. *Pacific Philosophical Quarterly, 89*, 204–225.
Douven, I. (2009). Uniqueness revisited. *American Philosophical Quarterly, 46*, 347–361.
Elga, A. (2007). Reflection and disagreement. *Noûs, 41*, 478–502.
Feldman, R. (2007). Reasonable religious disagreements. In L. Antony (Ed.), *Philosophers without gods: Meditations on atheism and the secular life* (pp. 194–214). Oxford: Oxford University Press.
Foley, R. (1992). Being knowingly incoherent. *Noûs, 26*, 181–203.
Foley, R. (1993). *Working without a net*. Oxford: Oxford University Press.
Fortunato, S. (2004). The Krause–Hegselmann consensus model with discrete opinions. *International Journal of Modern Physics, C15*, 1021–1029.
Gaylord, R. J., & D'Andria, L. J. (1998). *Simulating society*. New York: Springer.
Goldman, A. (Forthcoming). Epistemic relativism and reasonable disagreement. In R. Feldman, & T. Warfield (Eds.), *Disagreement*. Oxford: Oxford University Press.
Hartmann, S. (1996). The world as a process: Simulations in the natural and social sciences. In R. Hegselmann, U. Mueller, & K. G. Troitzsch (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Dordrecht: Kluwer.
Hegselmann, R. (1996). Cellular automata in the social sciences. In R. Hegselmann, U. Mueller, & K. G. Troitzsch (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 209–233). Dordrecht: Kluwer.
Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence: Models, analysis, and simulations. *Journal of Artificial Societies and Social Simulation, 5*. http://jasss.soc.surrey.ac.uk/5/3/2.html. (Accessed 4 March 2008).
Hegselmann, R., & Krause, U. (2005). Opinion dynamics driven by various ways of averaging. *Computational Economics, 25*, 381–405.
Hegselmann, R., & Krause, U. (2006). Truth and cognitive division of labor: First steps towards a computer aided social epistemology. *Journal of Artificial Societies and Social Simulation, 9*. http://jasss.soc.surrey.ac.uk/9/3/10.html. (Accessed 4 March 2008).
Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science, 65*, 575–603.
Kelly, T. (Forthcoming). Peer disagreement and higher order evidence. In R. Feldman, & T. Warfield (Eds.), *Disagreement*. Oxford: Oxford University Press.
Lehrer, K. (1974). *Knowledge*. Oxford: Clarendon Press.
Lehrer, K. (1976). When rational disagreement is impossible. *Noûs, 10*, 327–332.
Lehrer, K. (1980). A model of rational consensus in science. In R. Hilpinen (Ed.), *Rationality in science* (pp. 51–62). Dordrecht: Reidel.
Lehrer, K., & Wagner, C. (1981). *Rational consensus in science and society*. Dordrecht: Reidel.

[23] It is also worth mentioning that the models used in our simulations are utterly simple and avoid the complexities usually associated with simulations that aim to model social phenomena; see Hegselmann (1996). Another fact worth noting is that simulations in which the agents were allowed to move more or less freely in a two-dimensional environment (in the way of Gaylord & D'Andria, 1998), and in which they updated in certain ways both on the data they received and on the opinions of those peers they happened to meet (where 'meet' was defined as in ibid., Ch. 1), also gave essentially the same results as were obtained in the version of the Hegselmann–Krause model defined in Section 3. Lastly, to those who harbor more general doubts about the use of computer simulations in the study of social phenomena, I would like to recommend reading Lehtinen & Kuorikoski's excellent (2007). These authors convincingly argue that for the purposes of studying social phenomena, simulations can have real argumentative force, as opposed to a merely heuristic or illustrative value. In fact, they even go so far as to conclude that 'the information [computer simulations] provide is epistemically just as relevant as the information provided by an analytical proof' (ibid., p. 325; though I think it would have been more in line with what they show in their paper had they concluded that the information provided by simulations *may* epistemically be as relevant as that provided by an analytical proof). See also Hartmann (1996) for an illuminating account of the important role that computer simulations have come to play in both the natural and the social sciences.

Lehtinen, A., & Kuorikoski, J. (2007). Computing the perfect model: Why do economists shun simulation? *Philosophy of Science, 74*, 304–329.
Lorenz, J. (2007). Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics, C18*, 1819–1838.
Oddie, G. (2008). Truthlikeness. In S. Psillos, & M. Curd (Eds.), *The Routledge companion to philosophy of science* (pp. 478–488). London: Routledge.
Ramirez-Cano, D., & Pitt, J. (2006). Follow the leader: Profiling agents in an opinion formation model of dynamic confidence and individual mind-sets. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology* (pp. 660–667).
Rescher, N. (1973). *The coherence theory of truth*. Oxford: Clarendon Press.
Rosenkrantz, R. (1981). *Foundations and applications of inductive probability*. Atascadero CA: Ridgeview Publishing Company.
Sosa, E. (Forthcoming). The epistemology of disagreement. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Social epistemology*. Oxford: Oxford University Press.
Wagner, C. (1978). Consensus through respect. *Philosophical Studies, 34*, 335–349.
Weatherson, B. (2005). Can we do without pragmatic encroachment? *Philosophical Perspectives, 19*, 417–443.
Weisbuch, G., Deffuant, G., Amblard, F., & Nadal, J. P. (2002). Meet, discuss and segregate! *Complexity, 7*, 55–63.