Synthese
DOI 10.1007/s11229-012-0132-9

A puzzle about belief updating

Carlo Martini

Received: 25 April 2010 / Accepted: 24 May 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract In recent decades much literature has been produced on disagreement; the puzzling conclusion being that epistemic disagreement is, for the most part, either impossible (e.g. Aumann (Ann Stat 4(6):1236–1239, 1976)), or at least easily resolvable (e.g. Elga (Noûs 41(3):478–502, 2007)). In this paper I show that, under certain conditions, an equally puzzling result arises: that is, disagreement cannot be rationally resolved by belief updating. I suggest a solution to the puzzle which makes use of some of the principles of Hintikka's *Socratic epistemology*.

Keywords Disagreement · Consensus · Epistemology of disagreement · Belief updating · Linear updating · Equal weight view · Socratic epistemology

1 A brief survey on updating and disagreement

Recent epistemological literature has seen widespread interest in the topic of disagreement and its counterpart, consensus. Resolution of a situation of disagreement is, by definition, a consensus and is, according to some authors, an inevitable outcome, at least under certain conditions. Notorious defenses of the impossibility-of-disagreement thesis are Aumann's theorem (Aumann 1976) and the Lehrer-Wagner theory of consensus (Lehrer 1976; Lehrer and Wagner 1981). Moreover, several other theories have been presented for weaker possibility-of-agreement claims: among many, the split-the-difference view (Elga 2007; Frances 2010) and the Hegselmann-Krause model for consensus (Hegselman and Krause 2002).

C. Martini (✉)
Philosophy & Economics, University of Bayreuth, 95440 Bayreuth, Germany
URL: http://www.martinicarlo.net
e-mail: carlo.martini@uni-bayreuth.de

Published online: 16 June 2012

What makes those theories alike is their common claim that a situation of epistemic disagreement, under certain conditions (e.g. the disagreeing parties are epistemic peers), requires updating one's opinion accordingly. In other words, disagreement is a source of evidence on the basis of which one should update a previously held opinion.

The foregoing theories diverge on *how* the updating process should take place, that is, which updating function should be used. A class of functions that stands out among others is the class of linear updating functions.[1] Dietrich and List (2010) show that any updating function that satisfies a number of desiderata commonly accepted in social choice theory must be a linear one.[2]

Linear aggregation functions are thus a very important class of functions both because they satisfy standard desiderata in social choice theory, and because they provide a straightforward algorithm for updating. Such functions are those used in Lehrer and Wagner (1981), Elga (2007), Hegselman and Krause (2002), and several other models for reaching a consensus from an initial situation of disagreement.

I argue in this paper that a puzzle arises when looking at the problem of disagreement from the point of view of belief updating ('updating' will henceforth mean only linear updating, unless otherwise stated). From that viewpoint, agreement turns out to be an impossible goal to achieve, except for a limited number of cases; namely, when we simply happen to arrive at the same belief independently. In the following sections I present the argument (Sects. 2, 3), outline three possible objections, and respond (Sect. 4). The final sections develop on the philosophical consequences of the argument, in particular for belief updating methods in the light of disagreement. In Sect. 5 I argue that the puzzle arises only if we assume that resolving disagreement should be a matter of updating on beliefs. If instead the focus is on information, and consequently, on information pooling, rather than belief updating, the puzzle does not arise. Along the same lines of argument, Sect. 6 sketches an approach to resolving disagreement which follows closely the method of Socratic *elenchus* and "bracketing", of Jaakko Hintikka's *Socratic epistemology* (Hintikka and Harris 1988; Hintikka 2007).

2 Amateur birdwatchers and professional forecasters

Imagine a conversation between you and your friend Karl. You are planning to do some bird-watching at the Isle of May this coming weekend; you've been there a few times and you know that the trip is not worthwhile if the weather is bad. It just so happens that you and Karl are also professional meteorologists. You think that the probability of rain for the weekend is .8, whereas he thinks it's .4. Imagine that these probabilities

[1] Bayesian updating is another common updating method, but it is not particularly fruitful if disagreement is the kind of evidence on which we are updating our beliefs. For example, if I disagree with Tim and Tom on proposition A, Bayesianism would tell me to calculate, from my prior belief $P(A)$, the posterior belief $P(A|Tim_{disagrees} \wedge Tom_{disagrees})$, or similarly, the posterior belief $P(A|P_{Tim}(A) = X \wedge P_{Tom}(A) = Y)$. That is, after realizing the disagreement with Tim and Tom on A, my new belief A^* should be equal to the posterior belief A conditional on the fact that Tom's priors for A are equal to a certain value X, and similarly for Tim. It is far from clear, however, how exactly such updating should be done. The advantage of the linear methods discussed in this section is that they provide a clear algorithmic solution to disagreement.

[2] **Theorem 1**. An aggregation rule satisfies universal domain, collective rationality, independence and implication preservation if and only if it is linear. (Dietrich and List 2010, p. 11).

are needed for a decision table, on the basis of which you will decide whether to make the trip or stay home. In that case having a consensus on the probabilities involved is necessary for resolving your practical problem, but you and Karl are disagreeing.

What should you do? Being professionals, you decide that a weighted average of yours and Karl's forecasts is the best solution for your case. Weighted averaging implies that there should be a weight attached to your probability assessment. Let us leave aside, for the moment, the problem of where this weight comes from, and of what Karl should do, and focus instead on your strategy for updating your belief and thus moving closer to Karl's belief. You think that the probability of your forecast being correct is .7. Given your present state of belief, you are now claiming that (Q): "the probability of rain for the weekend is .8, and the probability of my statement being correct is .7". There is, in other words a certain probability (in this case .3) that the correct assessment of the probability of rain is not .8 but some other (confidence) value.

Let us formalize the previous assumptions. Q states that the probability of rain is equal to .8, but also that your assessment has a .7 probability of being correct; this part of Q translates as '$.8 \cdot .7$'. What Q also implies is that there is a certain probability that you may be wrong; in particular, there is a '$1 - .7$' probability that the probability of rain is some unknown value x, different from .8. Your full updated belief then, formally, is the following: '$.8 \cdot .7 + x \cdot (1 - .7)$' (where $x \neq .8$); let us call this R. There is now something awkward about the conjunction of your beliefs: You believe that the probability of rain is equal to .8, as part of the statement Q. At the same time, you concede that there is a .3 probability that you might be wrong, that is, you believe that the probability of rain is equal to R, above. The two beliefs, R and Q, should be equal, since they express the same thought, namely that "the probability of rain next week at the Isle of May is ......". If the two beliefs state the same, that is, they state your degree of belief in the event of rain, then $.8 \cdot .7 + x \cdot (1 - .7) = .8$. Given the latter equivalence, however, the only way for it to be true is if x is equal to .8, which was excluded in the definition of R because, intuitively, x should take a value that is different from the one reported as the probability of rain.

Let us generalize from the previous example: Call your belief in a certain event (e.g. rain-next-weekend) 'a', and your stated probability of it being correct 'y'. You should then believe that the probability of that event is '$a \cdot y + b \cdot (1 - y)$' (where b is some unknown probability value), and this should be equal to a. In other words, the following equivalence, let us call it E, should be true: $a \cdot y + b \cdot (1 - y) = a$. To solve $ay + b(1 - y) = a$ for y we need to distinguish two cases: (1) if $a \neq b$ then $y = \frac{a-b}{a-b} = 1$. Claiming that $y = 1$ is equivalent to claiming that the accuracy given to your initial estimate on the probability of rain (a) is 1.0 (100% accurate). (2) if $a = b$ then the equivalence reduces to $a \cdot y + a \cdot (1 - y) = a$.[3] This latter equivalence is tautological, and in that case there is no updating involved: if $a = b$, then I am claiming that there is a '$y + (1 - y)$' probability that my initial estimate is correct, that is, as before, there is a 1.0 probability that my estimate is correct. Elementary arithmetic then shows that

[3] This case is given here only for completeness, it should indeed be clear from the sentence in the natural language being formalized here (i.e. Q, above), that a should not be equal to b.

the only way for E to be true is if $a = b$, which goes against the intuition that in the formalization of Q the two cannot be the same.

Moreover, if $y = 1$, as it is the case when $a = b$, no updating of your beliefs can ever occur, since what you are saying is that "the probability of rain tomorrow is .8, and the probability that this statement is correct is 1.0". If you cannot coherently express any doubt about your own beliefs, you will never be in the position to reach an agreement with Karl, for instance by taking a weighted average of yours and his belief on the probability of rain. From what was argued so far, therefore, agreement seems impossible except in those cases when no updating is required, namely, when you and Karl simply happen to hold, independently, the same belief at the same time.

3 A principle of consistency

3.1 Formulation of the principle

The impossibility argument illustrated above relies on the principle that one cannot rationally hold two mutually inconsistent beliefs at the same time. When it is expressed in these terms the idea is simply a requirement of consistency, but the formulation of such principle is not as intuitive in the context of belief updating.

> *Consistency*: If your degree of belief in event e ('B(e)') is equal to r, then any conjunction of B(e) with other beliefs should either entail that $\mathrm{B}(e) = r$ is false, or be consistent with $\mathrm{B}(e) = r$.

The principle is not trivial. Suppose that I believe that the ferry service to the Isle of May does not run on rainy days. Any relevant information I receive will influence my present stock of beliefs, but any additional belief I may gain from that information must either disqualify that initial belief (e.g. "the ferry *does* run on rainy days!"), or be compatible with it.

Suppose, for instance, that I come to know that there is another service, a motor-boat, going to the island. In that case the new belief is consistent with the former. Suppose, instead, that I come to know that the administration has recently changed its policies, and now ferries run on rainy days too. This belief, if I acquire it, will disqualify the belief that ferries do not run on rainy days. What cannot happen is that I acquire a new belief that is in contrast with the former, and hold the former true at the same time. The latter case, however, is what happens in the amateur birdwatchers example.

3.2 Applications of the principle

Let us apply *Consistency* to the case of the amateur birdwatchers. In that example, you express your belief that it will rain with probability .8. However, after realizing that you are, after all, a fallible being, you restate your belief as the probability of rain being '$a \cdot y + b \cdot (1 - y)$' (the latter expression is the formalization of "the probability of rain is .8, but I might be wrong", where $a = .8$). The problem is that the latter belief does not eliminate the former, but rather makes use of it. Indeed, for the expression of fallibility to be somehow determinate, your belief in the probability of rain should take a specific, even if arbitrary, value, which in this example was chosen to be .8.

Let us also see how the principle is violated in the (Lehrer and Wagner 1981) updating process, one of the updating processes mentioned in Sect. 1. They provide an updating function that takes a weighted average of the opinions of, for instance, two agents, over a certain probabilistic event e. While the Lehrer-Wagner model is not limited to probabilistic events, for its relevance to the topics of this paper, in the following I will limit myself to the discussion of that case.

In a typical example of belief updating in the Lehrer-Wagner model, agent K' assigns a weight to herself and to her fellow agent K'', and so does agent K''. All weights w_{ij} are then normalized in the stochastic matrix W.

$$W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$$

In addition, agents K' and K'' express their opinion regarding the event on which they disagree: as shown in column-vector V, P_1 and P_2 are, respectively, K''s and K'''s opinions.

$$V = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}$$

The consensual value P_C will be obtained by the function $W^n \cdot V$ where $n \to \infty$. At each step ($W^1 \cdot V$, $W^2 \cdot V$, ..., $W^n \cdot V$) of the updating process, agent K''s belief will be $w_{11} \cdot P_1 + w_{12} \cdot P_2$. It is evident by pure comparison of this latter expression with its equivalent in the amateur birdwatchers case of Sect. 2 that the two situations are identical: the expressions '$w_{11} \cdot P_1 + w_{12} \cdot P_2$' and '$a \cdot y + b \cdot (1 - y)$' can be substituted with one another, and *Consistency* applies to the Lehrer-Wagner model as well.

It should be clear from the foregoing remarks that both the example of the amateur birdwatchers and the Lehrer-Wagner model are only instances of a general problem of belief updating, when the evidence on which an epistemic agent updates is evidence of disagreement. In fact, the problems are more general and elicit a puzzle, as I will show in Sect. 5; but first, the next section will deal with a number of possible objections to what has been claimed so far.

4 Objections and replies

4.1 Objection 1

Objection 1: *Consistency* does not apply to the amateur birdwatchers example, or to any analogous example, because my initial belief on the probability of rain ("the probability of rain tomorrow is .8") occurs *before* I add the claim "I may be wrong", and thereby implies that my revised belief that it will rain is '$.8 \cdot .7 + x \cdot (1 - .7)$'. If that is the case the equivalence $.8 \cdot .7 + x \cdot (1 - .7) = .8$ does not hold, blocking the proof that x must be equal to .8 (see Sect. 2).

Reply 1: The objection misses the point. It is clear that I first formulate a prediction on the probability of rain and, only afterwards, I revise it. Nonetheless, as it was claimed in Sect. 3.2, the revised belief must make use of the initial belief, and thus hold it true at the same time with the revised one.

If one would like to deny this, one needs to start from an *empty* belief, for example "the probability of rain is 'some specific value between 0 and 1', but I may be wrong". Even disregarding the fact that there is a sense in which such a sentence is absurd, it is clear that starting from an indeterminate belief leaves us no ground for updating, *a fortiori* for disagreement. In order to update on evidence I need to have some initial information, but when such information is my own belief on a certain issue (e.g. the probability of rain tomorrow), provided I do not give myself full confidence, I cannot avoid the contradiction of holding, at the same time, inconsistent beliefs.

4.2 Objection 2

Objection 2: The argument mixes up two kinds of probability, namely *objective* and *subjective* probabilities. On the one hand, the statement "the probability of rain is .8" is about objective probabilities, the probability that it will rain. On the other hand, the statement "the probability of rain is .8, but I might be wrong" is about subjective probabilities, that is, my *degree of confidence* as a forecaster. If the statements on the two sides of the equation '$.8 \cdot .7 + x \cdot (1 - .7) = .8$' refer to different types of probabilities, then, as before, *Consistency* does not apply, and the equation does not hold.

Reply 2: The objection is mistaken, because the argument is not committed to a specific type of probability and, in fact, holds for both subjective and objective probabilities. In the following, I illustrate how both kinds of probability can be justified in the argument, starting with subjective probabilities.

Suppose you and Karl use forecasting models to arrive at your predictions by means of your expertise in the field. You know perfectly well that often forecasting models are messy and unrealistic, though they may have good predictive power. Moreover, the final probabilistic judgment is also a function of the forecaster's background information, that is, expert information which comes neither from mathematical models nor from statistics or any other mechanical method for collecting information. The situation just depicted is quite common in several forecasting situations, and any interpretation using objective (or frequentist) probabilities would be highly misguided.

The probability of rain in this case has only to do with your "degree of knowledge", and is thus your subjective probability of the event 'it will rain next weekend'. Clearly, there is a straightforward justification for the claim that also the statement "I may be wrong with probability x" involves subjective probabilities. In this case both probabilities are subjective and the equivalence E of Sect. 2 holds: '$.8 \cdot .7 + x \cdot (1 - .7) = .8$'. Moreover, it could be argued that when the probabilities involved are subjective, by claiming that "the probability of rain is .8" we have already factored in the possibility that we might be wrong. In that case, the complete sentence Q would make little sense, if any at all.

In contrast, suppose now that your expertise rests entirely on your capacity to analyze and compute statistical data, so that no personal or unstructured information ever enters your judgments. In that case, the best interpretation for 'it-will-rain-next-weekend' is that the statement expresses objective probabilities. In addition, suppose that all weather forecasters are registered at the National Bureau of Forecasting (henceforth, NBF), where they keep track of all successful and unsuccessful forecasts, yielding a database of success and failure rates by which they can weigh a forecaster's predictions. Given this situation, it seems legitimate to claim that the probability that you are right or wrong is also objective, like the probability of rain. In this case then both probabilities are objective, and the equivalence E holds here as well.

I have claimed that there are plausible interpretations for the birdwatchers case to be either about subjective or about objective probabilities, and it is possible to construct numerous similar examples. Clearly, the two kinds of probability should not appear together on the two sides of the equivalence E; but, whereas *objection 2* claims that the two kinds of probabilities must appear together, therefore making the equivalence E false, the examples described in *reply 2* have shown that the two kinds of probabilities *need not* appear together. In other words, there are suitable examples in which probabilities of the same kind occur on the two sides of E, and that is what gives rise to the puzzle outlined in the previous sections.

4.3 Objection 3

Objection 3: Obviously something has gone awry. Why can I not change my mind, given the meta-evidence that my belief on the probability of some event is, with a certain probability, mistaken? My belief that the probability of rain is .8 is a first-order belief, whereas my updated belief that the probability of rain is '$a \cdot y + b \cdot (1 - y)$' is a second-order belief. As in the previous two examples, if this is the case, then the equivalence E of Sect. 2 does not hold.

Reply 3: Let us assume that the beliefs on each side of the equivalence are indeed different, being first- and second-order, respectively. The following situation is conceivable: An International Bureau of Forecasting (henceforth, IBF) keeps track of the reports of the National Bureau of Forecasting. Since the latter is also fallible, like any forecaster, the IBF should help the NBF correct its estimates, as well as those of other countries' bureaux of weather forecasting.

Given this situation, it is easy to build an argument for a second-order paradox as follows: "I believe that the probability of rain is .8, and I think that I am right in believing that with probability .7". Call the entire sentence A. A report from the IBF would prompt me to revise my belief. For example, A has .9 probability of being correct, so that I am now claiming "A, but I might be wrong with probability .1". The situation is analogous to the initial situation, and a new equivalence is prompted: "$A \cdot .9 + X \cdot (1 - .9) = A$.

It is easy to see why the latter equivalence should hold for exactly the same reasons that made the previous equivalence, $a \cdot y + b \cdot (1 - y) = a$, hold (see Sect. 3). In order to block this second-order argument, it seems natural that one would call for a third order of beliefs. After all, the International Bureau of Forecasting can be taken to be a meta-meta-forecasting institution. This will again not block the formulation of a third-order argument against the possibility of updating, and so on, as it is clear that *Consistency* should apply at all levels. Recourse to nth order beliefs triggers a theoretically infinite regress, where at each step n it is possible to argue for the impossibility of agreement.

At this point there are two possible conclusions, neither of which is satisfactory. In the first case, one may want to argue against assuming infinitely many levels of beliefs; i.e. the regress should be stopped at *some* level n_x. Stopping the regress at level n_x, however, will only solve the problem for the impossibility-of-agreement argument that can be formulated at level n_{x-1}. But it will still be possible to formulate an argument against updating at level n_x. Alternatively, one may want to argue that at *some* level n_x, the argument leading to the impossibility cannot be formulated. This premise, however, seems arbitrary; why not, then, just deny that it is impossible to formulate the argument for first-order beliefs? If *Consistency* holds at all, it should hold at all levels.

On the other hand, one may be willing to admit that the regress does not stop. If so, as explained before, at each level n *Consistency* must hold; but also with infinitely many levels of beliefs there will always be, at any specific level, an argument that leads to the impossibility of updating. Regardless, any theory that admits to infinitely many levels of beliefs becomes cognitively absurd. In the example of Sect. 2 the probabilities were to be used in a decision table—if the paradox is to be taken as a paradox of belief updating, and about disagreement resolution, it seems absurd to admit to ideal cognitive agents with infinitely many orders of beliefs. While one may still hold that there is nothing against postulating infinitely many levels of beliefs in a formal model, the assumption does not solve the problem because, as it was said before, the puzzle can be reformulated at any level n.

5 Impossibility of agreement?

Is it possible to agree on a certain topic, and reach a consensus? Most theories admit to different forms of updating and aggregation in order to combine individual probabilities (or other suitable values) and reach a unique value. The arguments presented in this paper, however, imply that I cannot update my beliefs in the light of the possibility that my beliefs are wrong. In other words, I cannot rationally update my beliefs by means of a weighted average of: (a) whatever I believe to be the right value on the issue that is being debated and (b) whatever value I assign to the probability that I may be wrong. After all, the probability that I may be wrong, so the principle of *Consistency* implies, must always be zero.

This conclusion is obviously puzzling for many views according to which it is rational to update one's beliefs in the light of disagreement, and this is because updating on disagreement typically implies to take some average over my own belief and the belief of my disagreeing party (in the case of disagreement between two epistemic agents only). This is the first consequence of the application of *Consistency* to the problem of disagreement (see Sect. 3.1). Similarly, the conclusion is just as puzzling because disagreement may derive from sources that are not other epistemic agents or the like. For instance, disagreement may derive from evidence which I myself construct while learning about a certain quantity. Imagine a situation where I calculate a bill twice, and end up with different results; one could imagine that there are, in this case, two discordant selves with which I have to deal epistemically.

The puzzle, however, arises only insofar as what is being updated are the beliefs of the agents in question. There is nothing wrong if what is being updated is, instead, information: Suppose I say, of Karl, the following: "he believes it will rain with probability .8, but there is a .7 probability that he might be wrong". In this case, Karl's belief that it will rain is .8, but my belief in the probability of rain, as reported by Karl, given that *he* is a fallible being ("*he* might be wrong"), is '$.8 \cdot .7 + x(1 - .7)$'. What changes from the first-person to the third-person perspective is that, when considering the probability of rain, in the former case I am speaking about my beliefs, or so it is most commonly assumed. But in the third-person perspective the report of Karl's belief in the probability of rain is only information, which I can freely feed as input of the updating function that weighs Karl's belief (.8) with his associated reliability score (.7).

In the following, I will argue that taking the updating to be 'information updating', rather than 'belief updating', both solves the puzzle that arises from updating in the light of disagreement, and suggests an alternative solution to updating in cases of epistemic disagreement.

6 Resolving a situation of disagreement

What we learn from a situation of disagreement like the one of the amateur birdwatchers of Sect. 2 is what is logically implied by the beliefs of the disagreeing agents. Assuming that there is an objective "state of affairs", what is implied is that *it must be the case that one of us* ['us' meaning you and Karl, or whoever is implied in the 'us'] *is wrong*.[4] What is not implied is that we should take the average of the two beliefs to be the new rational belief in which we should put our confidence. In fact, what I have argued is that any such average is in conflict with the principle of *Consistency* above, contrary to some alternative conciliatory views on disagreement (e.g. Elga's equal weight view).

Once we have ascertained that *it must be the case that one of us is wrong*, it seems reasonable that you should treat yours and Karl's probability-of-rain-reports as information, rather than as beliefs which you would commit yourself to. This approach, only introduced in this section, would seek to move past a situation of disagreement by means of further inquiry into the causes of the disagreement. For simplicity, let us imagine the extreme case where the disagreement is caused by two ideal selves: the scenario where I run a calculation twice and come up with a different result each time. In this case it seems absurd (and in fact it is) to say that the most reasonable belief to hold is a weighted average of my previous calculations. If I have any mathematical competence whatsoever, chances are that one of the two calculations I have performed,

[4] Clearly, it *could* be the case that we are both wrong.

rather than a third and completely unrelated one, will be the correct one. What any reasonable epistemic agent would do in this case is to go over the calculation once again, and try to find the mistake.

This approach, it is suggested here, would allow an epistemic agent to deal with a situation of disagreement with other agents while, at the same time, avoiding the puzzle discussed in the earlier sections. Clearly, this approach also implies that disagreement is not irreducible, but the foregoing sections should have proven that *genuinely* irreducible disagreement, if it were to occur, would not be rationally resolvable by updating.

In the space of this paper, this approach can only be sketched, and its full specification has to be left for future work. It is however possible to identify two main features that an interrogative model for disagreement would imply[5]: (a) a mechanism for acquiring new information in a situation of disagreement; and (b) a mechanism for stopping a certain piece of information from being open to further inquiring (in the work of Hintikka this is called *bracketing* (Hintikka 2007, Chaps. 1, 7)). A brief description of the rationale for the two mechanism is given in the following.

A. A method for gaining new information from a situation of disagreement would make sure that one can move past an otherwise unresolvable epistemic deadlock, similarly to the strategy of Hintikka's "logic of inquiry", by inquiring about the *causes* of disagreement, that is by asking further questions on the reasoning behind each of the disagreeing beliefs, and by adding the answers to the stock of available evidence. To be sure, deadlocks may still be possible, if one could show that a certain situation of disagreement is truly irreducible, that is, if no further enquiring is possible, and thus no further evidence is, even in principle, available. Situations of so-called "persistent disagreement" or "irreducible disagreement" are commonly mentioned in the relevant literature (Kelly 2010, 2005), but it is often not clear whether such situations are common or even rationally possible, assuming of course that the subject of inquiry is on matters of fact, that is, that the disagreement is *factual disagreement*.[6]

B. One might want to guarantee that certain items of information are less open to "revision by further inquiry" than others (see Hintikka 2007, p. 20). For instance, when there are two results of the same calculation, and one of the two is confirmed by scores of mathematicians, whereas I calculated the other one myself by hand; in that case I might want to give full weight to the hypothesis that my calculation is the incorrect one. Similarly, admitting that further inquiry is not justified in the presence of disagreement may find a rationale in all those cases where the epistemic agents are on different grounds with respect to the subject matter that is the object of disagreement: when one of the agents is cognitively impaired, when she is clearly lacking relevant information, etc. This approach would have the advantage of treating cases of disagreement among both epistemic peers and non-epistemic peers under a unified model, whereas most of the literature on epistemic disagreement treats the two separately.

[5] While the idea for such a model is taken primarily from the work of Hintikka (2007), the analogy with his "logic of interrogative inquiry" (Hintikka and Harris 1988) is here only a loose one, which would have to be established on firmer grounds.

[6] It is indeed quite rational to think that there may be several (and very rational) cases of non-factual irreducible disagreement, but also that different arguments apply when considering the factual versus non-factual cases of disagreement, as Martini et al. (2012) argue.

It should be clear that the model of interrogative inquiry in the case of disagreement avoids the dilemma outlined in the preceding sections because it avoids any consideration about averaging between two (or more) agents' beliefs. All that the disagreeing epistemic agents need to accept, in this model, is the logical consequence of their situation of disagreement, namely the fact that one of them must be wrong (possibly both). In the process of further inquiring, just like in Hintikka's "logic of interrogative inquiry", all that one needs to assume is information, that is, propositional content and not propositional attitudes such as 'belief' or 'knowledge'. Hintikka (2007) claims that "the entire theory of knowledge acquisition can [...] be developed without using the notion of belief." While it was not my intention in this paper to defend that claim in general, Hintikka's statement can easily be applied to knowledge acquisition in the context of epistemic disagreement.

7 Conclusion

In this paper I have argued that a puzzle arises in the context of belief updating in the light of disagreement. The puzzle arises because it is irrational to update on the belief that 'an event will happen with a certain probability x' on the basis of an added fallibility clause, e.g. "I might be wrong", or "I might be wrong with probability y". The foregoing arguments should have cast some doubt on the rationality of several conciliatory positions on disagreement; among many, the equal weight view, the split-the-difference view, and others (Elga 2007; Frances 2010). At the same time, the puzzle was meant to highlight the fact that disagreement cannot be resolved by means of rational updating on beliefs, in contrast with Aumann's famous result (Aumann 1976). The puzzle shows that Aumann's argument is not sound if, in his Bayesian framework, what we are updating on are beliefs.

Aumann (1976) speaks explicitly about 'beliefs' and 'opinions', but the choice is unfortunate. In Sect. 6 I suggest that in order to move away from the apparently puzzling conclusions of this paper, we should focus our epistemology on the concept of information rather than belief. Moreover, focusing on the concept of information allows us to treat the problem of disagreement with a unified account, where disagreement among epistemic peers is on a par with disagreement among epistemic non-peers. What is needed for such a unified account are: (1) a criterion for establishing when further analysis of the *reasons* for holding a certain propositional content true is justified, and when it is not; and (2) a method for analyzing the reasons of disagreement. In this paper I could only sketch the general features of such an account, and a more precise formulation will have to be left for future work; but I suggested that a promising avenue for developing it can be found in Hintikka's logic of interrogative inquiry.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Aumann, J. R. (1976). Agreeing to disagree. *The Annal of Statistics, 4*(6), 1236–1239.

Dietrich, F., & List, C. (2010). The aggregation of propositional attitudes: Towards a general theory. In T. S. Gendler (Ed.), *Oxford studies in epistemology* (Vol. 3). Oxford: Oxford University Press.

Elga, A. (2007). Reflection and disagreement. *Noûs, 41*(3), 478–502.

Frances, B. (2010). Disagreement. In D. Pritchard & S. Bernecker (Eds.), *Routledge companion epistemology*. London: Routledge.

Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation, 5*(3), 1–33.

Hintikka, J. (2007). *Socratic epistemology—Explorations of knowledge-seeking by questioning*. Cambridge: Cambridge University Press.

Hintikka, J., & Harris, S. (1988). On the logic of interrogative inquiry. *PSA: proceedings of the biennial meeting of the philosophy of science association* Volume One—Contributed Papers, pp. 233–240.

Kelly, T. (2006). The epistemic significance of disagreement. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (Vol. 1). Oxford: Oxford University Press.

Kelly, T. (2010). Peer disagreement and higher order evidence. In R. Feldman & T. A. Warfield (Eds.), *Disagreement*. Oxford: Oxford University Press.

Lehrer, K. (1976). When rational disagreement is impossible. *Noûs, 10*(3), 327–332.

Lehrer, K., & Wagner, C. (1981). *Rational consensus in science and society*. Dordrecht: Reidel.

Martini, C., Sprenger, J., & Colyvan, M. (2012). Resolving disagreement through mutual respect. *Erkenntnis, 64*(2). doi:10.1007/s10670-012-9381-8.