

An interpretation of weights in linear opinion pooling

Jan-Willem Romeijn
University of Groningen

Abstract

This paper explores the fact that linear opinion pooling can be represented as a Bayesian update on the opinions of others. It uses this fact to propose a new interpretation of the pooling weights. Relative to certain modelling assumptions the weights can be equated with the so-called truth-conduciveness known from the context of Condorcet's jury theorem. This suggests a novel way to elicit the weights.

1 Introduction

Say that Raquel consults her peer Quassim about the proposition S . Raquel's belief is $P(S) = r$, Quassim's is $Q(S) = q$. We assume regularity for the agents, i.e., that neither agent is absolutely certain about S , so we have $r, q \in (0, 1)$. How must Raquel respond to the opinion of Quassim? One influential model for updating opinions in the light of disagreement, due to French (1956) and Stone (1961), is linear pooling.¹ Linear pooling determines that the updated opinion of

¹The model has been developed by DeGroot (1974) and numerous others. For many philosophers the classical treatment is Lehrer and Wagner (1981). Genest and Zidek (1986) offer an

Raquel is given by $P^*(S) = r^*$ with

$$r^* = wq + (1 - w)r. \quad (1)$$

The weight $w \in [0, 1]$ specifies to what extent the updated opinion of Raquel will move towards that of Quassim. It expresses something like the trust that Raquel has in Quassim.

A different model for Raquel's accommodating Quassim's opinion is based on Bayesian updating. Quassim's opinion is in that case taken to be a proposition, and Raquel accommodates learning this proposition by a Bayesian update. A rigorous theory of how Bayesians can treat the opinions of others as evidence was first given in the context of game theory by Harsanyi (1966/67).² We bypass these mathematical foundations, and simply take for granted that the opinions of Raquel and Quassim can be captured coherently in an algebra, even if those opinions are also expressed in probability assignments over the algebra. Accordingly, we treat Quassim's belief in S as a random variable. For the event that Quassim believes S to degree q we write $\ulcorner Q = q \urcorner$, suppressing the argument S .

With these preliminaries in place, we can express Raquel's belief in S conditional on the event that Quassim believes S to degree q by Bayes' theorem:

$$P(S|\ulcorner Q = q \urcorner) = P(S) \frac{f(\ulcorner Q = q \urcorner|S)}{g(\ulcorner Q = q \urcorner)}. \quad (2)$$

Here f and g are the probability density functions derived from the cumulative distributions $P(\ulcorner Q < q \urcorner|S)$ and $P(\ulcorner Q < q \urcorner)$ respectively, both of which are assumed to be continuous relative to the Lebesgue measure over Q .³ Following extensive mathematical review, and Dawid et al. (1995) provides numerous details on relations to Bayesian inference.

²See Brandenburger (2007) for an accessible exposition of so-called interactive epistemology, in particular of the basic idea of Harsanyi type spaces.

³This is the standard format for updating probability assignments when the likelihoods are probability density functions. See Gelman et al (2020) for a general introduction into Bayesian

Bayes' rule, Raquel's belief after learning Quassim's opinion is:

$$P^*(S) = P(S|\ulcorner Q = q \urcorner). \quad (3)$$

We saw above that, if the pooling operation of equation (1) is used as the update method, we have $P^*(S) = r^*$. This paper is concerned with the relation between these two methods of updating, and the lessons we can draw from this relation.

It is known that we can replicate linear pooling in a Bayesian model. In the simple setting sketched above, we can orchestrate the Bayesian model to obtain

$$P(S|\ulcorner Q = q \urcorner) = r^*.$$

In particular, Genest and Schervish (1985) consider a Bayesian model of a decision maker learning the opinion of an expert, and investigate when the result of the update coincides with a general version of the linear opinion pool. They show that certain constraints on the model suffice, and that we do not need to specify the Bayesian model in full detail to obtain an outcome in accordance with linear pooling. This is in a sense expectable: there are several results showing the wide range of belief update operations that can be subsumed under a Bayesian header (cf. van Fraassen, 1989; Halpern, 2003).

As will be discussed below, the relation between pooling and Bayesian updating has been the subject of both philosophical and mathematical study. Some of this literature has focused on how Bayesian updating and linear pooling may be combined, e.g., on the question when the order of these operations matters (cf. Genest et al., 1986; Dietrich, 2010; Leitgeb, 2017). Other contributions focus on the ways in which linear pooling fails to accommodate the specifics of learning multiple expert opinions (Bradley, 2006; Steele, 2012; Bradley, 2018). In general,

statistics. To indicate that we are updating on a particular event, namely that Quassim reveals his opinion to be $Q(S) = q$, we will retain the notation $\ulcorner Q = q \urcorner$ also in the density function, instead of using the more customary single variable q .

Bayesian models cover a wider range of social deliberations than models based on the linear opinion pool, and the latter models violate various natural constraints on social learning (e.g. Dawid et al., 1995; Bradley, 2018; Dawid and Mortera, 2019; Baccelli and Stewart, 2020).

It remains worthwhile to investigate the relation between linear pooling and updating. We can learn a lot from the specifics of the Bayesian model that replicates the linear pool. If we adopt linear pooling as our updating method, we can use the Bayesian model to bring out the assumptions that we buy into, and thereby obtain conceptual insights into the pooling operation. Insofar as linear pooling adequately captures opinion change in social deliberation, the relation between pooling and Bayesian updating contributes to our grasp of how we shape our opinions through social exchange.

The specific innovation of this paper is an interpretation of the pooling weights in terms of beliefs. In the philosophical literature on linear pooling, the weight is mostly an exogenous component of the models (cf. Lehrer and Wagner, 1981, pp. 21–22). It is taken to express something like the trust that a decision maker has in an expert, or that peers have in each other. But this is odd: pooling is concerned with beliefs, and it might be expected that mutual trust can be spelled out in terms of beliefs as well. The current paper proposes such an interpretation: it relates the pooling weight to the so-called truth-conduciveness that we know from Condorcet’s jury theorem. An important benefit of this interpretation is that it facilitates the empirical elicitation of weights.

2 Bayesian models of linear pooling

This section presents a corollary of one of the results in Genest and Schervish (1985). After that we briefly discuss how the use of this result in the current paper relates to the existing discussion on Bayesian models of linear pooling.

2.1 Likelihood functions for the linear pool

The primary concern in the paper by Genest and Schervish is to specify the results of a Bayesian update on the opinion of an expert in the absence of a complete specification of the Bayesian model. They show that we do not need a full probability assignment over all the opinions. Under fairly broad assumptions on the probability model of the decision maker, the update yields a form of linear pooling. Genest and Schervish further discuss how the result can be generalized to multiple agents, and how certain assumptions on the conditional independence of expert opinions hang together with generalizations of linear pooling.

The theorem below only represents a small part of that. It restricts attention to two agents, to a simple algebra of a proposition and its negation, and to straightforward linear pooling only. In this limited setting, it can be established that the linear pool implies the use of particular likelihood functions $f(\Gamma Q = q^\neg | S)$. Conversely, the use of these likelihoods in Bayesian conditioning guarantees the equality of the results of pooling and conditioning. The theorem is actually fairly basic and the proof comes down to a simple rewriting of terms in the formula that equates the Bayesian update with the linear pool. It is recounted here because the expressions for the likelihoods lead to a new way to interpret the weights, and thereby to a new method of eliciting them.

THEOREM (AFTER GENEST AND SCHERVISH 1985)

Let $P(S) = r$, $Q(S) = q$, and let $r^* = wq + (1 - w)r$ be the result of linearly pooling these opinions. Then for all values $r, q, w \in [0, 1]$ we have $P(S | \Gamma Q = q^\neg) = r^*$ if and only if the likelihood functions satisfy the following constraints:

$$f(\Gamma Q = q^\neg | S) = g(\Gamma Q = q^\neg) \left(1 - w + \frac{w}{r} q \right), \quad (4)$$

$$f(\Gamma Q = q^\neg | \neg S) = g(\Gamma Q = q^\neg) \left(1 + \frac{r}{1-r} w - \frac{w}{1-r} q \right), \quad (5)$$

where g is the density function of the cumulative distribution $P(\Gamma Q < q^\neg)$.

PROOF

We first prove the theorem left-to-right: we assume that $P(S|\Gamma Q = q^\neg) = r^*$ and then derive equations (4) and (5). We rewrite Bayes' theorem for continuous likelihood functions in the following way:

$$f(\Gamma Q = q^\neg|S) = g(\Gamma Q = q^\neg) \frac{P(S|\Gamma Q = q^\neg)}{P(S)}$$

Substituting $P(S) = r$ and $P(S|\Gamma Q = q^\neg) = r^* = wq + (1 - w)r$ we obtain equation (4):

$$f(\Gamma Q = q^\neg|S) = g(\Gamma Q = q^\neg) \left(1 - w + \frac{w}{r}q\right).$$

Equation (5) can be derived in a similar manner.

Next we prove the theorem right-to-left: we assume the likelihoods of equations (4) and (5) and then derive that $P(S|\Gamma Q = q^\neg) = r^*$. This follows straightforwardly by filling in the likelihoods in Bayes' theorem, and by rewriting the terms:

$$P(S|\Gamma Q = q^\neg) = P(S) \frac{f(\Gamma Q = q^\neg|S)}{g(\Gamma Q = q^\neg)} = r(1 - w) + wq = r^*.$$

It follows immediately that $P(\neg S|\Gamma Q = q^\neg) = 1 - r^*$. \square

Notice that the density functions $f(\Gamma Q = q^\neg|S)$ and $f(\Gamma Q = q^\neg|\neg S)$ entail specific constraints on the probability density function $g(\Gamma Q = q^\neg)$. By the law of total probability we must have

$$P(S) = \int_0^1 g(\Gamma Q = q^\neg) P(S|\Gamma Q = q^\neg) dq.$$

In words, the prior must be a convex combination of the posteriors. Filling in the prior and the posterior determined by linear pooling we obtain

$$r = \int_0^1 g(\Gamma Q = q^\neg) (wq + (1 - w)r) dq.$$

With some algebra we derive

$$\int_0^1 q g(Q = q) dq = r. \quad (6)$$

The interpretation of this constraint on g is rather intuitive (cf. Bonnay and Cozic, 2018). In a Bayesian rendering of pooling, Raquel's distribution for Quassim's opinion is centred on her own. This seems natural as a constraint if Raquel is indeed planning to pool linearly: because she puts some trust in Quassim, any deviation in expectations would force Raquel to adapt her opinion even before she learns his opinion.⁴

The constraint (6) still leaves a considerable amount of freedom for the density g and hence for the likelihood functions of equations (4) and (5). For instance, Raquel might use a Beta-distribution showing a peak at $q = r$, expressing that she thinks it more probable that Quassim's opinion sits close to her own, while assigning decreasing probabilities to Quassim's opinion being further removed from hers. Another possibility is that Raquel expects Quassim to be opinionated on S , so that the density g peaks towards 0 and towards 1.

The above result offers specific conceptual benefits. For one, it may serve to spell out the conditions of adequacy of linear pooling as a method of belief change. There are well-known defenses of the Bayesian model of belief change, and we may fall back on these defenses. That is, we may motivate linear pooling as an epistemic shortcut by pointing to the probabilistic assumptions that the Bayesian model brings out, and by ascertaining that those assumptions are met in a particular situation. However, in what follows we will focus on a benefit of an interpretative nature.

⁴Genest and Schervish (1985) show that violating this constraint still leads to a version of linear pooling, just with an additional offset. But in what follows we will not need those results.

2.2 Other Bayesian models of pooling

There are several discussions of the relation between Bayesian updates and pooling operations in the literature. We will review some of these in order to position the current paper. Specific attention will be given to the conception of peers adopted here, and the restriction to linear pooling for two peers. This will bring the goal of the current paper more clearly into focus.⁵

Dawid et al. (1995) investigate the consistency of the linear opinion pool under specific constraints on the Bayesian model that underpins it. They consider a decision maker taking advice from multiple experts on a single proposition. Following DeGroot (1988), they adopt a definition of experts in terms of the Bayesian model: experts and decision maker start with a common prior, and only experts subsequently collect evidence, so that the decision maker is rationally compelled to defer to their probabilistic judgments. These assumptions together entail various constraints on the weights of the experts, including the counter-intuitive constraint that the weight of the decision maker for herself must be non-positive.

The study of the coherence of pooling multiple experts is continued in Bradley (2006), and subsequently in Steele (2012). In their papers it is again apparent that the linear opinion pool is restrictive in comparison to the Bayesian model. Apart from constraints that stem from the assumption of deference to experts, Bradley observes that the relations among experts are not adequately captured in the linear pool. The construction of the Bayesian model involves choices on the dependence or otherwise of the experts, choices that are external to the pooling model. This leads Steele to propose that the application of pooling operations has to be relativized to a context.

Importantly, the converse relation is less problematic: we can always construct a Bayesian model to replicate a linear pool with multiple agents, e.g., by build-

⁵This section benefited greatly from discussions with [omitted for blind reviewing].

ing up a Bayesian model in which the learning events are composed of multiple, simultaneously revealed opinions, or by converting it to a sequence of pooling operations with individual agents and computing the corresponding Bayesian updates for them. In particular, for the agents Quiana, Quassim and Raquel with weights v_2 , v_1 and $v_0 = 1 - v_1 - v_2$ respectively, we can rewrite the linear pool as

$$r^* = v_2 q_2 + v_1 q_1 + v_0 r = w_2 q_2 + (1 - w_2) [w_1 q_1 + (1 - w_1) r],$$

where $w_2 = v_2$ and $w_1 = \frac{v_1}{v_0 + v_1}$. The linear pool for Raquel with two agents is thus split into two consecutive pools with one agent, both of which can be written down as Bayesian updates with specific likelihood functions.

Clearly, none of this resolves the problems for linear pooling with multiple agents. The likelihood functions for consecutive pooling operations will be different depending on the order in which the agents are accommodated by Raquel, and they will encode specific dependencies among the agents' opinions: the impact of Quiana's opinion on Raquel's is different depending on whether Raquel first accommodated Quassim's opinion or not.⁶ This is entirely in line with the point in Bradley and Steele that, when compared with the Bayesian model, the opinion pool entails specific assumptions and therefore offers limited means for representing social deliberation, especially when multiple agents are involved.

The theme that linear pooling is a restrictive format for social deliberation has been discussed more recently too. Bradley (2018) shows that assuming the linear pool, the common prior and the principle of deference entail that the expert opinions need to be equal, a point further refined and corrected in Dawid and Mortera (2019). Baccelli and Stewart (2020) develop these results and argue that the linear pool is inconsistent with Bayesian models that comply to deference. They go on to prove that under certain additional assumptions the geometric pool is in fact

⁶In the final section we return to this point and its potential implications for the reach of the result from the current paper.

the only coherent pooling method. Insofar as these additional assumptions accord with the nature of social deliberation, the conclusion might have to be that linear pooling is simply not the right modeling tool.

2.3 The Bayesian model of this paper

The approach of this paper differs from that of the afore-mentioned ones. It does not take a Bayesian model with some additional assumptions as starting point, to investigate if and when we arrive at some form of pooling. Rather it takes the linear opinion pool as a starting point, and then investigates its Bayesian representation, without adopting any constraints on the Bayesian model at the outset. Although it is not ruled out, the common prior assumption does not feature in the current paper. More importantly, the current discussion does not give Quassim the role of expert in the sense of Dawid et al. (1995). Raquel does not defer to him, but considers herself epistemically on a par, as his peer.

This choice of approach hangs together with the goal of this paper. It is to clarify the workings of linear pooling and uncover the assumptions implicit in it by writing down its Bayesian model, so as to interpret the pooling weights and facilitate their elicitation. The afore-mentioned critical discussions may seem to suggest that this is a somewhat academic exercise. Admittedly, the approach of this paper makes most sense when we independently attach importance, normative or descriptive, to linear pooling as a means to model social deliberation. However, this importance can be argued for, by a principled defense of linear pooling (e.g. Pettigrew, 2019) or simply by the long history of such models in philosophy and computational sociology (e.g. Hegselmann and Krause, 2002). Moreover, we might contest the constraints on the Bayesian models that make linear pooling

incoherent, like the common prior and the principle of deference, on the grounds that they are too strict, or even unrealistic.⁷

To highlight the approach of the current paper, it will be helpful to return briefly to the Bayesian models of geometric pooling. What the current paper achieves for the linear pool, namely a new interpretation of the pooling weights through a Bayesian model of the pooling operation, we can also imagine for other pooling methods. In the Bayesian representation of geometric pooling, following Baccelli and Stewart (2020), Russell et al. (2015), EaGIHiVe (2016),⁸ and Dietrich and List (2016), the weights show up in the likelihood functions for the revealed opinions of others, directly modifying their impact. A clean expression of the weights in terms of belief can perhaps be distilled from these likelihood functions. For linear pooling, by contrast, the role of the weights is less clear: it seems more difficult to interpret the likelihood functions of equations (4) and (5). To arrive at interpretable expressions, the Bayesian representation of linear pooling is in this paper studied in its most basic format, involving only two peers and a single proposition. This allows us to extract a clear role for the pooling weights, relating these weights directly to the beliefs of the pooling peer. As will become apparent, this allows us to distill expressions for the weights in terms of beliefs after all.

3 Trust and truth-conduciveness

In this section we briefly review Condorcet’s result and a Bayesian reformulation of it. This will bring out a particular notion, namely the truth-conduciveness of jurors, that is crucial for the adequacy of a jury verdict. It is then shown that,

⁷This point of view rings through in some of the discussions responding to Dawid et al. (1995), especially those by Lindley, McConway, and Winkler.

⁸The acronym stands for Easwaran, Glynn, Hitchcock and Velasco. Its use is recommended in their paper.

subject to certain modeling assumptions, the pooling weight w can be equated with truth-conduciveness of a peer.

3.1 Truth-conduciveness in voting

Consider the setting of the original jury theorem of Condorcet (cf. Romeijn and Atkinson, 2011). Jurors are asked for a categorical vote V on a proposition S . It is assumed that jurors are competent, meaning that if S is true, they are more likely to vote in support of it, $V = 1$, while if S is false they are more likely to vote against, $V = 0$. The competence of a juror can be expressed by the values of two competence parameters, c_S and $c_{\neg S}$:

$$\begin{aligned} c_S &= P(\lceil V = 1 \rceil S) > \frac{1}{2}, \\ c_{\neg S} &= P(\lceil V = 0 \rceil \neg S) > \frac{1}{2}. \end{aligned}$$

The assumption of juror competence is thus that they are better than a fair coin in determining the truth or falsity of S . The original result of Condorcet is that for competent jurors, the majority vote of the jury will tend to the truth with increasing jury size. The theorem is essentially a version of the law of large numbers. If S is true, then for juries with increasing size the probability that more jurors vote for S than against will tend to 1, and conversely for $\neg S$.

There is also a Bayesian version of Condorcet's result, to the effect that the posterior probability of the true proposition, either S or $\neg S$, will tend to 1 when updating on ever more juror votes. To arrive at this result we need not assume that the jurors are competent, i.e., that they perform better than a fair coin. The assumption needed here is that the jurors are truth-conducive: it is more probable that jurors cast a vote for S if it is true than if it is false, and conversely for casting a vote against S . Formally, we may express this requirement as follows:

$$\Delta = P(\lceil V = 1 \rceil S) - P(\lceil V = 1 \rceil \neg S) = c_S + c_{\neg S} - 1 > 0 \quad (7)$$

where Δ is the truth-conduciveness of jurors. Notice that the above requirement expresses the truth-conduciveness for $\neg S$ as well because

$$P(\lceil V = 0 \rceil \neg S) - P(\lceil V = 0 \rceil S) = P(\lceil V = 1 \rceil S) - P(\lceil V = 1 \rceil \neg S).$$

If $\Delta > 0$, a vote $V = 1$ will increase the posterior probability of S and a vote $V = 0$ will increase the posterior of $\neg S$. Importantly, the truth-conduciveness of a voter is not the same as her having competences of larger than half. One of the two competences c_S and $c_{\neg S}$ may drop below that. We may for example have that $c_S = 1/4$ and $c_{\neg S} = 4/5$, in which case a vote in favour still directs more probability towards the truth of the hypothesis, and a vote against still directs probability towards its falsity.⁹

3.2 Voting as pooling

With the notion of truth-conduciveness in place, we can now start working towards the main result of this paper, which relies on the relation between voting and opinion pooling. We approach this relation by representing voting in terms of the more fine-grained Bayesian model of pooling. This will lead to an expression for the truth-conduciveness of jurors in which the trust parameter of pooling is the central term.

To this aim we view Quassim as the sole member of a jury that advises Raquel. But rather than taking his categorical vote at face value, we frame Quassim's vote in terms of the probabilistic opinions that he might have. That is, we make a more fine-grained model of the opinion that Quassim is expressing by his vote, so that we can forge a connection with the Bayesian model of pooling. Concretely, we identify the votes $\lceil V = 1 \rceil$ and $\lceil V = 0 \rceil$ with distinct events in the Bayesian

⁹See footnote 1 of Romeijn and Atkinson (2011) for a general expression of what supermajority is needed to get the required limiting behaviour, relative to the values of c_S and $c_{\neg S}$.

model of pooling, namely $\lceil Q > r \rceil$ and $\lceil Q < r \rceil$ respectively. We suppose that Quassim knows Raquel’s prior opinion, and that he is trying to nudge her over, i.e., he votes for S if his degree of belief in S is higher than hers, and against S if his degree of belief is lower.

Raquel thus conceives of Quassim’s vote V as a shorthand for a range of probabilistic opinions that he might have. She sees the event $\lceil V = 1 \rceil$ as a shorthand for Quassim expressing the opinion that $\lceil Q > r \rceil$, and similarly she takes the event $\lceil V = 0 \rceil$ to express that $\lceil Q < r \rceil$. Raquel may then accommodate Quassim’s opinion much like she accommodates a categorical vote, namely by a Bayesian update. In the remainder of this section we investigate what this leads to if we specify this update in terms of the Bayesian model of pooling, i.e., by means of the likelihoods (4) and (5).

The above translation between voting and pooling seems natural, but we might ask whether the specific representation of voting in terms of more fine-grained beliefs is the only reasonable way to relate the two. What justifies setting it up in this way? We can answer this in section 4, after we have seen what role the pooling weight obtains in this setup.

3.3 Weight as truth-conduciveness

To fully specify the Bayesian update that Raquel performs, we need to determine the density function $g(\lceil Q = q \rceil)$, which expresses Raquel’s expectations about Quassim’s opinions. Raquel might choose:

$$g(\lceil Q = q \rceil) = \begin{cases} l(q) & \text{if } q < \epsilon r \\ h(q) & \text{if } q > 1 - \epsilon(1 - r) \\ 0 & \text{else.} \end{cases} \quad (8)$$

where $0 < \epsilon \leq 1$. Further, we assume for now that the functions l and h are constant. If we take $\epsilon = 1$, Raquel distinguishes between Quassim offering a lower or a higher degree of belief in S , but within these two ranges the distribution is uniform. For small ϵ , Raquel takes Quassim to offer a probability below ϵr and hence close to zero, or above $1 - \epsilon(1 - r)$ and hence close to one. A distribution with small ϵ expresses that Raquel takes Quassim's opinions to resemble categorical votes. In short, the parameter ϵ expresses how opinionated Raquel takes Quassim to be.

If we want the Bayesian update of Raquel to replicate pooling, the density g must comply to the constraint (6). Solving for this under the assumption that l and h are constant yields:

$$l = \frac{1 - r}{\epsilon r}, \quad h = \frac{r}{\epsilon(1 - r)}.$$

Notice that this entails

$$P(\Gamma Q < r^\neg) = 1 - r.$$

This means, as intimated earlier, that Raquel finds it more probable that Quassim will report a value $Q = q$ that is lower than her opinion when she is herself assigning a low probability r to S . This is not specific for the distribution above but follows from equation (6) itself, and hence holds for all Bayesian models that replicate pooling. To accord with pooling, Raquel must always expect that Quassim's opinions will reinforce what she is already tending towards.¹⁰

Together with the likelihoods of equations (4) and (5), the step function (8) offers a Bayesian representation of Raquel's pooling operation, if she were given a sharp probability by Quassim. However, Quassim merely offers his opinion in the form of a categorical vote, which is then interpreted as either of the events

¹⁰Depending on your point of view, this might again cast doubt on linear pooling as a method for adapting one's beliefs to others. At any rate it is a virtue of the Bayesian model of pooling that it brings such implicit assumptions to light.

$\lceil Q > r \rceil$ or $\lceil Q < r \rceil$. Fortunately the Bayesian representation still allows us to compute an updated probability for Raquel, using the marginal likelihoods of S and $\neg S$ for the event $\lceil Q > r \rceil$, to wit, $P(\lceil Q > r \rceil | S)$ and $P(\lceil Q > r \rceil | \neg S)$, as well as the corresponding marginal likelihoods for the event $\lceil Q < r \rceil$.

We now determine these marginal likelihoods. For the event $\lceil Q > r \rceil$ we find

$$\begin{aligned} P(\lceil Q > r \rceil | S) &= (1-w)r + w - (1-r)w\frac{\epsilon}{2}, \\ P(\lceil Q > r \rceil | \neg S) &= (1-w)r + rw\frac{\epsilon}{2}. \end{aligned}$$

Similar equations can of course be derived for the event $\lceil Q < r \rceil$. We can now observe that

$$P(\lceil Q > r \rceil | S) - P(\lceil Q > r \rceil | \neg S) = w \left(1 - \frac{\epsilon}{2}\right) \quad (9)$$

and that the same expression can be derived for $\lceil Q < r \rceil$. A certain measure for the impact that the event $\lceil Q > r \rceil$ has on Raquel's opinion of S and $\neg S$, namely the difference of the likelihoods for these events, is proportional to the weight. And if we take the limit of ϵ to zero, making Quassim more and more opinionated, this proportionality is replaced by an equality. We have thus arrived at a clean expression of the pooling weight in terms of two key opinions of Raquel.

Recall that in the context of voting, the truth-conduciveness of equation (7) showed up as a crucial quantity. It will be clear that equation (9), pertaining to weight in the context of a Bayesian reconstruction of pooling, shows similarity to equation (7). Drawing on the parallel between votes $V = 1$ and $V = 0$ on the one hand, and opinions $Q > r$ and $Q < r$ on the other, we can talk about truth-conduciveness in the context of pooling as well:

$$\Delta = P(\lceil Q > r \rceil | S) - P(\lceil Q > r \rceil | \neg S).$$

In words, we label the difference between the likelihoods of S and $\neg S$ for the event $\lceil Q > r \rceil$ as the truth-conduciveness that Raquel attributes to Quassim.

Hence we have

$$\Delta = w\left(1 - \frac{\epsilon}{2}\right). \quad (10)$$

For diminishing values of ϵ , the match between truth-conduciveness and pooling weight becomes exact, $\Delta = w$. We can therefore think of w as the truth-conduciveness that Raquel attributes to Quassim.

4 Interpreting the weights

The above translation between pooling and voting is set up to help us derive a clean expression for weights in terms of beliefs. The Bayesian representation of pooling allows us to locate it among the opinions of the peers. In this section we elaborate on a particular advantage of this, namely for eliciting weights. The result is then contrasted with other interpretations of the weights in the linear opinion pool.

4.1 Weight elicitation

We first consider a central assumption that was made in forging the link between the models of voting and pooling, namely the choice of a particular step function as distribution over $\lceil Q = q \rceil$. This choice may seem arbitrary. However, the exact shape of the functions $l(q)$ and $h(q)$, which were taken as constants in the foregoing, loses import if we take ever smaller values of ϵ . To satisfy constraint (6), the probability masses $P(\lceil Q < \epsilon r \rceil)$ and $P(\lceil Q > \epsilon(1-r) \rceil)$ must approximate $1-r$ and r respectively as ϵ gets smaller, irrespectively of the shape of $l(q)$ and $h(q)$. The average values of $l(q)$ and $h(q)$ over their respective intervals must therefore approximate $\frac{1-r}{\epsilon r}$ and $\frac{r}{\epsilon(1-r)}$. Computing the marginal likelihoods on this basis yields equation (9) as an adequate approximation.

The upshot is that if we push more and more probability mass in the distribution $P(\ulcorner Q < q \urcorner)$ towards the extremes, the exact shape of this distribution matters less and less. Moreover, it is precisely in this limit that the weights are best understood as expressing truth-conduciveness. For the purpose of interpreting pooling weights we may disregard the shape of the density $g(\ulcorner Q = q \urcorner)$, as long as we assume that it concentrates on extreme values, i.e., that almost all its probability mass sits close to 0 or 1. If we assume this, we can obtain a clean expression for the pooling weight of the agent, connecting the weight of the agent to their truth-conduciveness. To derive this expression and interpret the pooling weight accordingly, we must ensure that $g(\ulcorner Q = q \urcorner)$ is concentrated on the extreme values. Equivalently, we must ensure that Raquel thinks of Quassim as highly opinionated about S .

Recall that the foregoing Bayesian model of linear pooling was set up to resemble categorical voting. For such categorical votes it is entirely appropriate to assume that the juror is opinionated, i.e., committed to S or to $\neg S$, and therefore that the density g is peaked at values close to 0 and 1. In the background of this is the Lockean thesis: above a certain threshold for the probability, $Q(S) > t$, Quassim is warranted to express a categorical belief. By choosing $\epsilon = 1 - t$ in Raquel's distribution $P(\ulcorner Q < q \urcorner)$, all opinions of Quassim $Q(S)$ that have nonzero probability by Raquel's lights will surpass the threshold one way or another, so that Raquel may consider the categorical votes $V = 0$ and $V = 1$ to be adequate representations for Quassim's opinions. In short, we can use the model specified above as long as we ensure that Raquel imagines Quassim to be in the role of a categorical juror.

We can now say how we might use the above model in weight elicitation. We can determine the weight that Raquel attaches to Quassim by asking her assessments of his error probabilities for S and $\neg S$ under the assumption that he is opinionated. Specifically, we can ask Raquel to imagine that Quassim is a ju-

ror who returns a categorical judgment on S because he is quite sure, one way or another, and then we can read off the pooling weight from her assessment of Quassim's competences. Here is a simple vignette to do this:

Imagine that Quassim has little doubt about S . In line with his firm opinion, he is going to report either that it is true, or that it is false.

1. How probable is it that he will report S to be false, if actually S is true?
2. How probable is it that he will report S to be true, if actually S is false?

The probability judgments are related to the competences that Raquel assigns to her peer Quassim, with the first question returning $1 - c_S$ and the second returning $1 - c_{\neg S}$, so that Δ can be computed from them using equation (7). Moreover, because we have asked Raquel to imagine Quassim to be opinionated, we can employ the above model to determine w as well.

Let us pause briefly over what this brings us.¹¹ By way of comparison, consider the well-known elicitation method for credences through betting contracts. Raquel's latent credence is revealed in how she responds to a situation engineered for the purpose of elicitation, specifically, in her decisions to accept and reject various bets. The proposal above is much the same: we can determine a latent characteristic of Raquel's epistemic state, in this case the pooling weight she assigns to Quassim, through her responses to a pre-designed situation. The foregoing establishes that in particular situations the pooling weight is equal to the difference between two credences, and the elicitation consists in having Raquel imagine the situation and determining her conditional credences.

¹¹This paragraph and the next were added on the suggestion of an anonymous referee, whose help is gratefully acknowledged.

The fact that we need a specific situation to determine the pooling weight that Raquel gives to Quassim, for instance that we need Raquel to imagine Quassim to be opinionated, does not prevent us from applying this pooling weight in different situations, for instance in ones where she knows Quassim is not opinionated. If we assume that Raquel is coherently pooling opinions, we can arguably represent her epistemic state in the Bayesian model provided above. The pooling weight is a general characteristic of that epistemic state, and so is the equality of the pooling weight to truth-conduciveness in particular situations. Much like singular credences, after having been determined through betting, are applicable outside the betting context, the pooling weight is applicable in situations that differ from the one needed for elicitation purposes.

This is a admittedly just a first step towards a principled and experimentally feasible way of eliciting pooling weights. Apart from the usual problems with probability elicitation, we must somehow ensure that the imaginings of Raquel match the assumptions of the model, so that we can indeed equate w and Δ . For instance, we must ascertain that Raquel takes Quassim to be opinionated when told that he has “little doubt” about S and that he is “going to report on it [categorically]”. An actual experiment will likely involve a more worked-out vignette than the one given above, perhaps involving betting contracts rather than straightforward questions about credence. Nevertheless, the foregoing suggests a way forward for the problem of weight elicitation.

4.2 Existing interpretations of weights

The present paper offers a new interpretation of pooling weights in terms of beliefs, and an accompanying suggestion for how to elicit them. This subsection discusses several earlier attempts at interpreting linear pooling weights, following Genest and McConway (1990). It will be argued that the extant interpretations

do not connect the weights properly to epistemic attitudes, or else that they fail to offer a handle on determining the weights. Arguably the current proposal ticks both these boxes.

Genest and McConway review several interpretations of the pooling weight. Their simplest proposal is that the weight is a veridical probability, that is, the probability that an agent's opinion is correct. But this leads to problems with the truth conditions for probabilistic opinions, as they readily admit. Alternatively, the weight may be the probability that an agent outranks all others. Here again, we are left in the dark on what this ranking is, or how it might be determined. Moreover, while it is appealing that the interpretation connects weight directly to a belief, certainly in the context of probabilistic epistemology, it is far from clear how such a conception of the weight can be made to fit with a probabilistic model of a deliberating agent, like the Bayesian model above.

Other options in Genest and McConway (1990) are that the weight expresses a strictly proper score, a measure of accuracy, or a distance to the truth of the agent's opinion (cf. Regan et al., 2006). These proposals provide more of a mathematical handle on the weights, and arguably they allow for an objective determination of weights, and thereby a normative underpinning of them. From the point of view of probabilistic epistemology, however, these interpretations are less attractive. As indicated in the foregoing, the informal interpretation of weight is in terms of relations of trust, or conceived competence, among the agents. Ultimately it would seem that such relations will have to be spelled out in terms of beliefs that the agents have about each other, and not as functions of the probabilities expressing the opinions of the agents about the matter over which they disagree.

Because of their connection to beliefs, interpretations that rely on the Bayesian representation of pooling have an intuitive appeal. In the Bayesian representation, the weight shows up as the coefficient of a linear regression of $I(S)$ on Q , where $I(S)$ is the indicator function for the proposition S . While this proposal connects

the weights to probabilities that express beliefs, the correlation coefficient itself is not readily understood as pertaining to a belief. The shortcoming is therefore, once again, that this interpretation of weights fails to explicate them in terms of opinions that agents have about each other, e.g., as the probability that they are reliable or some such. Correlation coefficients cannot be understood in this way, not in any direct sense at least.

The final proposal in Genest and McConway (1990) is to interpret weights as subjective ratings of trust. This proposal is directly connected to Lehrer and Wagner (1981) and DeGroot (1974), and it is indeed the interpretation favored in the current paper. Genest and McConway correctly note that the interpretation leaves a lot to be desired. It is unclear how these subjectively determined weights can be given a normative or an empirical basis.

It is on these points that the current paper presents a step forward: it offers a new interpretation of weights in terms of opinions, explains how agents express trust through the weights, and facilitates the elicitation of these trust parameters. This particular notion of trust is identified with truth-conduciveness, an identification that can be derived by situating opinion pooling in a Bayesian model for voting known from Condorcet's jury theorem. Put roughly, the weight expresses the ability attributed to a peer of determining the truth or falsity of the proposition under scrutiny.

5 Future research

The Bayesian model of pooling makes possible a particular interpretation of pooling weights, and an associated method for eliciting them. This section mentions some of the most immediate ways of expanding on this result.

As discussed in section 2.2, there are complications with representing the social deliberation of multiple peers in terms of the linear opinion pool. This may

seem to suggest that the interpretation offered in the foregoing is restricted to contexts with two peers. But this is by no means obvious. We might adapt the Bayesian model to represent linear pooling for multiple peers directly, and then reconstruct expressions for the weights. Or else we might attempt to trace the interpretation of weights in pools with multiple agents back to weights for one-on-one exchanges.

Recall that pooling with multiple peers can be converted into consecutive but pairwise pooling operations. For the latter operations, the interpretation of weights proposed above can be maintained. Moreover, the weights for multiple peers can be expressed in terms of weights for pairs. This will go some way towards an interpretation of them. On the other hand, if we want to use this construction to elicit weights for multiple peers, we need to know how to express them exactly, and for this we have to fix the order in which the peers will be accommodated: the order will matter for the calculation of weights. Further research is needed to clarify the conversion to consecutive pairwise pools, and the dependence among peers that is inherent to it.

Apart from expanding on the result by generalisation and refinement, there are natural ways to expand it by applications to other domains. One natural choice for this is traditional epistemology. For a good decade, epistemologists have been interested in disagreement among peers (Christensen and Lackey, 2010). With some notable exceptions (e.g., Russell et al., 2015; EaGIHiVe, 2016; Mulligan, 2019), this debate has relatively few points of contact with the literature in probabilistic epistemology, statistics and epistemic game theory. The current paper might serve a constructive role in facilitating more extensive contact, by offering specific formal counterparts, and thus conceptual explications, for common-sense notions like trust and perceived competence.

By way of example, an influential view on how to resolve disagreement employs the strategy of splitting-the-difference: when two agents have different de-

degrees of belief about some proposition S , then they can resolve their disagreement by both adopting a, possibly weighted, average of the two degrees using a linear pool. Many have wondered what may justify this intuitive response (e.g., Jehle and Fitelson, 2009). The Bayesian reconstruction of pooling may provide such a justification, by stating precisely what beliefs the agents must have for splitting-the-difference to be probabilistically coherent, and by specifying what it means for peers to have equal trust in each other.

There are other bodies of literature that can fruitfully be related to linear pooling through its Bayesian reconstruction. A good example is provided by Romeijn and Roy (2018), who construct a relation between iterative linear pooling (DeGroot, 1974; Lehrer and Wagner, 1981) and Aumann's well-known agreement theorem (Aumann, 1976). That result and the current one invite further research into pooling as a particular form of information sharing. In a world increasingly determined by the wisdom, or foolhardiness, of information-sharing crowds, research into the rationality of such opinion dynamics is of crucial importance.

Acknowledgements

I am indebted to Olivier Roy, with whom I undertook the bigger project of relating consensus formation through iterative pooling to the dynamic version of Aumann's agreement result. Further thanks go to Jean Baccelli, Denis Bonnay, Richard Bradley, Mikael Cozik, Franz Dietrich, Simon Huttegger, Hannes Leitgeb, Julia Staffel, Rush Stewart, Kevin Zollman, and an anonymous referee for illuminating comments and discussions.

References

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

- Bacelli, J. and R. Stewart (2020), Support for Geometric Pooling, manuscript available from the authors.
- Bonnay, D. and M. Cozic (2018), Weighted Averaging, Bayesian update, and Jeffrey Conditioning, *Theory and Decision* 85: 21—39.
- Bradley, R. (2006), Taking Advantage of Difference in Opinion, *Episteme: A Journal of Social Epistemology* 3(3), 141–15.
- Bradley, R. (2007), Reaching a Consensus, *Social Choice and Welfare* 29, 609—632.
- Bradley, R. (2018), Learning from others: conditioning versus averaging, *Theory and Decision* 85, 5—20.
- Brandenburger, A. (2007), The power of paradox, *International Journal of Game Theory* 35, 465–492.
- Christensen, D. and J. Lackey (2010), *The Epistemology of Disagreement*, Oxford University Press.
- Dawid, A. P., DeGroot, M. H. and J. Mortera (1995), Coherent combination of experts' opinions, with comments, *Test* 4, 263—313.
- Dawid, A. P. and J. Mortera (2019), Resolving some contradictions in the theory of linear opinion pools, *Theory and Decision*, <https://doi.org/10.1007/s11238-019-09729-0>.
- DeGroot, M. (1974), Reaching a consensus, *Journal of the American Statistical Association* 69, 118–121.
- DeGroot, M. (1988), A Bayesian view of assessing uncertainty and comparing expert opinion, *Journal of Statistical Planning and Inference* 20, 295–306.

- Dietrich, F. (2010), Bayesian group belief, *Social Choice and Welfare* 35(4), 595–626.
- Dietrich, F. and C. List (2016), Probabilistic opinion pooling, in C. Hitchcock and A. Hajek (eds.), *Oxford Handbook of Probability and Philosophy*, Oxford: Oxford University Press.
- Easwaran, K. , L. Fenton-Glynn, C. Hitchcock and J. Velasco (2016), Updating on the Credences of Others: Disagreement, Agreement, and Synergy, *Philosopher's Imprint* 16(11), 1–39.
- French, John R. Jr. (1956), A formal theory of social power, *Psychological Review* 63(3), 181–194.
- Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, (2020), *Bayesian Data Analysis*, available at <http://www.stat.columbia.edu/gelman/book/>.
- Geanakoplos, J. D. and H. M. Polemarchakis (1982), We Cannot Disagree Forever, *Journal of Economic Theory* 28, 192–200.
- Genest, C. and M. J. Schervish (1985), Modelling Expert Judgments for Bayesian Updating, *Annals of Statistics* 13(3), 1198–1212.
- Genest, C. , K. J. McConway, and M. J. Schervish (1986), Characterization of Externally Bayesian Pooling Operators, *Annals of Statistics* 14(2), 487–501.
- Genest, C. and K. J. McConway (1990), Allocating the Weights in the Linear Opinion Pool, *Journal of Forecasting* 9, 53–73.
- Genest, C. and J. V. Zidek (1986), Combining Probability Distributions: A Critique and Annotated Bibliography, *Statistical Science* 1(1), 113–135.

- Harsanyi, J. (1966-67), Games of incomplete information played by Bayesian players, Parts I-III, *Management Science* 14, 159–182, 320-334, 486–502.
- Halpern, J. Y. (2003), *Reasoning about Uncertainty*, MIT press.
- Hegselmann, R. and U. Krause (2002), Opinion dynamics and bounded confidence: models, analysis and simulation, *Journal of Artificial Societies and Social Simulation* 5, no. 3.
- Jehle, D. and B. Fitelson (2009), What is the ‘Equal Weight View’?, *Episteme*, 6, 280–293.
- Lehrer, K. and C. Wagner (1981), *Rational Consensus in Science and Society: A Philosophical and Mathematical Study*, Springer Verlag.
- Leitgeb, H. (2017), Imaging all the people, *Episteme* 14(4), 463–479.
- Mulligan, T. (2019), The Epistemology of Disagreement: Why Not Bayesianism?, *Episteme*, doi:10.1017/epi.2019.28.
- Pacuit, E. and O. Roy (2014), *Interactive Rationality*, manuscript available from the authors.
- Pettigrew, R. (2019), On the Accuracy of Group Credences, in T. Szabó Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*, Volume 6, Oxford: Oxford University Press.
- Romeijn, J. W. and D. Atkinson (2011), A Condorcet Jury Theorem for Unknown Juror Competence, *Politics, Philosophy, and Economics* 10(3), 237–262.
- Romeijn, J. W. and O. Roy (2018), All agreed: Aumann meets Lehrer and Wagner, *Theory and Decision* 85: 41–60.

- Regan, H.M., M. Colyvan, L. Markovchick-Nicholls (2006), A formal model for consensus and negotiation in environmental management, *Journal of Environmental Management* 80, 167–176.
- Russell, J. S. , J. Hawthorne and L. Buchak (2015), Groupthink, *Philosophical Studies* 172(5), 1287–1309.
- Steele, K. (2012), Testimony as Evidence: More Problems for Linear Pooling, *Journal of Philosophical Logic* 41, 983–999.
- Stone, M. (1961), The linear opinion pool, *Annals of Mathematical Statistics* 32, 1339–1342.
- van Fraassen, B. C. (1989), *Laws and Symmetry*, Clarendon Press, Oxford.