

Forthcoming in *The British Journal for the Philosophy of Science*

PERSISTENT DISAGREEMENT AND POLARIZATION IN A BAYESIAN SETTING

MICHAEL NIELSEN AND RUSH T. STEWART

ABSTRACT. For two ideally rational agents, does learning a finite amount of shared evidence necessitate agreement? No. But does it at least guard against belief polarization, the case in which their opinions get further apart? No. OK, but are rational agents guaranteed to avoid polarization if they have access to an *infinite*, increasing stream of shared evidence? No.

Keywords. Consensus; dilation; disagreement; merging of opinions; polarization; the Bayesian Consensus-or-Polarization Law

1. INTRODUCTION

In politics, group deliberation, and interpersonal relationships, persistent disagreement and belief polarization are often seen as lamentable features of social life. In some cases, polarization can even be dangerous. “Terrorism itself is a product, in part, of group polarization,” as Cass Sunstein points out (2002, p. 187). In certain politicized areas of science—such as vaccination, minimum wage policy, and climate change—disagreement is commonly attributed to the irrationality of one party to the debate.¹ If only we were rational, one might think, all disagreement would be resolved by collecting and sharing evidence. Call this *the optimistic thesis about learning* (TOTAL).

TOTAL. Rational agents who learn the same evidence resolve disagreements.

Something like TOTAL, made suitably precise, seems to underwrite many of our practices in diverse areas like activism, argumentation, inquiry, and mediation.² It may be quite hard to explain our actions in those areas in a non-cynical fashion without TOTAL.³ The thesis also motivates positions in popular philosophical disputes. Conciliatory positions in the peer disagreement debate take it that “the peer’s disagreement gives one evidence that one has made a mistake in interpreting the original evidence, and that such evidence should diminish one’s confidence” on the issue about which there is disagreement (Christensen, 2009, p. 757). If TOTAL were false, why would mere disagreement diminish one’s confidence? Two agents could disagree in the face of shared evidence without either having made a mistake.

But TOTAL *is* false. In fact, even certain weaker theses are untenable. For example, versions of TOTAL that require shared evidence to guard against polarization—the case in which the extent of disagreement increases—cannot be maintained. Here we study persistent disagreement and belief polarization in a general Bayesian framework. Many formal studies of social opinion dynamics focus on reaching consensus (e.g., Blackwell and Dubins, 1962;

Date: June 24, 2018.

¹Not always. Sometimes disagreement is attributed to a profit motive, for example. But good faith and rationality attributions to opponents in such debates are fairly rare.

²An anonymous referee has suggested that proponents of TOTAL be called *totalitarians*.

³So-called “virtue signaling” is an example of a cynical explanation in the case of activism.

DeGroot, 1974; Lehrer and Wagner, 1981; Genest and Zidek, 1986). But persistent disagreement and polarization are interesting phenomena worthy of study in their own right. While modeling these phenomena has been addressed in the literature (e.g., Hegselmann et al., 2002), we want to study them in a Bayesian setting. The point of adopting this framework is that Bayesian learning represents a leading contender for an *ideal* standard of rational belief revision. Even in this setting, both persistent disagreement in general and belief polarization in particular are possible in the face of increasing, *shared* evidence.

We make our case against TOTAL in two parts. In Part 1, we study persistent disagreement and polarization in circumstances in which agents learn some finite amount of shared evidence. This is an important case since it resembles the contexts in which we observe actual belief polarization. Our primary focus in Part 1 is polarization, of which we distinguish two senses. Since polarization (in either of our senses) implies persistent disagreement, cases in which polarization is rationally permissible represent clear failures of TOTAL. The first sense of polarization that we investigate (Section 2) is local in that it involves increases in the extent of disagreement *with respect to a particular event of interest*. We provide some simple characterizations of precisely when polarization in this sense occurs (Theorem 1). From a Bayesian perspective, no irrationality is required. In Section 3, we discuss a close connection between polarization with respect to an event and dilation, a well-studied phenomenon in the theory of imprecise probabilities. In Section 4 we introduce a global notion of polarization that does not depend on a particular event. When agents polarize globally, the *overall* extent to which their probability distributions disagree increases. As with polarization with respect to an event, global polarization is sometimes permissible for Bayesian agents.

In Part 2, we argue against an even weaker version of TOTAL by turning to the general, idealized setting of Bayesian learning in which agents are able to learn an infinite amount of evidence. Our main mathematical contribution there is a result that we call the *Bayesian Consensus-or-Polarization Law*. That result generalizes the classic merging of opinions result due to Blackwell and Dubins (1962) by relaxing a heavy-handed assumption (absolute continuity) in a mild way with interesting consequences. Relative to our weaker assumption, while it is no longer the case that an agent must assign probability 1 to achieving consensus in the limit, she must assign probability 1 to either achieving consensus or polarizing. So not only is polarization consistent with rationality in the sorts of learning scenarios in which it is observed (Part 1), in some cases, *rationality demands assigning positive probability to polarizing* in the limit of inquiry (Part 2).

Part 1. The Finite Case

Both psychological evidence and common life experience attest to the reality of persistent disagreement and belief polarization. In a classic study on polarization, Lord et al. report that, when exposed to the same set of conflicting studies regarding the possible deterrent effects of the death penalty, subjects disagreeing initially strengthened their respective views, coming to disagree even more strongly (1979). Ross and Anderson claim that this behavior stands “in contrast to any normative strategy imaginable for incorporating new evidence relevant to one’s belief” (1982, p. 145).

In order to investigate the rational status of such observed behavior, we begin by looking at learning situations that fairly closely approximate those in which the behavior occurs. In particular, we first consider cases of learning *finitely many* events. We focus exclusively on Bayesian conditionalization—according to which *events* are learned with *certainty*—rather than Jeffrey conditioning or more general rules (Jeffrey, 2004). If Bayesian learning cannot

guard against persistent disagreement and belief polarization, then generalizations cannot either.⁴

2. POLARIZATION

We begin with the phenomenon of belief polarization. Since, as we will show, polarization is possible for rational agents who learn the same finite amount of shared evidence, it follows that persistent disagreement is also possible. Polarization represents a radical failure of TOTAL. While it may be familiar from debates about uniqueness and permissivism that standard varieties of Bayesianism allow for persistent disagreement of some form (e.g., White, 2005), it does not follow from that fact that Bayesianism allows for belief polarization.

Throughout the paper, let Ω be a set of *elementary events* or *possible worlds*. Let \mathcal{F} be a sigma-algebra on Ω , that is, a non-empty collection of subsets of Ω closed under complementation and countable unions. Elements $A \in \mathcal{F}$ are called *events*. To take a simple and common example, Ω may contain six points representing the outcomes of a roll of a die, and \mathcal{F} might contain all of the relevant die-rolling events such as $\{2, 4, 6\}$, the event that the die lands with an even number face up. A *probability measure* P on the *measurable space* (Ω, \mathcal{F}) is a countably additive set function $P : \mathcal{F} \rightarrow [0, 1]$ such that $P(\Omega) = 1$.⁵ We assume the standard ratio definition of conditional probability. For all $A, E \in \mathcal{F}$,

$$P(A|E) := \frac{P(A \cap E)}{P(E)}, \text{ whenever } P(E) > 0.$$

According to Bayesian conditionalization, when an agent learns an event E , she should revise her probabilities by setting her new probabilities equal to her old probabilities conditional on E . Where P^E is the probability measure that the agent adopts after learning E , conditionalization says that

$$P^E(A) = P(A|E),$$

for all $A \in \mathcal{F}$. We call P^E the *posterior* and P the *prior*.

With those few preliminaries out of the way, consider the following simple example.

Example 1. *Suppose two polling experts provide opinions about the outcome of an election between four candidates. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ be the set of candidates, and let $\mathcal{F} = 2^\Omega$. Let P_1, P_2 , given by Table 1, be probability measures on (Ω, \mathcal{F}) representing the opinions of the polling experts.*

TABLE 1. Priors

| | ω_1 | ω_2 | ω_3 | ω_4 |
|-------|------------|------------|------------|------------|
| P_1 | 1/6 | 1/4 | 1/3 | 1/4 |
| P_2 | 1/2 | 1/12 | 1/4 | 1/6 |

⁴One might ask, as an anonymous referee did, whether in contexts of *only* uncertain evidence—so that updates by Jeffrey conditioning don't reduce to standard conditionalization—persistent disagreement and polarization can be avoided. In general, the answer is no. One subtle issue in this context is how to understand shared evidence. See Huttegger (2015) for a study of conditions that secure merging of opinions for Jeffrey conditioning. In particular, see his Theorem 6.1 for a class of cases in which disagreement persists.

⁵We say that P is *countably additive* if, for any countable collection of disjoint events $\{A_i\}_{i \in I}$, $P(\bigcup_{i \in I} A_i) = \sum_{i=1}^{\infty} P(A_i)$. We note that our Theorem 3 makes essential use of countable additivity. For the simple examples considered in this section, however, finite additivity is sufficient.

Suppose that there are two political parties with ω_1, ω_2 in one party, and ω_3, ω_4 in the other. In order to advance to the general election, a candidate must win her party's primary. Suppose that we are interested in the event $A = \{\omega_1, \omega_2\}$ that a candidate from the first party will win the general election. Let $E = \{\omega_1, \omega_3\}$ be the event that candidates 1 and 3 win their respective primaries. Then, $P_1(\cdot|E)$ and $P_2(\cdot|E)$ are given by Table 2.

TABLE 2. Posteriors

| | ω_1 | ω_2 | ω_3 | ω_4 |
|---------|------------|------------|------------|------------|
| P_1^E | 1/3 | 0 | 2/3 | 0 |
| P_2^E | 2/3 | 0 | 1/3 | 0 |

Then, $P_1(A|E) = 1/3 < 5/12 = P_1(A) \leq P_2(A) = 7/12 < 2/3 = P_2(A|E)$. So, updating on E , P_1 and P_2 get further apart with respect to A . That is, the information that candidates 1 and 3 win their primaries pushes the two polling experts further apart with respect to their opinions about whether a candidate from the first party will win the general election. \triangle

The behavior in Example 1 is eminently reasonable. Given polling expert 1's prior, the information $E = \{\omega_1, \omega_3\}$ is tantamount to learning that the weaker candidate from the first party will run against the stronger candidate from the second party in the general election. Indeed, expert 1 considers candidate ω_1 antecedently weaker than *all* other candidates. For expert 2, however, the case is reversed. Not only is ω_1 the stronger candidate in the first party, she is the strongest candidate overall. So while E decreases expert 1's confidence in A , the event that a candidate from the first party wins the general election, it increases expert 2's confidence. Nothing in this example or the next (Example 2) depends on unreasonable or extreme (0-1 valued) prior probability assignments.

Example 1 suggests the following natural definition of polarization.

Definition 1. Let P_1 and P_2 be probability functions on (Ω, \mathcal{F}) , and let $A, E \in \mathcal{F}$. We say that evidence E polarizes P_1 and P_2 with respect to the event A if

$$P_1(A|E) < P_1(A) \leq P_2(A) < P_2(A|E).$$

Polarizing evidence leads disagreeing opinions to strengthen their initial attitude with respect to each other, resulting in even greater disagreement. Note that Definition 1 is given in terms of a particular event with respect to which polarization occurs. We will later (Section 4) distinguish this sense of polarization from a sense that does not depend on a particular event.

Example 1 already shows that polarization is rationally permissible from a Bayesian perspective. To shed more light on the situation, we would now like to provide some simple conditions that characterize when polarization with respect to an event occurs. We begin by introducing some auxiliary probabilistic concepts. First, when $P(A \cap E) = P(A)P(E)$, we call events A and E *stochastically independent* (according to P). Next, consider the function S defined by

$$S_P(A, E) := \frac{P(A \cap E)}{P(A)P(E)}.$$

The covariance of A and E is given by $Cov(A, E) = P(A \cap E) - P(A)P(E)$. S just puts covariance in ratio form. When $S_P(A, E) = 1$, A and E are stochastically independent. When

$S_P(A, E) > 1$, A and E are positively correlated. And A and E are negatively correlated when $S_P(A, E) < 1$. (Seidenfeld and Wasserman adopt the convention that $S_P(A, E) = 1$ when $P(A)P(E) = 0$ (1993, p. 1141).) Pedersen and Wheeler point out (2014, p. 1312, fn. 8) that S has been put to various uses in formal epistemology and philosophy of science, including as a measure of coherence (Shogenji, 1999). Finally, the quantity $P(E|A)/P(E|A^c)$ is called the *likelihood ratio* for data E and hypotheses A and A^c . Likelihood ratios are used throughout statistics and express the impact of the evidence in terms of how much it favors one hypothesis to another.

Together, S and the likelihood ratio allow us to state two very simple characterizations of polarization with respect to an event.

Theorem 1. *Suppose that $0 < P_1(A) \leq P_2(A) < 1$. Then the following are equivalent.*

- (i) *Evidence E polarizes P_1 and P_2 with respect to A ;*
- (ii) $S_{P_1}(A, E) < 1 < S_{P_2}(A, E)$;
- (iii) $\frac{P_1(E|A)}{P_1(E|A^c)} < 1 < \frac{P_2(E|A)}{P_2(E|A^c)}$.

The proof of this result uses only the probability axioms and algebra. We omit it, assured the reader can furnish it herself should she so desire.

Condition (iii) of Theorem 1 is mentioned in psychological literature directly on polarization (e.g., Jern et al., 2014). Jern et al. analyze previous empirical studies of belief polarization and offer a Bayesian rationalization of some such behavior. They also note that such likelihood ratios as included in condition (iii) determine “the direction in which [an agent’s] beliefs will change.” Condition (ii) has been exploited in another, related literature that we discuss below (e.g., Seidenfeld and Wasserman, 1993; Wasserman and Seidenfeld, 1994; Pedersen and Wheeler, 2014). Theorem 1, while neither particularly deep nor surprising, is more general than it may seem at first. That’s because conditionalizing on any finite string of evidence, E_1, \dots, E_n , can be reduced to learning just a single event, namely, $E = \bigcap_{i=1}^n E_i$. So Theorem 1 identifies necessary and sufficient conditions for any finite string of evidence to polarize P_1 and P_2 with respect to A .⁶ If (Ω, \mathcal{F}) is a sufficiently complex space relative to how much the agents learn—say there is a countable infinity of events, for example, and agents only learn a finite amount of evidence—the agents may be polarized by the event representing their total evidence.

We offer another, more extreme example of polarization adapted from (Herron et al., 1994; Pedersen and Wheeler, 2015).

Example 2. *Consider P_1 and P_2 such that $P_1(G) = 0.1$ and $P_2(G) = 0.9$. Consider the toss of a coin that is fair according to both P_1 and P_2 : $P_1(H) = P_2(H) = 1/2 = P_1(H^c) = P_2(H^c)$. Suppose that the outcomes of the coin toss are independent of the event G according to both P_1 and P_2 . For example, $P_1(G \cap H) = P_1(G)P_1(H)$ and $P_2(G \cap H) = P_2(G)P_2(H)$. Let A be the “matching” event that either both G and H occur or both do not. That is, $A := (G \cap H) \cup (G^c \cap H^c)$. Notice that $P_1(A) = 1/2 = P_2(A)$. Despite initial agreement concerning A , the coin toss polarizes $P_1(A)$ and $P_2(A)$. For $i = 1, 2$,*

⁶We could distinguish the case in which $\bigcap_{i=1}^n E_i$ polarizes P_1 and P_2 with respect A from the case in which each E_i polarizes P_1 and P_2 with respect to A . In the former case, we are concerned with the cumulative effect of the evidence. The latter case obtains when the characterizing conditions of Theorem 1 hold for each piece of evidence, E_i .

$$P_i(A|H) = \frac{P_i([(G \cap H) \cup (G^c \cap H^c)] \cap H)}{P_i(H)} = \frac{P_i(G \cap H)}{P_i(H)} = \frac{P_i(G)P_i(H)}{P_i(H)} = P_i(G).$$

So even though both P_1 and P_2 assign probability $1/2$ to A initially, learning that the coin lands heads yields $P_1(A|H) = 0.1$ and $P_2(A|H) = 0.9$. Hence, $P_1(A|H) < P_1(A) \leq P_2(A) < P_2(A|H)$. \triangle

Example 2 is a striking case. In a single step, learning the same evidence can transform agreement into significant disagreement. Moreover, the example points to an interesting and important connection between polarization, a topic in social epistemology, and dilation, a topic in the theory of individual imprecise probabilities. We turn to this connection now.

3. DILATION

The polarization phenomenon with which we were concerned in Section 2 bears resemblance to the phenomenon known as *dilation* in the theory of imprecise probabilities (IP). In fact, Example 2 is a borrowed example of dilation. In our opinion, there is a very fruitful exchange of ideas between social epistemology and the theory of imprecise probabilities (see, e.g., Levi, 1982, 1985a; Seidenfeld et al., 1989, 2010; Elkin and Wheeler, 2018; Stewart and Ojea Quintana, 2018).

IP allows for more general representations of uncertainty than standard Bayesian probability theory does. Several frameworks have been used to accomplish this task. We will consider *sets of probabilities*, a very general IP framework. Let \mathbb{P} denote a set of probability measures defined on the same measurable space (Ω, \mathcal{F}) . Levi marks an important distinction between two possible interpretations of \mathbb{P} (1985b). On the one hand, we might retain the standard Bayesian ideal according to which any rational state of uncertainty admits representation in terms of a single probability function. On this account, a set \mathbb{P} could be used to represent the possible values that an agent's precise probability judgments may take. Such a set might arise from failures of introspection or partial elicitation. In short, \mathbb{P} may represent imprecision in measuring a precise credal state. For example, in estimating the probability of some event, an agent may be in a position to specify no more than one or two decimal places. On the other hand, IP can be seen as offering an alternative (laxer) normative standard. On this account, a set \mathbb{P} might *completely describe* an agent's probability judgments in cases in which her uncertainty is not reducible to a unique probability function. Levi calls this interpretation *indeterminate* probability. Of course, both factors may operate simultaneously. There could be partial elicitation of an indeterminate state of uncertainty.

Dilation occurs, roughly speaking, when learning *increases* uncertainty. There are various formulations of dilation, depending on the choice of IP representation and whether certain inequalities are strict or not. But to facilitate comparison with the notion of polarization defined in the previous section, consider the following common definition.

Definition 2. Let \mathbb{P} be a set of probabilities on (Ω, \mathcal{F}) , let \mathcal{B} be a positive measurable partition of Ω ⁷, and let $A \in \mathcal{F}$. We say that the partition \mathcal{B} *dilates* A just in case, for each $E \in \mathcal{B}$,

$$\inf\{P(A|E) : P \in \mathbb{P}\} < \inf\{P(A) : P \in \mathbb{P}\} \leq \sup\{P(A) : P \in \mathbb{P}\} < \sup\{P(A|E) : P \in \mathbb{P}\}.$$

⁷The partition \mathcal{B} is positive and measurable if $E \in \mathcal{B}$ implies $E \in \mathcal{F}$ and $P(E) > 0$.

The mathematics of dilation has been extensively studied in a series of articles (Seidenfeld and Wasserman, 1993; Herron et al., 1994; Wasserman and Seidenfeld, 1994; Herron et al., 1997; Pedersen and Wheeler, 2014, 2015). For results related to our Theorem 1, see in particular Wasserman and Seidenfeld’s Result 1 (1994), and Pedersen and Wheeler’s Theorem 1 and its corollary (2015).

Dilation is like a virulent form of polarization. In the social setting, it amounts to agents’ views getting further from consensus no matter what outcome of a partition \mathcal{B} they observe. In Example 2, where the partition is given by the outcomes of a toss of a fair coin ($\mathcal{B} = \{H, H^c\}$), P_1 and P_2 will move the same significant distance from agreement on A whether the coin lands heads or tails. It is the fact that a more precise estimate for an event A is transformed into a less precise estimate *regardless of the event in the partition that is learned* that attracts interest to dilation. While dilation emerges sometimes for IP, precise credal states are immune to dilation since $\inf\{P(A|E)\} = \sup\{P(A|E)\}$ for all A and E in \mathcal{F} .

Some see dilation as a pathological feature that calls IP into question (e.g., White, 2010). While not central to the present study, the debate about the normative status of dilation, we suggest, might be further illuminated by consideration of dilation in a social setting. Similarly, another line of inquiry into polarization would be to explore whether arguments against the rational acceptability of dilation can be extended in some way to social settings. Does the (un)acceptability of dilation provide some defeasible consideration in favor of the (un)acceptability of (certain kinds of) polarization or *vice versa*? Or is there, contrary to the hopes of those seeking decision theories and theories of inquiry unified across individuals and groups, some significant disanalogy?

4. GLOBAL POLARIZATION

In the previous sections, our focus was a local sense of polarization. We were interested in situations in which sharing evidence increases the extent of disagreement between two Bayesian agents *with respect to a fixed event*. In some situations, however, there may not be a distinguished event around which inquiry centers, or one may be interested in whether a consensus can be reached about a whole collection of events. In such situations, a more global perspective is appropriate. In climate science, for example, there is a complex cluster of issues—from future sea levels and air temperature to the mechanisms underwriting decadal climate variability—that exercise researchers in the area. In some cases, consensus on appropriate measures to mitigate change or on adaptive policies may require consensus on a large number of other issues.

In order to adopt the global perspective, we need a way of measuring the extent to which two probabilities disagree that does not depend on a distinguished event. We will use the *total variational distance* d defined for any probabilities P_1 and P_2 by

$$d(P_1, P_2) = \sup_{A \in \mathcal{F}} |P_1(A) - P_2(A)|.$$

Note that if P_1 and P_2 are in complete agreement, in the sense that $P_1(A) = P_2(A)$ for all events A , then the total variational distance between them is 0. If, on the other hand, P_1 and P_2 disagree *maximally*, in the sense that there’s an event A such that $P_1(A) = 0$ and $P_2(A) = 1$, then $d(P_1, P_2) = 1$. The total variational distance finds use throughout probability theory and, as we explain in the next sections, has played a crucial role in Bayesian thought *via* the Blackwell-Dubins merging of opinions theorem (1962) and related results (Schervish and Seidenfeld, 1990; Huttegger, 2015). In the examples below, we will make use of the fact

that in finite probability spaces the total variational distance is given by

$$d(P_1, P_2) = P_1(A_0) - P_2(A_0), \quad (1)$$

where A_0 is the set of points $\omega \in \Omega$ such that $P_1(\omega) > P_2(\omega)$.⁸ Using total variational distance, we can now introduce a notion of polarization that is the global analogue of polarization with respect to an event, as defined in Definition 1.

Definition 3. We say that evidence E polarizes P_1 and P_2 globally if $d(P_1, P_2) < d(P_1^E, P_2^E)$.

Note that the notion of polarization in Definition 3 does not depend on a particular event, but rather is concerned with the effect that learning has on the *overall* disagreement between two probabilities.

Does global polarization imply irrationality? Just as with polarization with respect to an event, the answer is no when the standard of rationality is Bayesian. This can be seen by considering a slight variation of Example 1.

Example 3. Let (Ω, \mathcal{F}) , $A = \{\omega_1, \omega_2\}$, and $E = \{\omega_1, \omega_3\}$ be defined as in Example 1, and consider the following priors and posteriors.

TABLE 3

| | ω_1 | ω_2 | ω_3 | ω_4 |
|---------|------------|------------|------------|------------|
| P_1 | 1/4 | 1/8 | 1/2 | 1/8 |
| P_2 | 1/2 | 1/12 | 1/4 | 1/6 |
| P_1^E | 1/3 | 0 | 2/3 | 0 |
| P_2^E | 2/3 | 0 | 1/3 | 0 |

Using (1), we can see that we have global polarization because $d(P_1, P_2) = 7/24 < 1/3 = d(P_1^E, P_2^E)$. Note also that we still have polarization with respect to A , like in Example 1, because $P_1(A | E) = 1/3 < 3/8 = P_1(A) \leq P_2(A) = 7/12 < 2/3 = P_2(A | E)$. \triangle

Absent reason to believe that the above probability assignments are unreasonable, Example 3 shows that TOTAL is false under a global interpretation of “disagreements.” In fact, as was the case in Section 2, the present example falsifies a thesis that is even weaker than TOTAL because global polarization, like polarization with respect to an event, implies persistent disagreement.⁹

⁸*Proof.* Note that A_0^c is the set of points $\omega \in \Omega$ such that $P_2(\omega) \geq P_1(\omega)$. For all $A \in \mathcal{F}$ we have

$$\begin{aligned} P_1(A) - P_2(A) &= \sum_{\omega \in A} [P_1(\omega) - P_2(\omega)] = \sum_{\omega \in A \cap A_0} [P_1(\omega) - P_2(\omega)] + \sum_{\omega \in A \cap A_0^c} [P_1(\omega) - P_2(\omega)] \\ &\leq \sum_{\omega \in A \cap A_0} [P_1(\omega) - P_2(\omega)] \leq \sum_{\omega \in A_0} [P_1(\omega) - P_2(\omega)] = P_1(A_0) - P_2(A_0). \end{aligned}$$

Similarly, $P_2(A) - P_1(A) \leq P_2(A_0^c) - P_1(A_0^c)$ for all $A \in \mathcal{F}$. Since $P_1(A_0) + P_1(A_0^c) = 1 = P_2(A_0) + P_2(A_0^c)$, we have $P_1(A_0) - P_2(A_0) = P_2(A_0^c) - P_1(A_0^c)$, and it follows from the above that $|P_1(A) - P_2(A)| \leq P_1(A_0) - P_2(A_0)$ for all $A \in \mathcal{F}$. The equality in (1) now follows by taking the supremum (maximum) of the left-hand side of the last inequality. \square

⁹If $d(P_1, P_2) < d(P_1^E, P_2^E)$, then $d(P_1^E, P_2^E) > 0$. This implies that there is some event A about which P_1^E and P_2^E disagree, i.e. $P_1^E(A) \neq P_2^E(A)$.

Having defined two notions of polarization, it is natural to ask whether there are any interesting logical relations between them. Example 3 shows that both kinds of polarization can occur simultaneously. But does one notion imply the other? To begin to address this question, we return to Example 1. In that example, although E polarizes P_1 and P_2 with respect to A , it is not the case that E polarizes P_1 and P_2 globally. One can see this by using (1) to calculate $d(P_1, P_2) = 1/3 = d(P_1^E, P_2^E)$. By altering the probabilities in Example 1 a bit, one can find even more striking cases, in which there is polarization with respect to an event even though conditionalizing *decreases* the total variational distance between posteriors. For completeness, we have included such an example in the Appendix (Example 4).

So we cannot infer global polarization from polarization with respect to an event. How about the converse implication? If evidence E polarizes two probabilities globally, does it follow that E polarizes the two probabilities with respect to some event? Again the answer is negative by another slight modification of the previous examples. See Example 5 in the Appendix. We summarize the previous conclusions with the following proposition.

Proposition 1. *There are cases in which E polarizes P_1 and P_2 both globally and with respect to some event A . However, the following implications do **not** hold.*

- (i) *If E polarizes P_1 and P_2 with respect to some event A , then E polarizes P_1 and P_2 globally.*
- (ii) *If E polarizes P_1 and P_2 globally, then E polarizes P_1 and P_2 with respect to some event A .*

Having established that global polarization and polarization with respect to an event are logically independent, it would be convenient to state a simple characterization of global polarization along the lines of Theorem 1. Although there are various methods for computing and comparing the total variational distances between posteriors and priors, we have not discovered a method that is sufficiently simple and illuminating for the purposes of this paper. Rather than introduce more technical material for relatively little philosophical payoff, we prefer to leave the task of finding a simple characterization of global polarization as an open problem. In future work, we plan to investigate global polarization in the context of imprecise probabilities. Global polarization gives rise to a phenomenon that is similar to dilation in some ways, but, like local and global polarization for precise probabilities, this phenomenon is logically independent of dilation for imprecise probabilities.

The important point for our purposes is that global polarization does not imply irrationality. We have now shown two senses in which TOTAL fails when the standard of rationality is Bayesian. Bayesian agents who learn a finite amount of shared evidence can exhibit both local polarization with respect to a distinguished event and global polarization with respect to their entire probability distributions. Whether we interpret TOTAL as requesting local or global resolutions of disagreement, we find that the thesis is false.

Yet, faith in the ability of rationality and evidence to avoid polarization may remain. Sure, learning just *one* event (or finitely many for that matter) allows for polarization. But doesn't ongoing inquiry that allows for as many observations as we please avoid it? We show in Section 6 that, in fact, it does not. Taking inquiry to the limit does not save TOTAL.¹⁰ But

¹⁰In a sense, this is already clear even in the finite case. While Example 2 does not depend on extreme assignments, it could be adapted so that $P_1(G) = 0$ and $P_2(G) = 1$. In this case, evidence H maximally "polarizes" P_1 and P_2 with respect to A . And since 0 and 1 probabilities are not revisable under Bayesian conditionalization, such polarization is permanent. So there is no hope of undoing polarization or resolving disagreements concerning A for P_1 and P_2 . Note that A is still *not* an event for which either prior is extreme

it is worth pointing out that retreating to the limit of inquiry already drastically weakens any automatic inference from the mere fact of actual polarization to the irrationality of some polarized agent or other. Presumably, all actual behavior occurs in the context of just finitely many observations. In such a context, this line of response concedes that polarization is consistent with the rationality of both parties.

Part 2. The General Case

In the general setting, we consider cases of learning infinite amounts of evidence. Why bother looking at such artificial learning scenarios? In a way, we are pursuing TOTAL to its last retreat. Even allowing rational agents to learn an infinite amount of evidence, agreement cannot be ensured. Moreover, such learning scenarios are frequently considered in the context of Bayesian foundations. It is sometimes thought that we can test the mettle of a learning method by looking at its asymptotic behavior. Bengt Autzen gives recent voice to this idea. “Under the ideal scenario of an infinitely large data set,” he writes, “an inference procedure should show certain desirable features” (Autzen, 2017, p. 3). One desirable feature claimed for Bayesian methodology is known as *convergence to the truth*. Convergence has long been taken to be a metric by which to judge probabilistic inference methods (e.g., Reichenbach, 1938, §43). Our concern in this paper, though, is with consensus. Consensus is no less pedigreed a methodological concern than convergence is. In “The Fixation of Belief,” Peirce puts forward a picture of scientific method according to which a community of inquirers achieve consensus eventually by updating on shared evidence. As he puts it, “the method must be such that the ultimate conclusion of every man shall be the same” (Peirce, 1992a, p. 120).¹¹ Here, too, many Bayesians claim success. In the next section, we explain the basis for this claim. In short, then, we examine the general case because proponents of TOTAL may be tempted to appeal to such idealized scenarios and because such scenarios are standardly studied in Bayesian theory and even play foundational roles in justifications for the Bayesian point of view.

Before we can state the relevant facts about convergence and consensus for Bayesians, we need to bring a bit more machinery online. This machinery will help us in this second part of our case against TOTAL. As above, let (Ω, \mathcal{F}, P) be a probability space. We will be interested in situations in which agents anticipate learning an increasing amount of evidence that eventually settles every event of interest to them. The evidence is represented as a sequence of finite partitions, $\{\mathcal{E}_n\}_{n \in \mathbb{N}}$, such that \mathcal{E}_{n+1} refines \mathcal{E}_n for all $n \in \mathbb{N}$.¹² Because of the assumption that the partitions are increasingly fine, we say that the agent’s evidence is *increasing*. For example, a partition might represent the possible outcomes of an experiment that the agent plans to perform. In the case of repeated coin tosses, \mathcal{E}_1 represents the information about the first toss of the coin, while \mathcal{E}_2 would represent all of the information about the first two tosses, etc. So by the time the agent observes the outcome of the second coin toss, she knows whether the “actual” sequence is one that begins *HH* or not.

in this modified example. So the point is not merely that achieving consensus is frustrated for those events with prior 0-1 assignments. In any case, disagreeing on prior 0-1 probabilities does not count as polarization as we define it, even if it is a case of interminable disagreement for Bayesians.

¹¹Otherwise, a method will fail to fix belief because we will encounter those who disagree with us and, due to our “social impulse,” our confidence will be shaken. In other places, Peirce seems to identify truth and whatever opinion the community settles on in the limit (e.g., 1992b).

¹²The finite partition assumption aids exposition but is not necessary. In the Appendix, we relax it and work with general filtrations of sub-sigma-algebras.

We will also assume that the observations eventually settle every event in \mathcal{F} and say that the evidence is *complete*. Formally, we require that the collection of all evidential events, namely $\bigcup_n \mathcal{E}_n$, generate the sigma-algebra \mathcal{F} on which the agent’s prior is defined. That is, we will assume that \mathcal{F} is the smallest sigma-algebra containing $\bigcup_n \mathcal{E}_n$.

From her prior perspective, the agent is uncertain, for all n , which event in \mathcal{E}_n she will learn. If $\omega \in \Omega$ is the actual world, then $E_n(\omega)$ denotes the event in \mathcal{E}_n that the agent learns at stage n , namely, whichever member of \mathcal{E}_n contains ω . In this setup, a Bayesian agent’s posteriors are then $P(\cdot | E_n(\omega))$ for all $\omega \in \Omega$ and $n \in \mathbb{N}$.¹³ The aforementioned *convergence to the truth* theorem says that a Bayesian agent assigns probability 1 to the event that her posterior probabilities converge to the truth about every event in \mathcal{F} . This means that for every $A \in \mathcal{F}$ and every $\omega \in \Omega$ in a set with P -probability 1, if A is true, so that $\omega \in A$, a Bayesian’s posteriors $P(A | E_n(\omega))$ will get arbitrarily close to 1 as n increases. If A is false, so that $\omega \notin A$, then $P(A | E_n(\omega))$ will get arbitrarily close to 0 as n increases. See the [Appendix](#) for a more formal summary. We return now to our primary focus, consensus or the lack thereof.

5. MERGING OF OPINIONS

Merging of opinions is an important part of Bayesian lore. Relative to just a few assumptions, with probability 1, opinions get closer and closer together as they learn from a shared, increasing stream of data. Huttegger sees merging results as evidence that Bayesianism fulfills Peirce’s vision of a method that settles belief for a community on the basis of experimental evidence. He writes, “experience trumps any initial belief state; diverging opinions are just a sign that not enough evidence has accumulated yet” (2015, p. 613). Savage (1954), Blackwell and Dubins (1962), and Gaifman and Snir (1982) provide classic versions of merging of opinions theorems, which have since been generalized in various ways (Schervish and Seidenfeld, 1990; Huttegger, 2015; Stewart and Nielsen, 2018). Such classic versions of these results attained their prominent theoretical status due to the subjective nature of personal probabilities. According to many Bayesians, that agents reach consensus (almost surely, relative to the assumptions in the theorems) should allay concerns that Bayesianism robs science of any sort of objectivity. “This approximate merging of initially divergent opinions is, we think, one reason why empirical research is called ‘objective’,” write Edwards, Lindman, and Savage (1963, p. 197).

Let P and Q be two probability measures on (Ω, \mathcal{F}) . Above, we explained that if P anticipates learning increasing and complete evidence, then P assigns probability 1 to converging to the truth. In this section, we continue to assume that evidence is increasing and complete, but we now also assume that it is *shared* with Q . We say that P *shares evidence with* Q if for all $n \in \mathbb{N}$ and $E \in \mathcal{E}_n$, if $P(E) > 0$, then $Q(E) > 0$. This ensures that Q can conditionalize on any evidential event that P can—anything that P can learn, Q can learn, too.

For all $\omega \in \Omega$, let $P_n(\omega) = P(\cdot | E_n(\omega))$ and $Q_n(\omega) = Q(\cdot | E_n(\omega))$ be the posteriors for P and Q after conditionalizing on a member of the n^{th} partition. As an informal gloss on the merging of opinions results, we might say, if P is “sufficiently similar” to Q , then for all ω in a set with P -probability 1 the distance between the posteriors $P_n(\omega)$ and $Q_n(\omega)$ goes to 0 as n increases. We will use the same notion of distance that we used in Section 4, namely total variational distance. We say that P and Q *merge* if $d(P_n(\omega), Q_n(\omega))$ gets arbitrarily close to

¹³Technically, we may have $P(E_n(\omega)) = 0$ for some n and ω , in which case we can define $P(\cdot | E_n(\omega))$ arbitrarily and replace “for all $\omega \in \Omega$ ” with “for all $\omega \in \Omega$ in a set with P -probability 1.”

0 for all ω as n increases, and we say that P *expects to merge with* Q if this event occurs for all ω in a set with P -probability 1.

We now need to say what we mean by sufficient similarity. Call P *absolutely continuous* with respect to Q when $Q(A) = 0$ implies $P(A) = 0$ for all $A \in \mathcal{F}$. In other words, any extreme probability assignment of Q 's is an extreme probability assignment of P 's. It could still be the case, though, that $P(A) = 0$ but $Q(A) > 0$ for some $A \in \mathcal{F}$. If Q is absolutely continuous with respect to P also, then we say that P and Q are *mutually absolutely continuous*.

Theorem 2 (Blackwell and Dubins (1962)). *Suppose that P shares increasing and complete evidence with Q . If P is absolutely continuous with respect to Q , then P expects to merge with Q .*

Given that P is absolutely continuous with respect to Q , P assigns probability 1 to approaching consensus with Q when they share increasing and complete evidence. If, in addition, Q is absolutely continuous with respect to P , then both P and Q expect to merge with each other.

But wait. Doesn't Theorem 2 show that polarization is essentially inconsistent with Bayesian rationality? Some authors do indeed seem to think polarization is beyond the pale, not just for some vague theory of rationality in general, but for Bayesianism in particular.

Ample psychological evidence suggests that people's learning behavior is often prone to a "myside bias" or "irrational belief persistence" in contrast to learning behavior exclusively based on objective data. In the context of Bayesian learning such a bias may result in diverging posterior beliefs and attitude polarization even if agents receive identical information. Such patterns cannot be explained by the standard model of rational Bayesian learning that implies convergent beliefs. (Zimper and Ludwig, 2009, p. 181)

At least two points about Theorem 2 require careful consideration. First, the theorem has preconditions. To see that they are significant, note that Theorem 2 has a partial converse, which follows from the Bayesian Consensus-or-Polarization Law in the next part of the paper. It turns out that if P is *not* absolutely continuous with respect to Q , then either P does not expect to merge with Q or P does not share evidence with Q . So, provided P shares evidence with Q , if P is *not* absolutely continuous with respect to Q , then P assigns positive probability to the event that the posteriors P_n and Q_n persist in disagreeing despite access to an increasing and complete evidence. In section 6, we return to these issues with more precision. But we should pause now to think about the absolute continuity assumption.

Absolute continuity is not a *rationality* requirement. After all, with respect to which measure or measures ought a prior be absolutely continuous? Unless we have a principled answer to that question, it is difficult to even make sense of a proposal according to which absolute continuity is a normative constraint. And there is no trivial answer to the question in general. In large measurable spaces¹⁴, there is no measure with respect to which all measures are absolutely continuous. Of course, even if P and Q were both absolutely continuous with respect to a third, distinguished measure, neither need be absolutely continuous with respect to the other. It might be tempting to urge priors to avoid extreme assignments. For instance, *regular* probability measures—which assign positive probability to all non-empty events—enjoy wide-spread support among philosophical probabilists (Shimony, 1955; Lewis, 1980;

¹⁴Such spaces arise when considering random variable with continuous distributions, for example. A random variable representing the unknown duration of a prizefight would be such a quantity.

Skyrms, 1995). However, in large enough probability spaces, such measures are impossible and extreme prior probability assignments unavoidable.¹⁵

Similarly, absolute continuity is of dubious *descriptive* value. According to Miller and Sanchirico, the condition is even “difficult to interpret behaviorally” (1999, p. 171). But our concern is that absolute continuity just *assumes* a great deal of agreement out of the gate. When is that much initial agreement actually realized? Perhaps one could maintain that communities of scientists often endorse statistical models sufficiently similar so as to be absolutely continuous. As a descriptive claim, however, such a view needs empirical support. Earman considers something even stronger.¹⁶ Mutual absolute continuity may be in some sense *constitutive* of a scientific community: “it could be held that decisions on zero priors help to define scientific communities and that an account of scientific inference must be relativized to a community” (1992, p. 142). In a debate about the descriptive adequacy of absolute continuity, such a response would verge on question-begging. If the very definition of a scientific community implies absolute continuity of its members’ opinions, then no serious debate about the descriptive status of absolute continuity in research communities remains to be had.

Another suggestion is that an important *conceptual* distinction about disagreement can be explicated in terms of absolute continuity.¹⁷ The distinction is between disagreement and *radical* disagreement, with radical disagreement being modeled by failures of absolute continuity. First, we note that radical disagreement on this explication is not symmetric, which may be a questionable feature. One prior may be in radical disagreement with another that is only in modest disagreement with it. Second, failures of absolute continuity can arise even when opinions are, in one sense, intuitively “close,” just as absolutely continuous priors can be “far apart.” If $Q(A) = 0$ but $P(A) = 0.000001$, for example, then P is not absolutely continuous with respect to Q ; while $P(A) = 0.99$ and $Q(A) = 0.01$ does not preclude absolute continuity. Third, in order to avoid being a mere relabeling of when absolute continuity holds or not, the distinction should draw on some well-motivated, independent account of radical disagreement which may not be forthcoming in light of the sort of example just given.

The second issue about Theorem 2 requiring special attention appears in the consequent rather than the antecedent. P merges with Q *almost surely*. The distinction between *sure* and *almost sure* is not always kept firmly in mind when it comes to Bayesian convergence and merging results, leading some to make remarks apparently inconsistent with our general claim in this paper. More cynically, Earman writes, “‘almost surely’ sometimes serves as a rug under which some unpleasant facts are swept” (Earman, 1992, p. 148).¹⁸ We may be interested in what holds surely (for all $\omega \in \Omega$), and not in what is merely highly probable.¹⁹

¹⁵“Perhaps the reason that the absolute continuity assumption has gained such currency in the literature is that it is so plausible in a finite, or even countable setting. Even the stronger assumption that both players regard each state as at least possible [they have regularity in mind here] seems attractive, since all this rules out is dogmatism. But it would be a mistake to carry this intuition into the necessarily uncountable setting that is relevant here: obviously, in this case some events must receive zero measure” (Miller and Sanchirico, 1999, p. 179).

¹⁶Earman himself regards the fanfare concerning merging of opinions with a good deal of skepticism.

¹⁷Thanks to an anonymous referee for this suggestion.

¹⁸Glymour (1980) raises a similar concern about convergence to the truth, writing, “The theorem does not tell us that in the limit any rational Bayesian will assign probability 1 to the true hypothesis and probability 0 to the rest; it only tells us that rational Bayesians are certain that he will” (p. 73). More recently, Gordon Belot has argued that convergence to the truth results constitute a liability for Bayesians because they forbid a “reasonable epistemological modesty” (2013, p. 502).

¹⁹See Nielsen (2018).

Attending to the distinction between *merging*, on the one hand, and *merging almost surely*, on the other, we see that, if the “actual world” ω is outside the support of P , for all Theorem 2 says, P and Q may not *actually* merge. It is just that P assigns such points 0 probability. A number of examples have been used to motivate distinguishing certainty from 0-1 probability assignments. An agent might regard each toss in an infinite sequence of tosses of a fair coin as independent. Then, any infinite sequence of (outcomes of) coin tosses bears probability 0. Yet it would be a mistake to infer that she is certain that no such sequence is the actual one. An infinitely fine dart is thrown at the $[0, 1]$ interval. An agent’s opinions about the outcome of the throw may be representable by the Lebesgue measure. Then, each real number in the unit interval bears probability 0. But the agent is not certain that, for each real number, the dart will not hit it.

Both the assumption of absolute continuity and the almost surely hedge require us to exercise considerable care in interpreting Blackwell and Dubins’s theorem. In the remainder of the paper, we focus on absolute continuity. Even restricting our attention to probability “strong laws” (results that hold with probability 1), relaxing absolute continuity even to a very modest extent has significant ramifications.

6. THE BAYESIAN CONSENSUS-OR-POLARIZATION LAW

The asymptotic analogue of polarization occurs when the total variational distance between posteriors tends to the maximal value of 1. More precisely,

Definition 4. P and Q *polarize in the limit* if $d(P_n(\omega), Q_n(\omega))$ gets arbitrarily close to 1 for all ω as n increases.

Polarization in the limit is basically stronger than the notion of global polarization in Definition 3. To explain this relation, suppose that P and Q polarize in the limit, and let r be the total variational distance between P and Q . It cannot be that $r = 0$ because then P and Q would be identical and would not be able to polarize in the limit. On the other hand, if $r = 1$, then polarization in the sense of Definition 3 can not occur. But this is a rather trivial limiting case. In cases of interest we may therefore assume that $0 < r < 1$. Since P and Q polarize in the limit, for all ω there is some stage n such that $d(P_n(\omega), Q_n(\omega)) > r$. It follows that $E_n(\omega)$ polarizes P and Q globally. Hence, excluding the trivial limiting case in which $r = 1$, if P and Q polarize in the limit, then for all ω there is some stage n such that $E_n(\omega)$ polarizes P and Q globally. It is in this sense that polarization in the limit “basically” implies global polarization. Even if our notion of polarization in the limit is not the unique extension of Definition 3 to the general setting, it is a natural one in light of the relation just explained. Another reason to focus on it is that the total variational distance plays a prominent role in the theory of Bayesian merging, as evidenced by the common retort that “priors wash out.”

In this section, we state and discuss a generalization of the Blackwell-Dubins merging of opinions result that we call the *Bayesian Consensus-or-Polarization Law*. To do this, we make use of a deep but easily explained result in measure theory called the Lebesgue Decomposition Theorem. In order to explain that result, we need one more definition. Let P and Q be any two probability measures. We say that P and Q are *mutually singular* if P assigns probability 1 to an event to which Q assigns probability 0.²⁰ When P and Q are mutually singular, absolute continuity fails in the most radical way possible. In subjective

²⁰Mutually singular probabilities have already made a brief appearance in our study of polarization. See footnote 10.

terms, P is probabilistically certain that some event A will occur while Q is certain that it will not. It follows that when P and Q are mutually singular, the total variational distance between them takes its maximal value of 1.

Now, for any two probability measures P and Q , the Lebesgue Decomposition Theorem implies that for some $\delta \in [0, 1]$, P can be decomposed as

$$P = \delta P^a + (1 - \delta)P^s, \quad (2)$$

where P^a is a probability measure that is absolutely continuous with respect to Q and P^s is a probability measure that is mutually singular with Q . Furthermore, if δ is strictly between 0 and 1, then the decomposition is unique. One can check that P is absolutely continuous with respect to Q if and only if $\delta = 1$. And similarly, P and Q are mutually singular if and only if $\delta = 0$.

The decomposition given by equation (2) tells us that P can be viewed as a mixture of two probabilities, one of which is absolutely continuous with respect to Q and one of which is mutually singular with Q . The larger δ is the “more” absolutely continuous P is with respect to Q . For this reason, we will call δ the *degree of absolute continuity* of P with respect to Q , and we say that P is absolutely continuous with respect to Q to degree δ . In view of our earlier remarks, P is absolutely continuous with respect to Q *simpliciter* just in case P is absolutely continuous with respect to Q to degree 1.

The main result of this part of the paper demonstrates a tight connection between degree of absolute continuity and merging of opinions and generalizes the Blackwell-Dubins theorem that we discussed in the previous section.

Theorem 3 (Bayesian Consensus-or-Polarization Law). *Let P and Q be any two probabilities and suppose that P is absolutely continuous with respect to Q to degree δ . Let M be the event that P and Q merge, and let L be the event that P and Q polarize in the limit. If P shares an increasing and complete sequence of evidence with Q , then*

$$P(M) = \delta \quad \text{and} \quad P(L) = 1 - \delta.$$

Theorem 3 implies that an agent with probabilities given by P is certain, with probability 1, that either he and Q will merge or they will polarize: $P(M \cup L) = 1$. Under the assumption of shared, increasing, and complete evidence, it is incoherent for P to assign positive probability to the event that the distance between the posteriors P_n and Q_n tends to 0.6, for example, or that the distance oscillates forever. To see that our result generalizes the Blackwell-Dubins theorem, suppose that P is absolutely continuous with respect to Q . Then $\delta = 1$, so $P(M) = 1$, which is the conclusion of Theorem 2.

The idea of the proof of Theorem 3 is straightforward. Think of P as being determined by flipping a coin with bias δ . With probability δ , P is absolutely continuous with respect to Q and equal to P^a . With probability $1 - \delta$, P and Q are mutually singular and P is equal to P^s . If P ends up equal to P^a , then the Blackwell-Dubins theorem tells us that P expects to merge with Q . Hence, M occurs with probability δ , which is the first conclusion of our result. If, on the other hand, P ends up equal to P^s , then $P(A) = 1$ and $Q(A) = 0$ for some event A . No amount of conditionalizing can change these extreme probability assignments, so P and Q must polarize in the limit. Hence, L occurs with probability $1 - \delta$. Making this argument precise requires some careful management of probability 0 events, which is where the assumption that evidence is shared comes into play. The complete proof can be found in the [Appendix](#).

Theorem 3 also implies the converse to the Blackwell-Dubins that was advertised in the last section. Suppose that P shares an increasing and complete sequence of evidence with Q

and that P expects to merge with Q . Then $1 = P(M) = \delta$ by Theorem 3. So P must be absolutely continuous with respect to Q . We see that in the presence of shared, increasing, and complete evidence absolute continuity is necessary for merging of opinions to occur with probability 1. We record this fact as the following corollary.

Corollary 1. *If P shares an increasing and complete sequence of evidence with Q , and P expects to merge with Q , then P is absolutely continuous with respect to Q .²¹*

One thing that we would like to stress about Theorem 3 is that the way in which we relax absolute continuity is quite mild because we continue to assume that absolute continuity relation holds with respect to evidential events. Our assumption is that for any event E that can be expressed as $E = E_1 \star \dots \star E_k$ (where each E_i is an element of some partition \mathcal{E}_n and \star is some operation on sets), that is, any event that can be settled by a finite amount of evidence, $P(E) = 0$ if $Q(E) = 0$. In other words, absolute continuity is relaxed just for infinitary events. Relaxing absolute continuity *further* than we have would be to relinquish the shared evidence assumption: there may be events that P can learn at some stage n that Q cannot. So, our assumption is a motivated way to mildly relax absolute continuity. However, the conclusion of our theorem is importantly different from that of Blackwell and Dubins's. This suggests to us that the classic merging of opinions results, by failing to be robust even to our mild weakening of the assumptions, are something of an artifact of the under-motivated but strong assumption of absolute continuity.

7. DISCUSSION

About polarization, Thomas Kelly asks, “Given that You and I are responding to our evidence in such-and-such a way, is there any chance that our doing so is anything other than blatantly unreasonable?” (2008, p. 631). More generally, are there grounds to deny TOTAL? Our study here provides an affirmative answer when the standard of reasonableness is Bayesian. As we have shown, it is trivial to find instances of polarization and persistent disagreement when agents learn just a finite amount of shared evidence. In the general case, the guarantee of asymptotic consensus for Bayesians is an artifact of special auxiliary assumptions, like absolute continuity, that supplement Bayesian updating. Claims that rational learning leads to consensus require endowing the auxiliary assumptions with a normative status. We find no plausible channel for such an endowment.

Our investigation could be taken in two different ways. On the one hand, we could hold to the normative standard provided by Bayesian probability theory. On the other hand, we could deny that the standard Bayesian picture of learning presents us with a sound scientific methodology. In the former case, it seems that we must relinquish any requirement that sound scientific methodology provides investigators with resources to resolve disagreements through shared evidence. Again, somewhat confusingly, claims contrary to this view are routinely made from inside the Bayesian camp. Suppes, for example, writes,

It is of fundamental importance to any deep appreciation of the Bayesian viewpoint to realize the particular form of the prior distribution expressing beliefs held before the experiment is conducted is not a crucial matter [...] For the Bayesian, concerned as he is to deal with the real world of ordinary and

²¹We note that this result was first proved by Kalai and Lehrer (1994). In the Appendix we actually generalize their result because their proof assumes that evidence is represented by partitions while ours does not.

scientific experience, the existence of a systematic method for reaching agreement is important [...] The well-designed experiment is one that will swamp divergent prior distributions with the clarity and sharpness of its results, and thereby render insignificant the diversity of prior opinion (1966, p. 204).

Similar claims in the context of merging of opinions can be found even in the most recent literature.²²

Consider now the second way to construe our study. There is more to scientific methodology, this reply goes, than is dreamt of in the Bayesian philosophy of coherence and conditionalization.²³ If we are to retain a probabilistic epistemology, we must constrain the class of “rationally permissible” priors rather substantially. But in a broadly Bayesian paradigm, the consensus requirement is tied to absolute continuity. Not only is absolute continuity a necessary condition for merging (whenever evidence is shared) that lacks normative motivation; it is a condition which we have no reason to expect to be satisfied in many cases. Furthermore, Theorem 2 only secures merging almost surely.

Holding out hope that a case can be made for the normative status of the absolute continuity condition, one might think that perhaps there are other aspects of sound methodology that would imply absolute continuity, and thereby secure merging. What is the nature of these additional aspects of scientific methodology? Are they all agent-invariant, or are some of them more subjective? By focusing solely on revising judgments of subjective probability *via* conditionalization, we have made the case against TOTAL harder to make than it might have been—at least in one sense. Adding parameters that are *not* invariant across rational agents would only make achieving consensus a less likely outcome. According to Isaac Levi’s epistemological outlook, to take an example with many subjective parameters, not only is what an agent “learns” or comes to accept determined in part by a subjective *value for information*, an agent can also revise her prior independently of learning new evidence in some circumstances (Levi, 1980).²⁴ So eliminating disagreements *via* rational learning is even less plausible on an account like Levi’s. But what about additional aspects of good methodology that are objective? A popular objective sort of constraint on probability judgments is the Principal Principle. But it is difficult to see how such a principle would help. In general, the Principal Principle does not pin down a unique prior probability (if for no other reasons than that would require, first, too much shared “knowledge of the chances” of events and, second, that chances are reasonably attributable to *all* relevant events).²⁵ Typically, only small fragments of a distribution are determined by it. This, of course, leaves ample occasion for the failure of absolute continuity between different priors. So much the worse for TOTAL.

²²See, for example, “We follow, e.g., Peirce in requiring that sound Scientific methodology provides investigators with the resources to resolve interpersonal disagreements through shared evidence” (Cisewski et al., 2017). But for which priors, with respect to how much evidence? And surely or only almost surely?

²³We are not considering approaches that fully abandon Bayesianism or fully denounce its subjective elements.

²⁴In Levi’s terminology, an agent can revise her *confirmational commitment*, which is a function from states of full belief to sets of probabilities, should circumstances call for it—without changing her state of full belief (see, e.g., Levi, 2009, §2).

²⁵An anonymous referee points out that a result due to Deutsch in the context of the many-worlds interpretation of quantum mechanics can be interpreted as delivering a unique rational probability. See (Greaves, 2007, p. 115).

APPENDIX

Examples from Section 4. The following two examples were referred to in Section 4. The first is a case of polarization with respect to an event accompanied by a decrease in total variational distance.

Example 4. Let (Ω, \mathcal{F}) , A , and E be defined as in Example 1, and consider the following priors and posteriors

TABLE 4

| | ω_1 | ω_2 | ω_3 | ω_4 |
|---------|------------|------------|------------|------------|
| P_1 | 1/24 | 5/12 | 1/12 | 11/24 |
| P_2 | 1/2 | 1/12 | 1/4 | 1/6 |
| P_1^E | 1/3 | 0 | 2/3 | 0 |
| P_2^E | 2/3 | 0 | 1/3 | 0 |

We still have polarization with respect to A because $P_1(A | E) = 1/3 < 11/24 = P_1(A) \leq P_2(A) = 13/24 < 2/3 = P_2(A | E)$. And yet, $d(P, Q) = 5/8 > 1/3 = d(P_1^E, P_2^E)$. \triangle

The next example is a case of global polarization without polarization with respect to any event.

Example 5. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $\mathcal{F} = 2^\Omega$, and $E = \{\omega_1, \omega_2\}$. Consider the priors and posteriors in the table below.

TABLE 5

| | ω_1 | ω_2 | ω_3 |
|---------|------------|------------|------------|
| P_1 | 1/3 | 1/3 | 1/3 |
| P_2 | 4/15 | 2/5 | 1/3 |
| P_1^E | 1/2 | 1/2 | 0 |
| P_2^E | 2/5 | 3/5 | 0 |

We have $d(P_1, P_2) = 1/15 < 1/10 = d(P_1^E, P_2^E)$, so E polarizes P_1 and P_2 globally. However, there is no subset of Ω with respect to which E polarizes P_1 and P_2 . This is straightforward to verify and we omit the details. \triangle

Proof of Theorem 3. In the remainder of this appendix, we will provide a more formal and general presentation of the mathematical framework of Part 2. Then we will prove Theorem 3.

Let (Ω, \mathcal{F}, P) be a probability space. We say that an event $A \in \mathcal{F}$ occurs *almost surely* with respect to P when $P(A) = 1$. We also say A occurs *a.s.* (P), and if Q is another probability measure on (Ω, \mathcal{F}) with respect to which A occurs almost surely, we say that A occurs a.s. (P/Q). We denote the indicator function for A by $\mathbf{1}_A : \Omega \rightarrow \{0, 1\}$, defined as

$$\mathbf{1}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A; \\ 0, & \text{otherwise} \end{cases}$$

When \mathcal{F} and \mathcal{G} are both sigma-algebras of subsets of Ω , we call \mathcal{G} a *sub-sigma-algebra* of \mathcal{F} if $\mathcal{G} \subseteq \mathcal{F}$. Intuitively, sigma-algebras represent bodies of information. For example, consider tossing a coin N times. Prior to any tossing, one is uncertain what the actual sequence of tosses will be and has a prior distribution over all the binary sequences of length N . After observing the first toss one now knows whether the actual sequence begins with H or T . This information is represented by the sigma-algebra \mathcal{G} that partitions all the possible binary sequences into those beginning with H and those beginning with T . Formally,

$$\mathcal{G} = \{\emptyset, \{\omega \in \Omega : \omega \text{ begins with } H\}, \{\omega \in \Omega : \omega \text{ begins with } T\}, \Omega\}.$$

If \mathcal{F} is a sigma-algebra, then a real-valued function $X : \Omega \rightarrow \mathbb{R}$ is called \mathcal{F} -*measurable* if $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for every $x \in \mathbb{R}$. Intuitively, the function X is \mathcal{F} -measurable if every question about the values that X takes can be answered by the information in \mathcal{F} . To take a simple example, if \mathcal{G} is defined as above and $A = \{\omega \in \Omega : \omega \text{ begins with } H\}$, then $\mathbf{1}_A$ is \mathcal{G} -measurable.

In the main text, we assumed for simplicity that evidence is represented by a sequence of finite partitions and that at each stage an agent learns an event in a partition. We now relax that assumption. We now assume that evidence is represented by a *filtration* $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ on (Ω, \mathcal{F}) , which is defined to be a collection of sub-sigma-algebras of \mathcal{F} such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all $n \in \mathbb{N}$. To see that this generalizes the partition model used in the main text, note that every finite partition \mathcal{E}_n generates a sub-sigma-algebra, the members of which are unions of members of \mathcal{E}_n (as well as \emptyset). The inclusion relationship $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ is the generalization of the above requirement that later partitions refine earlier ones and captures the idea that evidence is *increasing*. Let \mathcal{F}_∞ be the smallest sigma-algebra containing $\bigcup_{n \in \mathbb{N}} \mathcal{F}_n$. In the general setting, we say that the filtration $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is *complete* if \mathcal{F} is generated by the sub-sigma-algebras \mathcal{F}_n in the sense that $\mathcal{F}_\infty = \mathcal{F}$. As before, the point of this assumption is that the evidential information contained in the filtration eventually captures all events of interest (events in \mathcal{F}).

Since evidence is represented by a filtration, we need a notion of conditional probability given a sub-sigma-algebra, whereas in the main text we were able to make do with the familiar notion of conditional probability given an event. We use the definition of conditional probability that is standard in modern probability theory. The conditional probability $P(A|\mathcal{F}_n)$ of A given the sub-sigma-algebra \mathcal{F}_n is an \mathcal{F}_n -measurable function that satisfies $P(A \cap E) = \int_E P(A|\mathcal{F}_n) dP$ for all $E \in \mathcal{F}_n$. The existence of such a function is guaranteed for any sub-sigma-algebra by the Radon-Nikodym Theorem and is unique up to sets of P -measure 0. In other words, any \mathcal{F}_n -measurable function X that satisfies the given integral equation is almost surely equal to $P(A | \mathcal{F}_n)$ with respect to P , and we say that X is a *version* of $P(A | \mathcal{F}_n)$. Note that in the simple case where each \mathcal{F}_n is generated by a finite partition \mathcal{E}_n the conditional probabilities given \mathcal{F}_n are given by

$$P(A | \mathcal{F}_n)(\omega) = P(A | E_n(\omega))$$

for all ω in a set with probability 1, which corresponds with the treatment of conditional probabilities in the main text.

The convergence-to-the-truth theorem, as stated in the main text, generalizes to

$$\lim_{n \rightarrow \infty} P(A|\mathcal{F}_n) = \mathbf{1}_A \text{ a.s.}(P).$$

for all $A \in \mathcal{F}$. As in the main text, all limits are pointwise in $\omega \in \Omega$. Another way of stating convergence to the truth, then, is

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} P(A | \mathcal{F}_n)(\omega) = \mathbf{1}_A(\omega)\}) = 1$$

Following our previous notation, we write $P_n(\omega) = P(\cdot|\mathcal{F}_n)(\omega)$ and $Q_n(\omega) = Q(\cdot|\mathcal{F}_n)(\omega)$. In order to state the Blackwell-Dubins merging of opinions theorem in full generality, we need to make a further assumption about P_n and Q_n . For all that we have said so far, for fixed $\omega \in \Omega$, either or both of $P_n(\omega)$ and $Q_n(\omega)$ may fail to be probability measures on (Ω, \mathcal{F}) . If there are versions of P_n and Q_n such that $P_n(\omega)$ and $Q_n(\omega)$ are probabilities for all $\omega \in \Omega$, then we say that these versions are *regular* conditional probabilities. Under the assumption that \mathcal{F}_n is generated by a finite partition, regular versions of P_n and Q_n exist. We can also guarantee the existence of regular versions of P_n and Q_n by making assumptions about the space (Ω, \mathcal{F}) . For example, if Ω is a Polish space and \mathcal{F} is its Borel sigma-algebra, then regular versions of the conditional probabilities exist. But without further assumptions about the probability space or filtration, regular versions may not exist, so we now add their existence as a further assumption of the framework.²⁶

If P is absolutely continuous with respect to Q , then we write $P \ll Q$. We say that P *shares evidence with* Q if $P|_{\mathcal{F}_n} \ll Q|_{\mathcal{F}_n}$ for all n , where $P|_{\mathcal{F}_n}$ and $Q|_{\mathcal{F}_n}$ denote the restrictions of P and Q to the sub-sigma-algebra \mathcal{F}_n . Note that this corresponds with the definition of sharing evidence in the main text when \mathcal{F}_n is generated by a finite partition. We say that P and Q *merge* when

$$\lim_{n \rightarrow \infty} d(P_n, Q_n) = 0.$$

The general version of Theorem 2 is

Theorem 4 (Blackwell and Dubins (1962)). *Let P and Q be probability measures on (Ω, \mathcal{F}) , and let $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ be a complete filtration. Suppose that P_n and Q_n are regular conditional probabilities for all n . If $P \ll Q$, then P and Q merge with P -probability 1, i.e.*

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} d(P_n(\omega), Q_n(\omega)) = 0\}) = 1.$$

The general version of Theorem 3, which we are now ready to prove, is

Theorem 5 (Bayesian Consensus-or-Polarization Law, General Version). *Let P and Q be two probability measures on (Ω, \mathcal{F}) , and let $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ be a complete filtration. Suppose that P_n and Q_n are regular conditional probabilities for all n , that P shares evidence with Q , and that P is absolutely continuous with respect to Q to degree δ . Then*

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} d(P_n(\omega), Q_n(\omega)) = 0\}) = \delta$$

and

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} d(P_n(\omega), Q_n(\omega)) = 1\}) = 1 - \delta.$$

Proof. If $\delta = 1$, then the result follows from Theorem 4, so assume $\delta < 1$. Let the Lebesgue decomposition of P with respect to Q be given by

$$P = \delta P^a + (1 - \delta)P^s,$$

as in equation (2) of the main text. Clearly, $P^a \ll P$ and $P^s \ll P$.

By the triangle inequality, $P^a \ll Q$, $P^a \ll P$, and Theorem 4,

$$d(P_n, Q_n) \leq d(P_n, P_n^a) + d(P_n^a, Q_n) \rightarrow 0 \text{ a.s. } (P^a) \quad (3)$$

With $M = \{\omega \in \Omega : d(P_n(\omega), Q_n(\omega)) \rightarrow 0\}$, (3) implies

$$P^a(M) = 1. \quad (4)$$

²⁶For more on regular conditional probabilities, see Seidenfeld (2001) and Durrett (2010, 4.1c).

Next, let $A^s \in \mathcal{F}$ be such that $Q(A^s) = 0$ and $P^s(A^s) = 1$. Such an event exists because P^s and Q are mutually singular. Since $Q(A^s) = 0$, we have $Q_n(A^s) = 0$ a.s. $(Q/Q|_{\mathcal{F}_n})$ since Q_n is \mathcal{F}_n -measurable. This implies $Q_n(A^s) = 0$ a.s. $(P|_{\mathcal{F}_n}/P)$, which then implies $Q_n(A^s) = 0$ a.s. (P^s) . Since $P^s(A^s) = 1$, we have $P_n^s(A^s) = 1$ a.s. (P^s) . In this way, for all n we find a P^s -probability 1 set on which $Q_n(A^s) = 0$ and $P_n^s(A^s) = 1$. It follows that $d(P_n^s, Q_n) = 1$ for all n a.s. (P^s) . Using this fact, the triangle inequality, $P^s \ll P$, and Theorem 4, we get

$$d(P_n, Q_n) \geq d(P_n^s, Q_n) - d(P_n^s, P_n) \rightarrow 1 \text{ a.s. } (P^s) \quad (5)$$

With $L = \{\omega \in \Omega : d(P_n(\omega), Q_n(\omega)) \rightarrow 1\}$, (5) implies

$$P^s(L) = 1. \quad (6)$$

Now, $M \subseteq L^c$ and $L \subseteq M^c$, so (4) and (6) imply

$$P^a(L) = 0 = P^s(M). \quad (7)$$

Using (4), (6), (7) and the Lebesgue decomposition (2) of P with respect to Q we compute

$$P(M) = \delta \text{ and } P(L) = 1 - \delta,$$

which is the desired result. \square

REFERENCES

- Autzen, B. (2017). Bayesian convergence and the fair-balance paradox. *Erkenntnis* 83(2), 253–263.
- Belot, G. (2013). Bayesian orgulity. *Philosophy of Science* 80(4), 483–503.
- Blackwell, D. and L. E. Dubins (1962). Merging of opinions with increasing information. *The Annals of Mathematical Statistics* 33(3), 882–886.
- Christensen, D. (2009). Disagreement as evidence: The epistemology of controversy. *Philosophy Compass* 4(5), 756–767.
- Cisewski, J., J. Kadane, M. Schervish, T. Seidenfeld, and R. Stern (2017). Standards for Modest bayesian credenes. *Philosophy of Science* 85(1), 2018.
- DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association* 69(345), 118–121.
- Durrett, R. (2010). *Probability: Theory and Examples*. Cambridge University Press.
- Earman, J. (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
- Edwards, W., H. Lindman, and L. J. Savage (1963). Bayesian statistical inference for psychological research. *Psychological Review* 70(3), 193–242.
- Elkin, L. and G. Wheeler (2018). Resolving peer disagreements through imprecise probabilities. *Noûs* 52(2), 260–278.
- Gaifman, H. and M. Snir (1982). Probabilities over rich languages, testing and randomness. *The Journal of Symbolic Logic* 47(03), 495–548.
- Genest, C. and J. V. Zidek (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science* 1(1), 114–135.
- Glymour, C. (1980). *Theory and Evidence*. Princeton University Press.
- Greaves, H. (2007). Probability in the everett interpretation. *Philosophy Compass* 2(1), 109–128.
- Hegselmann, R., U. Krause, et al. (2002). Opinion dynamics and bounded confidence: Models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation* 5(3).
- Herron, T., T. Seidenfeld, and L. Wasserman (1994). The extent of dilation of sets of probabilities and the asymptotics of robust bayesian inference. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1994, pp. 250–259. Philosophy of Science Association.
- Herron, T., T. Seidenfeld, and L. Wasserman (1997). Divisive conditioning: Further results on dilation. *Philosophy of Science* 64(3), 411–444.
- Huttegger, S. M. (2015). Merging of opinions and probability kinematics. *The Review of Symbolic Logic* 8(04), 611–648.
- Jeffrey, R. (2004). *Subjective Probability: The Real Thing*. Cambridge University Press.
- Jern, A., K.-M. K. Chang, and C. Kemp (2014). Belief polarization is not always irrational. *Psychological review* 121(2), 206–224.
- Kalai, E. and E. Lehrer (1994). Weak and strong merging of opinions. *Journal of Mathematical Economics* 23(1), 73–86.
- Kelly, T. (2008). Disagreement, dogmatism, and belief polarization. *The Journal of Philosophy* 105(10), 611–633.
- Lehrer, K. and C. Wagner (1981). *Rational Consensus in Science and Society: A Philosophical and Mathematical Study*, Volume 21. Springer.
- Levi, I. (1980). *The Enterprise of Knowledge*. MIT Press, Cambridge, MA.
- Levi, I. (1982). Conflict and social agency. *The Journal of Philosophy* 79(5), 231–247.
- Levi, I. (1985a). Consensus as shared agreement and outcome of inquiry. *Synthese* 62(1), pp. 3–11.
- Levi, I. (1985b). Imprecision and indeterminacy in probability judgment. *Philosophy of Science* 52(3), 390–409.
- Levi, I. (2009). Why indeterminate probability is rational. *Journal of Applied Logic* 7(4), 364–376.
- Lewis, D. (1980). A subjectivist’s guide to objective chance. In W. L. Harper, R. Stalnaker, and G. Pearce (Eds.), *IFS*, pp. 267–297. Springer.

- Lord, C. G., L. Ross, and M. R. Lepper (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37(11), 2098–2109.
- Miller, R. I. and C. W. Sanchirico (1999). The role of absolute continuity in “merging of opinions” and “rational learning”. *Games and Economic Behavior* 29(1-2), 170–190.
- Nielsen, M. (2018). Deterministic convergence and strong regularity. *The British Journal for the Philosophy of Science*, Forthcoming.
- Pedersen, A. P. and G. Wheeler (2014). Demystifying dilation. *Erkenntnis* 79(6), 1305–1342.
- Pedersen, A. P. and G. Wheeler (2015). Dilation, disintegrations, and delayed decisions. In *ISIPTA '15: Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications*. Aracne Editrice.
- Peirce, C. S. (1992a). The fixation of belief. In N. Houser and C. Kloesel (Eds.), *The Essential Peirce, Volume 1: Selected Philosophical Writings (1867–1893)*, pp. 109–123. Indiana University Press.
- Peirce, C. S. (1992b). How to make our ideas clear. In N. Houser and C. Kloesel (Eds.), *The Essential Peirce, Volume 1: Selected Philosophical Writings (1867–1893)*, pp. 124–141. Bloomington: Indiana University Press.
- Reichenbach, H. (1938). *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge*. University of Chicago Press.
- Ross, L. and C. A. Anderson (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Chapter 9, pp. 129–152. New York: Oxford University Press.
- Savage, L. (1972, originally published in 1954). *The Foundations of Statistics*. New York: John Wiley and Sons.
- Schervish, M. and T. Seidenfeld (1990). An approach to consensus and certainty with increasing evidence. *Journal of Statistical Planning and Inference* 25(3), 401–414.
- Seidenfeld, T. (2001). Remarks on the theory of conditional probability: Some issues of finite versus countable additivity. In V. F. Hendricks (Ed.), *Probability Theory: Philosophy, Recent History and Relations to Science*. Synthese Library, Kluwer.
- Seidenfeld, T., J. B. Kadane, and M. J. Schervish (1989). On the shared preferences of two bayesian decision makers. *The Journal of Philosophy* 86(5), 225–244.
- Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2010). Coherent choice functions under uncertainty. *Synthese* 172(1), 157–176.
- Seidenfeld, T. and L. Wasserman (1993). Dilation for sets of probabilities. *The Annals of Statistics* 21(3), 1139–1154.
- Shimony, A. (1955). Coherence and the axioms of confirmation. *The Journal of Symbolic Logic* 20(01), 1–28.
- Shogenji, T. (1999). Is coherence truth conducive? *Analysis* 59(264), 338–345.
- Skyrms, B. (1995). Strict coherence, sigma coherence and the metaphysics of quantity. *Philosophical Studies* 77(1), 39–55.
- Stewart, R. T. and M. Nielsen (2018). Another approach to consensus and maximally informed opinions with increasing evidence. *Philosophy of Science*, Forthcoming.
- Stewart, R. T. and I. Ojea Quintana (2018). Probabilistic opinion pooling with imprecise probabilities. *Journal of Philosophical Logic* 47(1), 17–45.
- Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy* 10(2), 175–195.
- Suppes, P. (1966). A bayesian approach to the paradoxes of confirmation. *Studies in Logic and the Foundations of Mathematics* 43, 198–207.
- Wasserman, L. and T. Seidenfeld (1994). The dilation phenomenon in robust bayesian inference. *Journal of Statistical Planning and Inference* 40(2), 345–356.
- White, R. (2005). Epistemic permissiveness. *Philosophical Perspectives* 19(1), 445–459.

- White, R. (2010). Evidential symmetry and mushy credence. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 3, pp. 161–186. Oxford University Press New York.
- Zimper, A. and A. Ludwig (2009). On attitude polarization under bayesian learning with non-additive beliefs. *Journal of Risk and Uncertainty* 39(2), 181–212.