Episteme (2025), 1–17
doi:10.1017/epi.2025.10049

episteme

ARTICLE

Arguing in the Face of Disagreement: Is It Worth It?

Marc Jiménez-Rolland

Departamento de Filosofía, Universidad Autónoma Metropolitana, Iztapalapa, Mexico
Email: mjrolland@izt.uam.mx

(Received 19 December 2024; revised 16 April 2025; accepted 13 May 2025)

Abstract

Argumentation is often conceived as a rational response to disagreement, even when it does not resolve differences of opinion. Arguing in the face of disagreement has, however, distinctive epistemic effects. Sometimes argumentation achieves convergence of opinion or at least the mutual recognition that a more thorough inquiry is required. But facing disagreement, participants of argumentative exchanges quite often remain steadfast in their initial views or even radicalize them. Can we make sense of these latter situations? To account for their occurrence, it is common to point out that people's ability to argue is flawed, that an "argumentative culture" is lacking, and that emotional and other non-rational factors often interfere in confrontative situations. But these suggestions do not amount to a thorough satisfactory explanation. In this paper, I provide the outline of a purely epistemic account of these peculiar effects of argumentation in the face of disagreement. I argue that probabilistic models of degrees of confidence (or "credences") can shed light on the conditions that give rise to several of these effects. This could provide some guidance on how to avoid them.

Keywords: Bayesian epistemology; credal and doxastic disagreements; epistemic models of argumentation; polarization; steadfastness

1. Introduction

Among its many uses, argumentation is often pursued aiming to provide a rational response to disagreement. "*En parlant on se comprend,*" as the saying goes.[1] By means of rational exchange, we might understand each other better, overcome our differences, and converge in our views, fostering peaceful coexistence and promoting cooperation. That is the rosy picture.[2] However, this kind of convergence seldom occurs in

[1]"By speaking we understand each other." According to Louis-Philippe Geoffrion "this proverbial phrase . . . means approximately that those who neither speak loudly enough, nor often enough, nor at enough length are misunderstood . . . Let us therefore speak, if we want to be understood" (Geoffrion 1925: 86–87). Unless otherwise indicated, the translations of texts originally in foreign languages are mine.

[2]Although with some nuances, something akin to this rosy picture is present in many contemporary influential approaches to argumentation theory (Aikin and Talisse 2019; 2020; Eemeren 2010; 2018; Gilbert 1995; Walton 1998). Even those who contest it this rosy picture recognize its significant theoretical

argumentative reality. Instead, rather peculiar situations typically arise. Sometimes participants remain steadfast in their initial position, even while recognizing that the reasons being offered to them are relevant to the issue at hand. When no new evidence is being offered, some participants often intensify their views. If this happens to antagonists in an exchange, their extreme initial positions are accentuated, thus widening the gap that existed before the conversation took place.

It is a platitude that argumentation does not often lead to the rational resolution of disagreement, as the rosy picture suggests. This fact is widely recognized, to the point of being the subject of an encyclopedic entry: "The usual *result* of arguing [. . .] is that everyone remains more attached to their convictions than before"[3] (Diderot and D'Alembert 1765: 196). While arguing in response to disagreement, some beliefs are rendered impervious to change in the light of evidence and views are radicalized in the absence of new evidence for them. Worse still: in many cases, argumentation seems to be *the driving force* for these peculiar effects on people's thinking. Although argumentation has all kinds of effects – e.g., emotional, social, political, and so forth – here I focus exclusively on its *epistemic* effects: the impact of argumentation in beliefs or in degrees of confidence (credences) of participants in the exchange. My aim in this paper is to provide the broad outline of a purely epistemic account of how argumentation can *typically* produce epistemic effects like the ones described above. For the kind of understanding that this account aims to provide, it is not enough to hint at some indefinite flaw or deficiency. Rather, specific conditions under which these effects are to be expected should be identified. Insofar as these results are not only oddities but also undesirable, this model should offer some advice on interventions to avoid them; this, however, might not always translate into personal advice to improve argumentation.

Against the prevailing view in contemporary argumentation theory, I deploy tools from logic and formal epistemology, contending that they are furnished with descriptive and normative resources to account for a wide range of argumentative effects. I start by presenting some clarifications concerning my use of the terms "argumentation," "rationality," and "disagreement." Next, I outline a simple model for argumentative phenomena. The epistemic effects of argumentation in "ideal" situations are illustrated according to this model. Then, I discuss how some distinctive epistemic effects of argumentation in the face of disagreement are prone to come about. I conclude with some remarks about the kind of understanding we get from models like the one outlined in this paper, as well as their potential use for improving the quality of argumentative discourse.

2. Argumentation, rationality, and disagreement

Since these notions play a heavy role in a model which aims to account for the epistemic effects of argumentation in response to disagreement, in this section I clarify my use of "argumentation," when its epistemic effects are deemed to be "rational," and what can be counted as a "disagreement."

Several perspectives for the study of argumentation focus on non-overlapping phenomena, amenable to different kinds of evaluations. Roughly, for rhetorical perspectives argumentation involves strategic attempts to persuade an audience – e.g., ways of "influencing others' opinions and expectations without damaging

momentum within argumentation studies (Aikin and Casey 2022; Doury 2012; Goodwin 2007; Paglieri 2009).

[3]"Le *résultat* ordinaire des disputes [. . .] c'est que chacun demeure plus attaché à son sentiment qu'auparavant" (Diderot and D'Alembert 1765: 196).

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

cooperative bonds" (Bernache Maldonado 2025: 130). In dialectical perspectives, participants must perform certain roles in several stages for argumentative exchanges to occur (Eemeren 2018: 34–36). Here, I embrace a logical perspective. From this standpoint, argumentation requires *arguments:* inferentially structured sets of claims. Typically, communicative exchanges containing them involve people making statements that purport to evidentially contribute to the topic under discussion. The assessment of these episodes involves *argument* evaluation,[4] which considers the qualities of constituent claims as well as those of inferential structure. On a positive evaluation, epistemic "goods" or "assets" can be associated with statements (e.g., truth, knowledge, justification, among others) and/or with their inferential link (e.g., truth-preservation, coherence, confirmation, understanding, among others).[5]

By way of illustration, consider a sound deductive argument. According to this logical perspective, the positive assessment of such an argument relies on the quality of its constituent statements (i.e., its premises are *true*, perhaps *known to be true*) and of their evidential link to its conclusion (i.e., its structure is *valid*, or *truth-preserving*). Thus, a sound deductive argument can provide knowledge or understanding of its conclusion, by competently deducing its truth from that of the premises. Its statements and its structure are jointly "truth-conducive" at the highest possible degree. Although this might be used as template for evaluation from a logical perspective, it is not a feature of this approach that *only* sound deductive arguments are good arguments. In a broader sense, "truth-conduciveness" could be generalized to uncertain deductions and non-deductive arguments. Thus, the features of arguments that are relevant to their evaluation from this logical perspective concern "epistemic rationality."

Let us start with a slightly restrictive and highly idealized[6] characterization of epistemic rationality in a communicative exchange. An epistemically rational agent engaged in argumentation seeks to maximize her epistemic assets. Starting from a coherent epistemic situation, she incorporates the "evidence" produced during the exchange. This includes the content of what her interlocutors say, the very fact that they make a claim, as well as the information available or salient in the conversational context. Two constraints govern this process of incorporating evidence.[7] First: the agent's degree of confidence in statements that she does not recognize as logical truths should be responsive to (some) evidence. Second: new information should be incorporated in a systematic way, seeking to preserve (synchronic) coherence in beliefs and degrees of confidence, and (diachronically) updating them in ways that are truth-

[4]This perspective is compatible with incorporating evaluative considerations from other approaches, but they are subordinated to the logical evaluation of arguments: to the extent that they lack good arguments, those exchanges that have other merits cannot be considered good argumentative episodes.

[5]Broncano-Berrocal and Carter (2021) offer a broader catalog of epistemic assets that might be expected from argumentation in the face of disagreement.

[6]By claiming that this characterization is "highly idealized" it is not meant that it is a description of "ideal" epistemic agents (i.e., of exemplars we should try to emulate). Rather, what is meant is that it offers a simplified sketch to which details could be added and from which distortions could be removed to describe *real* epistemic agents. If "most" people —most of the time— turn out to be "good approximations to" epistemically rational agents, it would not be surprising that the logical perspective is able to describe and, to some extent, predict the epistemic effects of argumentation occurring in *actual* disagreements. Although this assumption is not directly tested in this paper, the main argument addresses a similar claim: insofar as it offers an empirically adequate model of (apparently defective) epistemic effects of argumentation in the face of disagreement, the logical perspective has the descriptive and evaluative resources to understand (apparently imperfect) arguers.

[7]These constraints are not meant to be action-guiding precepts that the agents attempt to follow; rather, they provide a standard to which (consciously or not) rational agents adhere to.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

conducive upon adequate evidence. This is a highly idealized characterization of epistemic rationality, since it abstracts away from a constellation of factors that do matter in people's lives and that are often present in argument. Furthermore, it introduces an unrealistic degree of precision, which we do not use to describe the chaotic real world.

Since argumentative exchanges in the face of disagreement are the focus of our attention, before presenting a simple model of their epistemic effects, I offer a more precise characterization of "disagreement." In talking about disagreements, their subject matter is presented as claims or statements, which could be the premises or conclusions of arguments. Consider a statement p, e.g., the claim that God exists. Three doxastic attitudes toward the truth of p are usually distinguished: (1) Believing that p is true (theism); (2) Believing that p is false (atheism); and (3) Suspending judgment concerning the truth of p (agnosticism). Thus, disagreement can be thought as:

> DOXASTIC DISAGREEMENT (D-DISAGREEMENT). People D-DISAGREE about p if and only if they hold different doxastic attitudes toward p. (Frances and Matheson 2018: §1)

Under this characterization, an "argumentative resolution" of D-DISAGREEMENTS requires that at least one of the parties changes their initial doxastic attitude toward p as a result of argument. However, even if there are changes in doxastic attitude, D-DISAGREEMENTS can be "argumentatively stagnated" if the participants end up with different doxastic attitudes.

Instead of speaking about doxastic attitudes towards p, we can talk about the degree of confidence or "credence" of a person toward p. As in our previous example, someone could be highly confident in p's truth, assigning its credence a value close to 1 (theistic fundamentalism); she could be more confident in p's truth than in its falsity, assigning its credence a value above the threshold of 0.5 but quite distant from 1 (fallibilistic theism); she could be equally confident in p's truth as in its falsity, with a precise credence of 0.5; or she could be more confident in p's falsity than in its truth, assigning its credence a value below the threshold of 0.5 (fallibilistic atheism), perhaps even close to 0 (atheistic fundamentalism). This allows another characterization of disagreement:

> CREDAL DISAGREEMENT (C-DISAGREEMENT). People C-DISAGREE about p if and only if they hold different credences[8] toward p. (Rowbottom 2018: §2)

For the "argumentative resolution" of a C-DISAGREEMENT, both parties must converge in assigning the exact same value to their credences toward p. For that to occur – or, more generally, for an argumentative "decrease" of C-DISAGREEMENT – one or both parties must "shift" their degree of confidence toward p "in the direction" of the other party's credence (i.e., they must "converge" with her on a value or on a narrower threshold than their starting point). C-DISAGREEMENTS can also be "argumentatively augmented" when, because of argument, one (or both) of the parties changes their degree of confidence toward p in an "opposite direction," in a way that is not compensated by a change in the degree of confidence by the other party (i.e., the distance in their degrees of confidence increases, widening the threshold). Even if there are changes in credence, C-DISAGREEMENT can be "argumentatively stagnated" if the distance in the degrees of confidence remains equal.

[8]As Rowbotton (2018: 224–227) discusses at length, one might think that there is a threshold —which perhaps varies contextually— for the difference in degrees of confidence to count as a case of disagreement.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

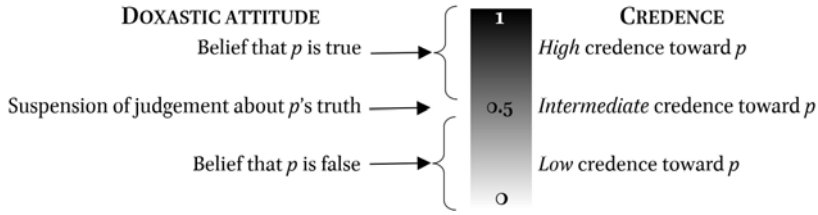

Figure 1. A possible mapping of doxastic attitudes to (intervals of) credences.

The distinction between D-DISAGREEMENT and C-DISAGREEMENT might appear unnecessary, and perhaps redundant.[9] By identifying doxastic attitudes with (intervals of) credences, we could dispense of one of the characterizations of disagreement just outlined, by means of a mapping as in Figure 1.

There are, however, several theoretical objections to identifying beliefs with (intervals of) credence.[10] For instance, retaining the distinction between credence and belief might be crucial to account for some apparent exceptions to plausible rational constraints on the logic of belief (Titelbaum 2019: 2–3; 2022: 8–9). To avoid the Lottery Paradox, even extremely low credences toward p should not be identified with belief in the falsity of p. If a 0.0001 credence in the statement "Ticket i will win the lottery" is not identified with the belief that ticket i will not win the lottery for each ticket of a fair lottery, there is no inconsistency in believing that at least one ticket will win the lottery. Analogously, to eschew the Preface Paradox, extremely high credences toward p should not be identified with belief in p's truth. An author who has carefully revised the contents of her book might believe in each claim therein included. Assigning a very high credence to the claim that "At least one statement in my book is false" (i.e., acknowledging the possibility of error in her Preface) is not inconsistent with belief (and high credence) in every single claim.

In addition to these normative advantages of preserving the distinction between credence and belief, it might be descriptively useful to distinguish both kinds of disagreements. Two agents with the same doxastic attitude toward p do not D-DISAGREE about p. They might, however, assign different credences to p (e.g., one might be much more confident than the other of p's truth); thus, they C-DISAGREE about p. Although this is a distinction compatible with mapping doxastic attitudes to (intervals of) credences, *it does not require it*. On the other hand, there could also be cases where two agents D-DISAGREE without there being some C-DISAGREEMENT involved. For that to occur, agents with the exact same credence toward p should have different doxastic attitudes toward p. This possibility is not preserved by a mapping as the one we presented in Figure 1. It might depict agents who adjust their degree of confidence to the same value, while retaining a prior commitment to their beliefs.

[9]Some authors suggest that these two models of disagreement only differ in how they represent belief: one "as a full state and the other as a graded one. [. . . B]oth models operate in debates over disagreement in epistemology, although the second one has several advantages" (Gensollen 2022: 95) —e.g., "levels of confidence, or degrees of belief [. . .] allow for more precision in discussing cases" (Frances and Matheson 2018: §1).

[10]There is a substantial literature addressing the descriptive and normative links between belief and credence (see Jackson 2020 for an overview). In what follows, I will assume that it is useful to represent agents as having both attitudes. This aligns with Weisberg's (2020) belief-credence dualism —which considers that both attitudes are equally fundamental, and neither reduces to the other. It is also compatible with the idea —endorsed by Leitgeb (2017)— that there are normative connections between belief and credence. However, the simplified model outlined in the next section does not assume that there are general, systematic connections between belief and credence, either descriptive or normative.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

This possibility will be explored in Section 4. In the meantime, both *doxastic* and *credal* will be treated as (potentially) genuine cases of disagreement.

We can now ask: What are the epistemic effects of arguing in the face disagreement? To answer this question, I will sketch a simple model using probabilities.

3. A simple model of argumentative epistemic effects

To explore the epistemic effects of argument, in this section I outline a simplified model[11] of some epistemic features of the participants in an argumentative exchange. Before attempting to account for peculiar cases as the ones mentioned at the outset, it will be useful to illustrate how this model represents the epistemic effects of argumentation in "ideal" situations.

Our model assumes a finite set of epistemically rational agents $\{a, b, c, \ldots, o\}$, representing the individuals taking part in the argumentative exchange. There is also a set of atomic propositions, denoted by letters $p, q, r, \ldots, u$ (with or without subscripts), which is closed under truth-functional operations. These are used to represent claims that are made within the conversational exchange, but they might also represent the information available to the participants at the outset or during the exchange. For each agent α, there is a probability function, $\mathrm{Pr}_\alpha(\Phi)$, which assigns real numbers between 0 and 1 to each (simple or complex) proposition Φ. These are credences reflecting α's degree of confidence in Φ: they can be high (close to 1), low (close to 0), or intermediate. In addition to credences, an agent α might also have (qualitative) beliefs toward propositions, $\mathrm{B}_\alpha(\Phi)$.

In this model, epistemic rationality imposes (synchronic) *consistency* constraints. At any time t, the contents of α's beliefs should not include or entail contradictions. Although each agent might assign different values to the same propositions, the consistency of their credence functions is equally constrained by the mathematical properties of the probability calculus. Additionally, all epistemically rational agents (diachronically) *update* their credences toward Φ by Conditionalization:[12] assuming that Ψ represents everything the agent learns in the interval of time t_1 and a later time t_2, the value in t_2 of $\mathrm{Pr}_\alpha(\Phi) = \mathrm{Pr}_\alpha(\Phi|\,\Psi)$ in t_1, given $\mathrm{Pr}_\alpha(\Psi)>0$; when nothing is learned in the interval, Ψ is considered to be a tautology. This reflects the agent's "balance of evidence": the epistemic effect of the agent rationally assimilating information relevant for her prior credences. But epistemic rationality does not only demand deductive and probabilistic (synchronic) consistency, as well as systematic (diachronic) evidential updating. It also requires the choice of beliefs considered as courses of action (i.e., as "cognitive options")[13] that *maximize expected epistemic utility*. In this model, epistemically rational agents seek two *epistemic goals:* (1) to gather truths and (2) to avoid falsities. An agent's beliefs considered as courses of action have an associated epistemic utility, $u\mathrm{B}_\alpha(\Phi)$, which depends on the truth and informativeness of the believed claim. When true propositions are more informative, they provide a greater utility for the agent who believes them. Some courses of action might also have an

[11]This credal model is adapted from a similar framework described in Titelbaum (2019; 2022); the extension to an epistemic decision theory is heavily inspired by Buchak (2013; 2021; 2023).

[12]Or a suiting generalization, like Jeffrey Conditionalization (Titelbaum 2019: 13–15; 2022: 156–161).

[13]Following Lara Buchak, it is worth emphasizing that this representation of beliefs as courses of action does not require viewing belief as something under our voluntary control. It is compatible with the view that we do not have *any* control over our beliefs, in which case the model "will tell us how to evaluate various automatic mechanisms that produce belief" (Buchak 2021: 199).

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

associated additional quasi-*epistemic benefit*, β, which may add to the epistemic utility of a course of action, when the believed proposition turns out to be true.

For our purposes, this model aims to perform two tasks concerning the *epistemic effects* of argumentation in the face of disagreement:

> EXPLANATION TASK: If the effects follow a *recurring pattern*, the model should explain *why they do occur.*

> OPTIMIZATION TASK: If the effects are *undesirable*, the model should provide guidance on *how to avoid them.*

Before examining the peculiar situations described at the outset, it is worth considering how this model would represent the epistemic effects of argumentation in "ideal" cases. From some additional assumptions, Robert Aumann obtained the following result:

> *Aumann's Theorem.* If two agents have the same prior credence toward p, and the result of their conditionalizing upon evidence is "common knowledge," then their final credence toward p is the same even if their posteriors are based on different information.[14]

Aumann points out that, "once one has the appropriate framework, [this result] is mathematically trivial. Intuitively, though, it is not quite obvious" (1976: 1236). Adapted to our terminology, the theorem describes situations where two agents do not C-DISAGREE about p prior to an exchange and share their evidential assessments concerning p. It concludes that the epistemic effect of this exchange cannot be a C-DISAGREEMENT about p in conditions of "full disclosure."[15] Linked to our topic, Aumann's theorem suggests that argumentation itself does not *generate* C-DISAGREEMENT when those involved agree and share their rational assessments of the evidence. This result is in line with the rosy picture's normative intuitions about the epistemic effects of argumentation. However, it would be hard to find situations that fit such a description. One reason for this is that agents often assign different credences to the disputed claim: their starting point is a C-DISAGREEMENT. The model seems to offer some additional encouraging prospects for the argumentative *resolution* of such disagreements, in light of various convergence theorems (e.g., Gaifman and Snir 1982). As Igor Douven explains,

> What these theorems say, at bottom, is that the difference between prior [credences] do not mean all that much, as these priors will get "washed out" by the evidence, that is to say, as we go, the influence of the [credences] we once started out with will gradually diminish, and our [credences] will come to be increasingly determined by the evidence we receive. (Douven 2011: 259)

Even if they start from a pronounced C-DISAGREEMENT, by updating from *shared evidence* rational agents will gradually converge in assigning the same value (or interval) to their credences toward a proposition.[16] Since this convergence responds directly to evidence, it seems to contribute to the epistemic goals of gathering truths and avoiding

[14]I have adapted the terminology from Aumann's (1976) formulation.

[15]Although we left this possibility open in the previous section, we have no grounds here to suspect that this would be a case of D-DISAGREEMENT.

[16]As in the previous case, we have no grounds here to suspect that this would be a D-DISAGREEMENT.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

falsities. These features of the model provide an optimistic forecast for the rational argumentative resolution of disagreement. For our current purposes, however, they pose a challenge. The robustness of these results (they are theorems!) suggests that the peculiar phenomena described at the outset of this paper are incompatible with the model. For convenience, I state them succinctly:

> STEADFASTNESS. While acknowledging the relevance of evidence presented in the argumentative exchange, a participant does not change her mind.
>
> ESCALATION. With no new supporting evidence (beyond the fact of disagreement itself), a participant intensifies her degree of confidence in her view.
>
> POLARIZATION. With no new supporting evidence (beyond the fact of disagreement itself), participants intensify their degrees of confidence in opposite directions.

These epistemic effects of argumentation diverge from the patterns suggested by the aforementioned theorems. In these scenarios, argumentation does not contribute to the rational resolution of disagreements; in some cases, it even aggravates them!

A popular suggestion is that the EXPLANATION TASK for these patterns lies beyond the reach of simple models like the one sketched above. Insofar as these epistemic effects are not adequately represented by the behavior of epistemically rational agents, they must be attributable to some other facts about real flesh-and-blood human beings. Models dispense of our cognitive biases and abstract from the emotional richness of our psychological life; social ties and affiliations are omitted too; the plethora of rhetorical niceties that flood our conversations is also left out. These factors could play a crucial role in describing the epistemic effects of argumentation in the face of disagreement.[17] Besides, idealized epistemic models place too many and too severe demands on lazy, careless, and clumsy creatures – as many of us (and our interlocutors) are. Thus, the forces preventing the argumentative resolution of disagreement might derive from deviations from epistemic rationality.

If we must resort to these kinds of alternative explanations, our model cannot fulfill the tasks we assigned to it. In that scenario, explaining why argumentation has these effects and offering prescriptions for how to avoid them would be beyond the scope of epistemic rationality, as characterized by the model. It would also require paying attention to features of argumentation that are not considered by the logical perspective. However, in the following sections I will show that these apparently irrational epistemic effects can be melted down and poured into the behavior molded for epistemically rational agents. The rigid cast resulting from this process will show that "the irrational is not merely the non-rational, which lies outside the ambit of the rational; irrationality is a failure within the house of reason" (Davidson 1982: 169). To achieve this, I will heavily rely on previous work.

4. Rational steadfastness in the face of disagreement

Lara Buchak describes situations in which one is prone "to adopt a conviction and to maintain that conviction when counterevidence arises [. . .] even if one's credences would not make it rational to adopt that belief anew" (2021: 210). Some beliefs a person

[17]That these factors do play a role in producing the epistemic effects of argumentation is not necessarily incompatible with the logical perspective. It would not seem problematic if they were mere anomalies or coincidences, or if they reinforced or causally overdetermined an underlying epistemic pattern.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

holds in such circumstances do not seem to sensibly respond to her balance of evidence. This seems incompatible with our model: this person is not behaving as one would expect from an epistemically rational agent, or these epistemic effects might be due to further non-epistemic factors excluded from the model. These kinds of situations are not unusual: they are typical, especially when it comes to controversial claims about important issues.[18] Buchak suggests that these situations involve "faith," understood as some kind of sustained resistance against counterevidence; but she also argues that such an attitude can be epistemically rational, under certain conditions. By specifying those conditions within the model, its ability to explain epistemic effects of argumentation in the face of disagreement can be tested.

An agent α having "faith" toward p, in an epistemic situation q with possible additional evidence E, is subject to the following conditions:

(i) α must believe that p and assign it an initially high (but less than 1) credence, conditional on q.
(ii) α must commit[19] to take a risk on p, regardless of counterevidence E.
(iii) α must follow through on her commitment even when receiving counterevidence E against p.

As we have seen, epistemically rational agents can have different initial beliefs and credences, so condition (i) seems compatible with the model. What seems problematic is (iii): How can such an agent *believe* that p while also acknowledging counterevidence that *diminishes her credence* toward p? This is a case of STEADFASTNESS.

To illustrate, imagine *you* (represented by α) and *I* (represented by ε) are arguing about an issue (p) on which we hold different doxastic attitudes and toward which we hold different degrees of confidence: we are both in D-DISAGREEMENT and in C-DISAGREEMENT (see Table 1).

We argue. As a result of the exchange, *you* and *I* come to agree on the balance of evidence, converging on the *exact same value* for our respective credence $\Pr_x(p)$, thus solving our C-DISAGREEMENT. Nevertheless, "it might still be the case that each of us should remain committed to our initial belief, despite now sharing the same credence" (Buchak 2021: 211). The possibility of such a situation is incompatible with the existence of a general principle governing synchronic rational relations between beliefs and credences.[20] It is worth noting that the model we presented does not *demand* such a principle. Thus, it is open to the possibility of epistemically rational agents having "faith" toward certain claims. The explanatory work of the model consists in detailing *under what conditions* having "faith" would maximize expected epistemic utility for an agent, thus rendering these situations typical. It is there that the ideas of "commitment to take a risk" and "following through on a commitment" play a crucial role in Buchak's explanation. Roughly, she shows general sets of conditions under which – given an initial epistemic situation compatible with (i) – committing to a belief is a course of action that maximizes expected epistemic utility compared to reevaluating that same belief after

[18]This is equivalent to Buchak's "FREQUENCY" condition (2021: 193, 195, 211).

[19]In line with notes 7 and 13, this commitment need not be interpreted as a voluntary (or even conscious) choice. It might result from automatic cognitive mechanisms that align with the standards of epistemic rationality. I thank an anonymous referee for pointing this out to me.

[20]Buchak rejects such a principle by denying that belief *after* the encounter of disagreement must solely depend on the balance of evidence *at that time*. She calls this assumption "SYNCHRONICITY" (2021: 196, 211, 213–214). This suggests that "we can be conciliationist about our *credences*, and steadfast about (some of) our *beliefs*" (Buchak 2021: 214).

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

Table 1. Initial disagreements concerning *p*.

α: "*you*"	ε: "*I*"	Agents
$B_\alpha(p)$	$B_\varepsilon(\neg p)$	D-DISAGREEMENT
$Pr_\alpha(p) = 0.9$	$Pr_\varepsilon(p) = 0.1$	C-DISAGREEMENT

Table 2. Utility assigned to our cognitive options toward *p*.

	State of affairs	
Cognitive options	p	¬p
$u(B_x p)$	1	0
$u(\neg B_x p \ \& \ \neg B_x \neg p)$	0.75	0.75
$u(B_x \neg p)$	0	1

receiving counterevidence. To continue with our previous example, we could make the following assumptions for our cognitive options toward *p*: for any of us, "believing *p*" has a positive utility of 1, if *p* is true – 0 if not; "believing ¬*p*" has a positive utility of 1, if *p* is false – 0 if not; "remaining agnostic towards *p*" has a positive utility of 0.75, regardless of *p*'s truth-value – since it avoids being mistaken, but is less informative (see Table 2).

Assuming that counterevidence *E* might be misleading or correctly lead to the current state of affairs, we can compare the utility of "Committing" to a course of action (e.g., "to believe *p*") and that of "Reevaluating" our initial course of action after receiving counterevidence *E* (see Table 3).

As Buchak observes, committing might maximize expected epistemic utility compared to reevaluating after receiving counterevidence when there are quasi-epistemic benefits,[21] β, that go beyond the utility of believing something true at time *t*. For example, consider the scenarios presented in Table 4.

Such benefits are common in various circumstances: (a) when *continuously believing* the truth has higher epistemic value than intermittent patterns of belief (e.g., for gaining knowledge while sustaining interpersonal relationships); (b) when a claim is a *high-level proposition* and is deeply embedded in the web of reasoning (e.g., it is at the core of a paradigm or worldview); or (c) when a claim is at the base of a *long-term intellectual project* (e.g., in science, morality, religion, and so forth) that requires actions over time whose rationality depends on believing that claim. It seems that some of these conditions are clearly met in ordinary cases of STEADFASTNESS.[22] Our example would provide the cases of *rational* STEADFASTNESS represented in Table 5.[23]

Insofar as these conditions account for the pervasiveness of *rational* STEADFASTNESS in many ordinary cases, the model might perform the EXPLANATION TASK therein. Also,

> The result here goes beyond the obvious point that if there are benefits to commitment, then it can be good to maintain a belief even if it wouldn't have been

[21]These benefits may include *avoiding costs* (with negative value) of breaking commitment (Buchak 2021: 204, n. 28).

[22]Thus, the model "can help us understand an important phenomenon arising in epistemic life: allegiance to, and break from, a tradition" (Buchak 2023: 741).

[23]These numerical values (as well as those in Tables 2 and 5) are just meant as an illustration and are drawn from Buchak (2021: 205, n. 29).

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

Table 3. Utility of "Committing" and "Reevaluating," upon receiving counterevidence E (Buchak, 2021: 201).

	State of affairs			
Responses to E's truth-value	p & E "correctly leading positive evidence"	p & ¬E "misleading negative evidence"	¬p & E "misleading positive evidence"	¬p & ¬E "correctly leading negative evidence"
u(Committing to $B_x p$*)*	1	1	0	0
u($B_x p$ *& Reevaluating in response to E)*	1	0.75	0	0.75

Table 4. Quasi-epistemic benefits of Committing.

α: *"you"*	ε: *"I"*	Agents
$\beta B_\alpha(p) = 0.9$		*Scenario 1:* High benefits for believing *p*, if *p* is true.
	$\beta B_\varepsilon(\neg p) = 0.2$ $\delta \neg B_\varepsilon(\neg p) = -0.2$	*Scenario 2:* Medium benefits for believing ¬*p*, if ¬*p* is true; and additional costs of relinquishing belief in ¬*p*, if ¬*p* is true.

Table 5. Examples of *rational* Steadfastness with D-Disagreement, but without C-Disagreement.

Thus far, we have assumed that α and ε were in D-Disagreement: $B_\alpha(p)$ and $B_\varepsilon(\neg p)$. Upon receiving counterevidence E, α and ε come to agree in their credences: $Pr_\alpha(p\|E) = 0.5$ and $Pr_\varepsilon(p\|E) = 0.5$.		
Scenario 1	*Scenario 2*	
$Pr_\alpha(p\&\neg E) = 0.85$ $Pr_\alpha(p\& E) = 0.05$ $Pr_\alpha(\neg p\&\neg E) = 0.05$ $Pr_\alpha(\neg p\& E) = 0.05$	$Pr_\varepsilon(p\&\neg E) = 0.05$ $Pr_\varepsilon(p\& E) = 0.05$ $Pr_\varepsilon(\neg p\&\neg E) = 0.85$ $Pr_\varepsilon(\neg p\& E) = 0.05$	Credence partitions
EU(*Committing to* $B_\alpha p$) = (0.9)(1.6)+(0.1)(0) = **1.44** EU($B_\alpha p$& *Reevaluating in response to E*) = (0.85)(1)+(0.05)(0.75)+(0.05)(0)+ (0.05)(0.75) = **0.925** EU(*Upholding commitment*) = (0.5)(1.6)+(0.5)(0) = **0.8** EU(*Abandoning commitment*) = **0.75**	EU(*Committing to* $B_\varepsilon(\neg p)$) = (0.9)(1.2)+(0.1)(0) = **1.08** EU($B_\varepsilon\neg p$& *Reevaluating in response to E*) = (0.85)(1)+(0.05)(0.75)+(0.05)(0)+ (0.05)(0.75) = 0.**925** EU(*Upholding commitment*) = (0.5)(1.2)+(0.5)(0) = **0.6** UE(*Abandoning commitment*) = 0.75 – 0.2 = **0.55**	Expected utility
In both scenarios, remaining faithful to their initial course of action maximizes the agent's expected utility.		

good to believe it in the absence of these benefits. What the model contributes is showing which things, exactly, we ought to be weighing against each other, and how the magnitude of the new probability matters. (Buchak 2021: 206)

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

Although by itself the model does not resolve this issue, it portrays situations in which STEADFASTNESS would not be an *undesirable* outcome. Thus, the OPTIMIZATION TASK should consider additional dimensions. If we manage to isolate the undesirable cases, what does the model tell us about how to avoid them? Roughly speaking, the model is compatible with an *argumentative resolution* of these cases of D-DISAGREEMENT. Since faith does not constrain the balance of evidence, agents might converge on their credence, resolving their C-DISAGREEMENT. Even if belief can be rationally maintained against counterevidence for some time, *this does not occur indefinitely or in the face of any amount of counterevidence:* the effect of benefits "anchoring" belief might be outweighed by the balance of evidence, especially over time.[24] Thus, commitment is eventually abandoned and D-DISAGREEMENT argumentatively resolved.

Some valuable clues can be gleaned from the model about what happens in such situations. To the extent that an agent's beliefs interact with her balance of evidence, upholding a commitment along with certain credences will be an epistemically *awkward position* for the agent: she believes something that it "would not be rational to believe if [she] approached the question anew. [. . . T]his tension is in fact felt when we respond to disagreement by maintaining our own beliefs, and so it is an advantage of the [model] that it can explain this tension" (Buchak 2021: 214). Facing STEADFASTNESS argumentatively will foster this kind of epistemic tension, perhaps for considerable periods of time. Additionally, relinquishing faith will differ from other kinds of doxastic change: once credence exceeds a certain threshold, the motivational effect of commitment fades away, and the agent experiences a sudden "doxastic openness to evidence." Due to the kinds of beliefs that are prone to be subject of commitment, this change will have large-scale repercussions in the epistemic life of a person: it will be akin to an episode of religious "conversion." Thus, according to the model, the *argumentative resolution* of disagreements in cases of STEADFASTNESS might *not always* be *desirable*, and it is probably *not easy*.

5. Rational escalation and polarization

Let us now consider cases of ESCALATION and POLARIZATION. Our description of these situations seems to involve a peculiar weighting of evidence in response to an argumentative exchange in the face of disagreement. Regardless of whether there is a D-DISAGREEMENT, the net effect of argumentation in these cases seems to be an increase in the "distance" between the agents' credences, thus *aggravating* a C-DISAGREEMENT. Unlike Lara Buchak's account of STEADFASTNESS, this response to disagreement seems epistemically irrational mainly at the level of the balance of evidence. We have assumed that in these cases the only relevant information that emerges from the argumentative exchange is the fact of the disagreement itself. So, why should a participant modify her credence in the *opposite direction* of that of her interlocutor, as occurs in ESCALATION?[25] Why would both participants *drift apart*, as they do in POLARIZATION? These phenomena seem at odds with the optimistic theorems of the model. As far as I know, there is no completely general explanation of these phenomena within the model. Some preliminary results, however, seem encouraging for the EXPLANATION TASK.

[24]As Buchak points out (2023: 744), the model offers an alternative rational explanation to "incommensurability" for this class of epistemic effects.

[25]I follow Pallavicini, Hallsson and Kappel (2021: 3–6), as well as Olsson (2021: 215), in calling this phenomenon "escalation" to distinguish it from polarization in a narrower sense. Both epistemic effects can be associated.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

Kenny Easwaran, Luke Fenton-Glynn, Christopher Hitchcock and Joel Velasco (2016) have investigated how an epistemically rational agent should respond to the discovery of the credence that other agents assign to a claim.[26] They aim to identify a computationally undemanding heuristic (one that is psychologically plausible) which emulates the epistemic effect of conditionalizing on the credence of other agents. In the process of devising such a heuristic, they demand that it allows agents to respond "synergestically" upon learning the prior credences of other agents, i.e., that the result of α's updating her credence toward p after learning ε's credence toward p must lie "outside of the interval" of their prior credences $[min\{\Pr_x(p)\}, max\{\Pr_y(p)\}]$. This property, which they call "synergy," sometimes in fact occurs by performing the (computationally more complicated) procedure of applying Conditionalization upon learning the credence of other agents and assessments of their reliability (Easwaran, Fenton-Glynn, Hitchcock and Velasco 2016: 36–37). Thus, although "in some cases of disagreement, the credences of one's peers are evidence against one's view; here, they are evidence for it" (Easwaran, Fenton-Glynn, Hitchcock and Velasco 2016: 10). Though the cases they analyze are not exhaustive of situations in which synergy occurs, they suggest that "such a response will be rational in any case in which, despite the fact that one's credence in $[p]$ differs from those of one's peers, the different credences nevertheless represent a kind of agreement about whether the evidence favors $[p]$ or $[\neg p]$" (Easwaran, Fenton-Glynn, Hitchcock and Velasco 2016: 10). In our terms, this is a case of ESCALATION: arguing about an initial C-DISAGREEMENT results in (at least) one of the participants intensifying her degree of confidence (modifying her credence to an interval outside the range of the original disagreement). Although under this description it seems peculiar, in some cases it would appear to be intuitively rational to reason synergistically. Consider the following situation:

> A doctor wonders what dosage of a drug to give a patient. She is initially very confident (e.g., with credence 0.97) that she should inject 5 ml. She consults with one of her colleagues. She discovers that her colleague is also very confident (e.g., with credence 0.92) that the correct dose is 5 ml. As a result, she increases her credence (e.g., up to 0.98) that the correct dose is 5 ml.[27]

In such cases, ESCALATION seems a rational response to the exchange. Roughly, this seems to be the case when, despite there being a C-DISAGREEMENT among participants considered as reliable sources of information, there is no D-DISAGREEMENT among them. However, this crude generalization is insufficient to account for many apparent cases of ESCALATION. A more careful examination of the model could provide better insights of the conditions that account for this phenomenon.

The most relevant results concerning POLARIZATION in the model come from experiments with computer simulations. As Erik Olsson (2021) discusses in a literature review, computer simulations of Bayesian agent networks (i.e., that satisfy the constraints of the model) offer a robust "argument for the rationality of polarization. Not only is polarization compatible with Bayesian updating; polarization is [...] omnipresent in social networks governed by such updating" (Olsson 2021: 221–222). These results come from performing experiments on a social communication network

[26]As they discuss in their paper, this issue is linked to —but different from— issues surrounding testimony, judgment aggregation, and peer disagreement (Easwaran, Fenton-Glynn, Hitchcock and Velasco 2016: 3–6).

[27]I have adapted the example, which the authors take from Christensen (Easwaran, Fenton-Glynn, Hitchcock and Velasco 2016: 10).

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

represented by means of a simulation model called "Laputa," written by Staffan Angere (Olsson 2011). The outcome of these experiments depends on feeding the model with some arbitrary values (e.g., for initial credences of agents and for sets of functions for assigning and updating of their estimates on trustworthiness of their interlocutors). By performing the experiment on multiple parameters for these values (285 groups), Josefine Pallavicini, Bjørn Hallsson and Klemens Kappel (2021) concluded that: "The very surprising result of this simulation is *all of the groups polarized to some degree*. In fact, most groups polarized to the maximum level. There were no conditions under which depolarization occurred" (Pallavicini, Hallsson and Kappel 2021: 22). There is no general explanation of why this happens for agents of the model.[28] An attractive suggestion, however, is that POLARIZATION is caused by

> differences in background beliefs and the effects those differences have, [. . . since] those effects essentially mean, in the case of credence updating, that evidence coming from sources believed to be trustworthy is taken at face value, whereas evidence coming from sources believed to be biased is taken as "evidence to the contrary." (Olsson 2021: 218)[29]

If we focus our attention on real situations that resemble POLARIZATION, these kinds of conditions often seem to be fulfilled. According to Thi Nguyen (2020), in addition to epistemic bubbles (where you do not hear the other side), there are echo chambers: social structures which *discredit* all outsiders. Their testimony is not just ignored but taken as evidence for the falsity of its content. In light of the robustness of computational results, and the fact that "depolarization" does not occur, perhaps an *argumentative resolution* of this kind of disagreement is not possible, or it is extremely unlikely.

An outcome like this might not be so bleak. After all, sustained disagreement might be accompanied by other epistemic benefits for groups cultivating it, such as "the revelation of a complex of interrelated assumptions, values, inferential standards, and cognitive goals [. . .] through critical interaction among researchers, [which is] disagreement under a more benign name" (Longino 2022: 179). Another moral that could be drawn from the model is that, even if disagreements that exhibit POLARIZATION cannot be *argumentatively* resolved, they are susceptible to other kinds of *epistemically rational interventions:*

> The route to undoing their influence is not through direct exposure to supposedly neutral facts and information; those sources have been preemptively undermined. It is to address the structures of discredit – to work to repair the broken trust between [members of such structures] and the outside social world. (Nguyen 2020: 159)

6. Final remarks

I explored the resources of the logical perspective, equipped with probabilistic tools, for the description and assessment of various epistemic effects of argumentation in response to disagreement. I used a simple model to illustrate how this approach allows us to

[28]Kevin Dorst (2023) has argued that ambiguous evidence (i.e., evidence that is hard to know how to interpret) might be predictably polarizing in disagreements.

[29]Something similar occurs in simulations performed by Cailin O'Connor and James Owen Weatherall (2018), which are based on a different model that considers evidence coming from sources with whom agents disagree as *more uncertain* than evidence coming from sources with whom they agree.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

identify very general sets of conditions under which an argumentative resolution of disagreements can be expected. Unsurprisingly, this seems possible under the constraints of epistemic rationality in "ideal" situations: when agents share initial credences and assessment of evidence, or when different initial credences are repeatedly assessed considering shared evidence. More controversially, drawing on several recent formal results, I suggested that the logical perspective is also capable of describing *sets of conditions under which argumentation does not contribute to the resolution of disagreements, or it even aggravates them*. Under these conditions, it might be epistemically rational (although, sometimes, undesirable) for disagreement to endure argumentation. By modeling these situations, it is possible to identify interventions to avoid these conditions; *the simplest or most effective interventions might not always be argumentative*.[30]

The highly idealized model I sketched has several limitations. It leaves out much of what is important in argumentative exchanges. Many details could be added to apply it to real situations. Nevertheless, even this outline offers recognizable diagnoses of what seems to be happening in various argumentative exchanges in the face of disagreement. It provides an outline of the factors that play a crucial role in these situations, and of the normative constraints that seem to guide them. This is compatible with other factors (e.g., choice of words, appeal to emotions, tone of voice, virtuous behavior, and so forth) playing a significant role in producing such epistemic effects in argumentative exchanges. It is not essential, however, to take them into account to explain *the possibility* of such effects: they might occur when an agent's epistemic rationality maintains its course. Requiring that "most" people – most of the time – be "good approximations to" epistemically rational agents should not strike one as a *flaw* of the model. This assumption makes it easier to deal with people; furthermore, it is explanatorily more fruitful than assuming that they are victims (or executioners) of rampant irrationality. However, the model employs a degree of precision that unrealistically represents epistemic capabilities and duties. At this point, it is important to remember that it is simply a model. It "can be thought of as an imaginary, simplified version of the part of the world being studied, one in which exact calculations are possible" (Gowers 2002: 4). Still, it can provide a very general understanding of why argumentation in the face of disagreement typically produces certain kinds of epistemic effects.

Acknowledgements. This paper is part of the project "Modelos epistémicos de argumentación científica," funded by UAM-I. Thanks to Mario Gensollen, Nancy Nuñez, Fabián Bernache, Fernanda Clavel, José de Teresa, Yolanda Torres, Juan Carlos Sánchez, Alejandra Tovilla, Christian Omar Méndez, José Carlos Rojas, Luis Gustavo Guzmán, Pablo Rangel, Marcos Iván Ramos, and Joaquín Galindo for insightful conversations on different aspects of the topic of this paper. Earlier versions were presented at the V Coloquio Internacional de Argumentación y Retórica (Mexico 2022) and as an invited lecture at the Universidad Autónoma de Guerrero (Chilpancingo 2024); I thank the audiences at both events for their questions and comments. I am also grateful to an anonymous referee for extensive and incredibly helpful comments. The remaining mistakes and misconceptions are mine.

References

Aikin S.F. and Casey J. (2022). 'Argumentation and the Problem of Agreement.' *Synthese* **200**(2), 1–23.

Aikin S.F. and Talisse R.B. (2019). *Why We Argue (and How We Should): A Guide to Political Disagreement*, 2nd ed. New York: Routledge.

[30]If arguing makes things worse, we do not have to consider our interlocutor to be a vegetable —as Aristotle (*Met*, IV, 1006a13–14)— to recognize that it is absurd to argue under such conditions.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

Aikin S.F. and Talisse R.B. (2020). *Political Argument in a Polarized Age: Reason and Democratic Life*. Cambridge: Polity Press.
Aumann R.J. (1976). Agreeing to Disagree. *The Annals of Statistics* **4**(6), 1236–39.
Bernache Maldonado F. (2025). 'Argumentation, Cooperation, and Disagreement.' *Informal Logic* **45**(1), 105–37.
Broncano-Berrocal F. and Carter J.A. (2021). 'Deliberation and Group Disagreement.' In F. Broncano-Berrocal and J.A. Carter (eds), *The Epistemology of Group Disagreement*, pp. 9–45. New York: Routledge.
Buchak L. (2013). *Risk and Rationality*. New York: Oxford University Press.
Buchak L. (2021). 'A Faithful Response to Disagreement.' *Philosophical Review* **130**(2), 191–226.
Buchak L. (2023). 'Faith and Traditions.' *Noûs* **57**(3), 740–59.
Davidson D. (1982). 'Paradoxes of Irrationality.' In *Problems of Rationality*, pp. 169–87. New York: Oxford University Press.
Diderot D. and D'Alembert J.-B.R. (eds). (1765). *Encyclopédie, ou Dictionnaire Raisonné des Sciences, des Artes et des Métiers, par un Société de Gens de Lettres*. Vol. **14**. Neufchastel: Samuel Faulche & Compagnie.
Dorst K. (2023). 'Rational Polarization.' *The Philosophical Review* **132**(3), 355–458.
Doury M. (2012). 'Preaching to the Converted: Why Argue when Everybody Agrees?' *Argumentation* **26**, 99–114.
Douven I. (2011). 'Relativism and Confirmation Theory.' In S.D. Hales (ed), *A Companion to Relativism*, pp. 242–63. New York: Blackwell.
Easwaran K., Fenton-Glynn L., Hitchcock C. and Velasco J.D. (2016). 'Updating on the Credences of Others: Disagreement, Agreement, and Synergy.' *Philosophers' Imprint* **16**(11), 1–39.
Eemeren F.H. (2010). *Strategic Maneuvering in Argumentative Discourse: Extending the Pragma-Dialectical Theory of Argumentation*. Amsterdam: John Benjamins Publishing Company.
Eemeren F.H. (2018). *Argumentation Theory. A Pragma-Dialectical Perspective*. Cham: Springer.
Frances B. and Matheson J. (2018). 'Disagreement.' In E.N. Zalta (ed), *The Stanford Encyclopedia of Philosophy*. Stanford: Stanford University.
Gaifman H. and Snir M. (1982). 'Probabilities over Rich Languages, Testing and Randomness.' *Journal of Symbolic Logic* **47**(3), 495–548.
Gensollen M. (2022). *Argumentación y Desacuerdo*. México: Universidad de Guadalajara.
Geoffrion L.-P. (1925). *Zigzags Autor de nos Parlers. Simple Notes*. Québec: Chez l'auteur.
Gilbert M. (1995). 'What is an Emotional Argument? Or Why do Argument Theorists Quarrel with their Mates?' In F.H. van Eemeren, R. Grootendorst, J.A. Blair and C.A. Willard (eds), *Analysis and Evaluation: Proceedings of the Third ISSA Conference on Argumentation*, pp. 3–12. Amsterdam: Sic Sat.
Goodwin J. (2007). 'Argument has no Fuction.' *Informal Logic* **27**(1), 69–90.
Gowers T. (2002). *Mathematics. A Very Short Introduction*. New York: Oxford University Press.
Jackson E.G. (2020). 'The Relationship Between Belief and Credence.' *Philosophy Compass* **15**(6), e12668.
Leitgeb H. (2017). *The Stability of Belief. How Rational Belief Coheres with Probability*. New York: Oxford University Press.
Longino H.E. (2022). 'What's Social About Social Epistemology?' *Journal of Philosophy* **119**(4), 169–95.
Nguyen C.T. (2020). 'Echo Chambers and Epistemic Bubbles.' *Episteme* **17**(2), 141–61.
O'Connor C. and Weatherall J.O. (2018). 'Scientific Polarization.' *European Journal for Philosophy of Science* **8**, 855–75.
Olsson E.J. (2011). 'A Simulation Approach to Veritistic Social Epistemology.' *Episteme* **8**(2), 127–43.
Olsson E.J. (2021). 'Why Bayesian Agents Polarize.' In F. Broncano-Berrocal and J.A. Carter (eds), *The Epistemology of Group Disagreement*, pp. 211–29. New York: Routledge.
Paglieri F. (2009). 'Ruinous Arguments: Escalation of Disagreement and the Dangers of Arguing.' *OSSA Conference Archive 121*. https://scholar.uwindsor.ca/ossaarchive/OSSA8/papersandcommentaries/121.
Pallavicini J.; Hallsson B. and Kappel K. (2021). 'Polarization in Groups of Bayesian Agents.' *Synthese* **198**, 1–55.
Rowbottom D.P. (2018). 'What is (Dis)Agreement?' *Philosophy and Phenomenological Research* **97**(1), 223–36.
Titelbaum M. (2019). 'Precise Credences.' In R. Pettigrew and J. Weisberg (eds), *The Open Handbook of Formal Epistemology*, pp. 1–55. Ontario: The PhilPapers Foundation.
Titelbaum M. (2022). *Fundamentals of Bayesian Epistemology, 2 vols*. New York: Oxford University Press.

https://doi.org/10.1017/epi.2025.10049 Published online by Cambridge University Press

Walton D. (1998). *The New Dialectic: Conversational Contexts of Argument*. Toronto: University of Toronto Press.
Weisberg J. (2020). 'Belief in Psyontology.' *Philosopher's Imprint* **20**(11), 1–27.

Marc Jiménez-Rolland is a professor of epistemology at the Autonomous Metropolitan University-Iztapalapa, Mexico. He is a member of the editorial board of the journal *Signos filosóficos*. He is also a member of Mexico's National System of Researchers and is affiliated to the Ibero-American Society of Argumentation (SIbA). His research focuses on formal approaches to epistemology and the philosophy of science.

Cite this article: Jiménez-Rolland M. (2025). "Arguing in the Face of Disagreement: Is It Worth It?" *Episteme* 1–17. https://doi.org/10.1017/epi.2025.10049