# The Veritistic Merit of Doxastic Conservatism in Belief Revision*

Gregor Betz (Karlsruhe)

May 2016

*Abstract*. There are different varieties of conservatism concerning belief formation and revision. We assesses the veritistic effects of a particular kind of conservatism commonly attributed to Quine: the so-called maxim of minimum mutiliation, which states that agents should give up as few beliefs as possible when facing recalcitrant evidence. Based on a formal bounded rationality model of belief revision, which parametrizes degree of conservatism, and corresponding multi-agent simulations, we eventually argue against doxastic conservatism from the vantage point of veritistic social epistemology.

*Keywords*. Doxastic conservatism; Epistemic conservatism; Verificationism; Maxim of minimum mutilation; Degree of justification; Belief revision; Bounded rationality; Veritistic value; Argumentation; Dialectical structure.

## Contents



*I'm grateful to Claus Beisbart and Georg Brun for extremely helpful feedback on an earlier version of this text.

# 1 Introduction

To which extent should epistemic agents be conservative and stick to their beliefs? That question seems to be particularly salient whenever novel evidence or arguments speak against an agent's previous views. Accordingly, we suggest to investigate the merit of being conservative by comparing two strategies of belief formation, which can preliminarily be put as follows:[1]

**Conservatism (maxim of minimum mutilation).** Agents strive to preserve their previous beliefs in belief change processes.[2] Verificationism (maxim of maximum warrant).
Agents strive to acquire (internally) well-justified beliefs in belief change processes.

Epistemic agents can adhere to these maxims to different degrees, and case studies in the history of science actually reveal that some scientists are cautious 'verificationists' while others are bold 'conservatives.' Consider for instance two major proponents of the so-called Great Devonian Controversy, which was sparked off by attempts to map the geology of Devonshire in South West England (cf. Rudwick 1985): Roderick Murchison and Adam Sedgwick initially agreed on key hypotheses in the stratigraphic debate; but as recalcitrant evidence came up, Sedgwick temporarily suspended judgment while Murchison maintained his far-reaching, yet—in view of the counter-arguments—less and less plausible position. Only as major changes in the background theory alleviated the anomalies did Sedgwick's and Murchsion's beliefs converge again. Obviously Murchison was doxastically more conservative than Sedgwick, who in turn put greater weight on having well-justified beliefs.

There seems to be a tension between the two maxims, as the example illustrates: An agent may stick to her beliefs although incoming evidence

[1] Foley (1983), in contrast, compares *conservatism* with *epistemic liberalism*, according to which the negation of every belief of mine is prima facie justified (for me). But *epistemic liberalism* seems to be outright implausible, and a corresponding comparison will not tell us anything about conservatism's epistemic merits.

[2] This maxim has been attributed to Quine, who alludes to the principle of conservative belief revision in (Quine 1951, 41, 43; 1992, 15); see also Rott (2000).

renders it less and less plausible. Or she may be prepared to sacrifice many of her beliefs in order to gain a well-justified position.[3]

Striking a balance between conservatism and verificationism, this paper assumes, is not just a matter of subjective preference. It is, in contrast, a question amenable to a normative epistemological analysis. More precisely, we will veritistically assess the two strategies by comparing them in view of the twofold goal of acquiring many (substantial) beliefs, and (mainly) true beliefs.[4]

Outline: Section 2, discerning different ways in which agents can be conservative in belief formation processes, clarifies the focus of this study and its relation to other work. Sections 3 and 4 develop a bounded rationality model of belief revision which parametrizes degree of doxastic conservatism. By means of multi-agent simulations, the results of which are visualized in a truth-content space (introduced in Section 5), we assess the truth-tracking-ability of agents who revise their beliefs in more or in less conservative ways (Section 6). While Section 6 argues that the simulation results, as seen from an individual agent, favor neither conservatism nor verificationism, Section 7 advances an argument against conservatism in view of collective strategies an epistemic community may employ to improve its members' beliefs.

## 2 Varieties of doxastic conservatism

We can discern, in the context of reasoning and belief formation, different types of conservatism, which allows us to clarify the assumptions of our study and its relation to similar work.

**Basic conservatism (doxastic revisionism).** A rational epistemic agent *revises* her previous beliefs in light of a new epistemic situation—rather than starting afresh and constructing her beliefs independently from previous beliefs and based on the novel epistemic situation (e.g., the modified total evidence) only.

Belief revision theory and Bayesianism *unisono* assume basic conservatism. Such revisionism is certainly a tenet of any epistemology for hu-

[3] Popper's critique of inductive logics (cf. Popper 2002, 269) is a pointed version of this trade-off: Given a probabilistic notion of justification, the maxim of maximum warrant urges agents to give up all substantial beliefs and maintain but tautologies.

[4] As an instance of *normative*, *veritistic*, *social* epistemology, this study may be subsumed under the research program set up by Goldman (1999).

man beings, because of our inability to re-process the total evidence whenever an epistemic situation changes.[5]

**Inferential conservatism (deductivism).** Whatever a rational epistemic agent comes to belief given novel evidence is *deductively entailed* by her previous beliefs and the evidence that triggered the belief revision.

Inferentially conservative agents don't perform ampliative inferences, they take no inductive risks. The corresponding belief revision is, in the words of Schurz (2011), "corrective," not "creative." Standard AGM-style belief revision theories (cf. Hansson 2009) assume inferential conservatism, since an agent's novel beliefs are entailed by some subset of what she previously believed (which is required to be consistent with the new evidence) plus the new evidence (which triggered the belief change).[6]

**Content conservatism (maxim of minimum mutilation).** A rational epistemic agent gives up as few of her beliefs as (logically) possible when revising, and especially when contracting her system of beliefs.

While Levi (1980, 61–62) explicitly holds this view, AGM-style belief revision theories don't presume the maxim of minimum mutilation, leaving it by and large to an agent's preferences to which extent she sticks to her beliefs or not (Rott 2000; 2001, 72–75). However, different (contested) postulates stipulate axiomatic lower boundaries for rational belief suspension (cf. Hansson 1993, 2013). Thus, *recovery* prescribes that agents, when coming to disbelieve p, will only shrink their belief system to the extent that, upon adding p again, they would re-acquire their original belief state. According to *maximality*, a belief state may only be adopted as a result of withdrawing p if there is no alternative belief state which contains the former while still excluding p.

**Atomistic conservatism (dogmatism).** An epistemic agent is *prima facie* justified in believing p simply in virtue of the fact that she believes p.[7]

[5]But are ideal (cognitively unbounded and epistemically perfect) agents necessarily revisionists, too? Why should they consider previous beliefs as something relevant at all? That seems to be an important question, which is rarely addressed. Without explicitly stating the radical alternative to doxastic revisionism, the literature discussing dogmatism (see below) however seems to approach this problem from time to time (e.g. Foley 1983, 170).

[6]More precisely, inferential conservatism follows, in terms of standard AGM notation, from the *inclusion* postulate for revision, $K * p \subseteq K + p$, and the *definition of expansion*, $K + p = \mathrm{Cn}(K \cup \{p\})$.

[7]Or, in other words, the fact that one believes p is a pro tanto reason for p.

Dogmatism concerns the justificatory status of individual statements (or beliefs)—and not the wholesale, dynamic process of belief system modification. It is chiefly considered by authors, such as, for example, Fumerton (2007), who are ultimately interested in spelling out a notion of knowledge which adequately addresses the skeptical challenge.[8] In that context, dogmatism has been misleadingly associated with Quine's maxim of minimum mutilation (Christensen 1994, 70). Foley (1983) and Christensen (1994) construct convincing counterexamples against atomistic conservatism;[9] but it is unclear whether those counterexamples bear on other versions of doxastic conservatism, too.

In the following investigation, we take revisionism and deductivism for granted, we remain silent about dogmatism, and we probe the veritistic merit of content conservatism.[10]

# 3 The statics of belief: A bounded rationality framework

In order to scrutinize content conservatism, we deploy a formal model of opinion dynamics which is developed in close analogy to theories of belief revision (specifically AGM-style theories).[11] Yet our model differs from standard belief revision theories in terms of its idealizations: Agents are not assumed to be logically omniscient.[12] The theory of dialectical

[8] An exception is Harman (1986, 32–42), who presents an argument for dogmatism in belief revision. But he assumes that an agent's beliefs to-be-revised are *not* embedded in a justificatory structure.

[9] For a discussion and defence of a more nuanced version of atomistic conservatism see McCain (2008). Vahid (2004) identifies further varieties of atomistic conservatism.

[10] Schurz (2011) reviews and proposes ways for merging ampliative modes of reasoning with belief revision theories. Such hybrid theories study agents who can learn more than what is deductively entailed by novel pieces of evidence and their previous beliefs (so-called creative revision). Although the model presented below makes belief revision dependent on non-deductive relations of justification (which come in degrees), it stays—for the time being—firmly within AGM-style theories inasmuch as it models *corrective* (not *creative*), *input-driven* (not *deliberate*) revision (see Schurz 2011 and references therein).

[11] We are going beyond AGM belief revision theories in specifying alternative choice functions (for different degrees of conservatism) and assessing them veritistically. While there is a literature which studies whether (AGM-style) belief-revision is necessarily truth-conducive (cf. the special issue edited by Kuipers and Schurz (2011)), we are not aware of any veritistic assessments of *conservative belief revision strategies* in particular.

[12] In this sense, we are proposing a *bounded rationality model* in the spirit of Simon (1982), which resembles other simple models of opinion dynamics and rational debate that take an intermediary position (and serve a mediating function) between theories of

structures provides the conceptual framework within which the model is spelled out (cf. Betz 2012a).[13]

The starting point of the model is a finite set of basic sentences, the *sentence pool*, which is closed under negation.[14] A *finite inferential net* specifies deductive inferential relations that hold between those sentences. The model principally allows for changes in the finite inferential net (discovery of novel arguments) and hence enables one to investigate "inferential drivers" of rational controversy (cf. Betz 2015). In this study, however, we consider opinion dynamics on *fixed* finite inferential nets, driven by evidence accumulation only.[15]

An agent's beliefs are represented by a *position*, i.e. by a truth-value assignment to (some) sentences in the pool. Complete positions assign truth-values to all sentences in the debate. A position $P$ *extends* a position $Q$ ($Q$ is a *sub-position* of $P$) if and only if all truth-value assignments by $P$ are also truth-value assignments of $Q$. Let the *size* of a position denote the number of truth-values it assigns. The *degree of content preservation* of position $Q$ relative to position $P$ (assuming $P$ extends $Q$) is defined as

$$\textsc{Cpr}(Q, P) := (N - k)/N,$$

where $N$ is the size of a complete position and $k$ is the difference in the size of positions $P$ and $Q$. We use the notion of content preservation in order to operationalize, below, the maxim of minimum mutilation.

A *complete position is dialectically coherent* (relative to a finite inferential net) if and only if it (i) assigns complementary truth-values to contradictory sentences and (ii) satisfies the deductive constraints imposed by the finite inferential net. An *incomplete position is dialectically coherent* if and

perfect epistemic agents and our actual epistemic practice, such as: the bounded confidence model (Hegselmann and Krause 2002), the theory of defeasible argumentation (Pollock 1987), explanatory networks (Thagard 1992), or the basic feature approach (Cevolani, Crupi, and Festa 2011; Cevolani 2015). On the need for such intermediaries compare also Spohn (2012, 258–62); Christensen (2007, 3–8), too, argues for pluralism regarding rationality ideals.

[13] That framework has originally been proposed as a systematization of argumentation-theoretic analyses of real debates.

[14] In other words, the sentence pool consists in pairs of contradictory sentences.

[15] Hansson (2016) proposes to modify AGM belief revision theory by restricting the model to *finite-based* belief states, i.e. belief states which are the deductive closure of some finite set of sentences. While this is certainly a move towards de-idealizing AGM theory, which puts Hansson's theory somewhere in the broad spectrum of bounded rationality models described in footnote 12, we go a step further than Hansson, who still assumes that agents possess potentially infinite inferential capacities, by letting agents adopt beliefs in view of a finite number of sentences *and* a finite number of inferential constraints only.

only if it can be extended to a dialectically coherent complete position.

A position is *dialectically closed* (relative to a finite inferential net) if and only if every singular extension of it (i.e., every single truth-value assignment to a so-far-unassigned sentence) yields a dialectically coherent position. The *dialectic* (or *inferential*) *closure* of a position $P$ is the the intersection of all complete and dialectically coherent positions which extend $P$.[16]

We require that rational agents hold dialectically closed and coherent positions.

In formal epistemology, an agent's belief state is usually represented as a logically closed set of sentences or, equivalently, as a set of possible worlds. According to that standard model, the set of belief states exhibits an algebraic structure. This is however not the case in the framework we employ here, which also explains why the belief revision model presented in Section 4 cannot be developed as a special case of standard belief revision theory, but only in analogy to AGM-style models.[17]

The *conditional degree of justification* of a position $P$ given position $Q$ is defined as the ratio of the number of complete and dialectically coherent positions that extend $P$ and $Q$ ($\sigma P,Q$) over the number of complete and dialectically coherent positions that extend $Q$ ($\sigma Q$):

$$\textsc{Doj}(P|Q) := \frac{\sigma_{P,Q}}{\sigma_Q}.$$

Degrees of justification, thus defined, capture essential (dynamic) properties of our pre-theoretic notion of strength of justification (Betz 2012b), enable one to reconstruct explanatory and hypothetico-deductive reasoning (Betz 2013a, 2013b), and seem to be truth-indicative (Betz 2015).[18] We use degrees of justification in order to operationalize, below, the maxim of maximum warrant.

[16]Note that the dialectic closure of a position is itself dialectically closed, and that every dialectically closed position is identical with its dialectic closure. Moreover, every complete and dialectically coherent positions is dialectically closed.

[17]More specifically, dialectically coherent and complete positions can be seen as the 'possible worlds' given a finite inferential net. But while there exists, for every dialectically closed and coherent position $P$, a unique set of possible worlds which extend $P$, not every set of possible worlds can be 'represented' by a dialectically closed and coherent position in this way. See also Hansson (2016, 13).

[18]Degrees of justification take the same value as the uniform probability measure which is defined over the power set of the set of all dialectically coherent & complete positions (~ possible worlds), which assigns every dialectically coherent & complete position the same probability.

# 4 Adding the dynamics: Belief revision on finite inferential nets

Let's assume that an agent holds a dialectically closed and dialectically coherent position $P$ given a fixed *finite inferential net*, and that $P$ expands the total evidence $E$ revealed so far, which we also model as a position. Next she learns, qua some item of evidence, that sentence $e$ from the sentence pool is true. Her way of processing this information depends on her previously held position $P$. If $e$ has been maintained by the agent before, her position remains unchanged. If expanding the agent's previous position by $e$ yields a dialectically coherent position, then the previous position is expanded accordingly and inferentially closed. The remaining cases, in which $e$ conflicts with the previous position $P$, are the interesting ones. Here, the agent proceeds as follows:

1. Contracting. Determine all sub-positions of $P$ that can be coherently expanded by $e$. Mark the *degree of content preservation*, $\textsc{Cpr}(Q,P)$, for each such contraction-option $Q$.

2. Expanding. Expand each sub-position identified in step 1 by $e$ and inferentially close the resulting position. Mark the *degree of justification* of the expanded and closed position, $Q$+$e$, relative to the total evidence revealed so far, including $e$ ($E$+$e$).

3. Optimizing. Choose a sub-position $Q$ so as to maximize the convex combination

$$\alpha \cdot \textsc{Cpr}(Q, P) + (1 - \alpha) \cdot \textsc{Doj}(Q + e|E + e),$$

and adopt $Q$+$e$ as novel position.

Agents hence chose a novel position in view of the content-loss they'd suffer and the degree of justification of the new position.[19] The parameter

[19]Ranking theory, too, closely connects degrees of justification and belief revision. Spohn (2012, 481) identifies degrees of justification with positive ranks (degrees of belief); a ranking function in turn represents a rational agent's subjective disposition to revise her beliefs given novel input. Hence, in ranking theory, it seems to be analytic that rational agents revise beliefs so as to obtain well-justified positions: rational agents adhere to the maxim of maximum warrant qua being rational. We follow Spohn in connecting degrees of justification and belief revision strategies. But in the model developed here, degrees of justification stem from objective properties of the state of a debate and are conceptually

$\alpha$ determines the relative weight of these two aspects in the belief revision process. The higher $\alpha$, the more conservative the agent.

In steps 1 and 2, the revision mechanism superficially replicates the basic structure of AGM-style belief revision (as expressed in the Levi identity). But these steps are, at closer inspection, merely preparatory computational steps, with the revision being ultimately performed in step 3. We really have a one-step choice mechanism, which effectively implements a monoselective choice function (in the sense of Hansson 2015) in step 3.[20]

# 5 A veritistic visualization of beliefs and their trajectories

A veritistic assessment gauges the epistemic value of belief states by comparing them to the truth (i.e., a presumably correct truth-value assignment). Inasmuch as such an analysis strives to *quantify* the veritistic value of belief states, it has to strike a balance between the antagonistic goals of acquiring many beliefs and mainly true beliefs, or, equivalently, between the epistemic value of true beliefs and the disvalue of false beliefs. (E.g., is a belief state with 60 correct and 40 incorrect beliefs veritistically better than a belief state with 30 correct and 10 incorrect beliefs?)

In this paper, we employ a novel method for analyzing the veritistic merits of belief states which is not committed to a specific weighing of correct against incorrect beliefs. Rather than expressing the veritistic value of a belief state in a single number, we map the belief state in a two-dimensional truth-content space.

Specifically, we plot belief states according to their content size (x-axis) and their *net truth content*, i.e. the difference between the number of correct and incorrect truth-value assignments they make (y-axis). Content size ranges from zero (maximal judgment suspension) to the size of the sentence pool (adoption of a complete position). A position's net truth content cannot be greater than its content size $s$ and not be smaller than $-s$. Thus the triangular shape of the truth-content space, as shown in Figure 1. The left corner of the triangle represents full judgment suspension.

independent from the agents' belief revision dispositions; accordingly, agents *may or may not* adhere to the maxim of maximum warrant when revising their beliefs – as represented through the parameter $\alpha$.

[20]Note that we don't implement AGM's select-and-intersect method. In addition, the revision procedure described in steps 1-3 obviously violates the postulates of recovery and maximality (see Section 2).

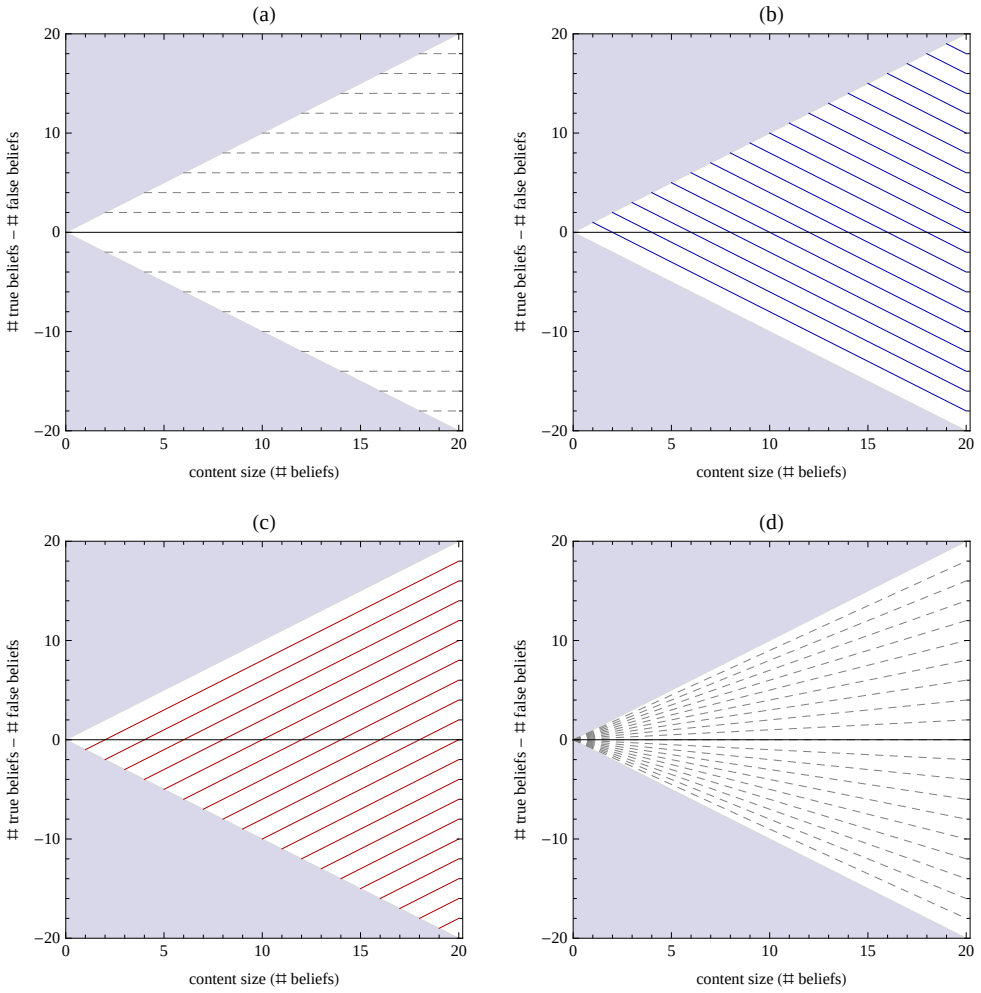


Figure 1: Isolines in the truth-content space. Belief states on the same horizontal line display the same difference between correct and incorrect truth-value assignments (a). Belief states with an equal number of true beliefs lie on downward sloping lines, parallel to the lower edge of the triangle (b), whereas belief states with an equal number of false beliefs lie on upward sloping lines, parallel to the upper edge of the triangle (c). Belief states which possess the same ratio of true beliefs lie on one and the same beam from the origin (d).

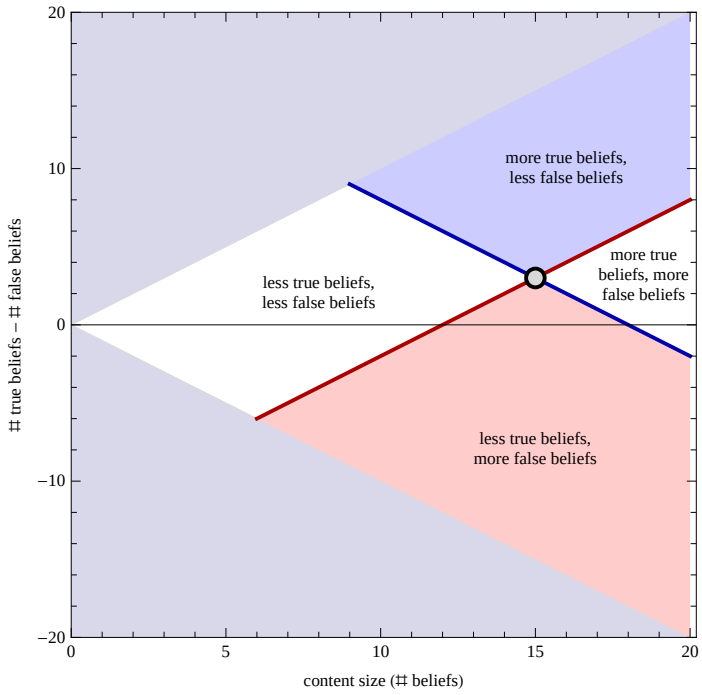


Figure 2: Division of truth-content space according to absolute numbers of true and false beliefs

The upper (lower) right corner marks the belief state which assigns correct (incorrect) truth-values to all sentences in the pool. Figure 1 illustrates the expressive power of the truth-content space by plotting various isolines.[21]

Belief states, when visualized in the truth-content space, can be assessed from different veritistic perspectives. Figure 2, for example, plots the illustrative position which correctly assigns 9 and incorrectly assigns 6 truth-values. It divides the truth-content space into three parts, containing (i) veritistically preferable positions ("more true beliefs, less false beliefs"), (ii) veritistically inferior positions ("less true beliefs, more false beliefs"), and (iii) veritistically ambiguous positions (both more or both less correct and incorrect truth-value assignments than the illustrative position). Figure 3, plotting the very same illustrative position, divides the truth-content space in a slightly different way (and hence visualizes an alternative veritistic perspective).

[21]Goldman's measure of a belief system's veritistic value (Goldman 1999, 89) corresponds to horizontal isolines (Figure 1a). Contrast measures of verisimilitude (Cevolani, Crupi, and Festa 2011, 188) induce, depending on their parameter, verisimilitude-isolines of different angles, ranging between those depicted in plots (a) and (b) in Figure 1.

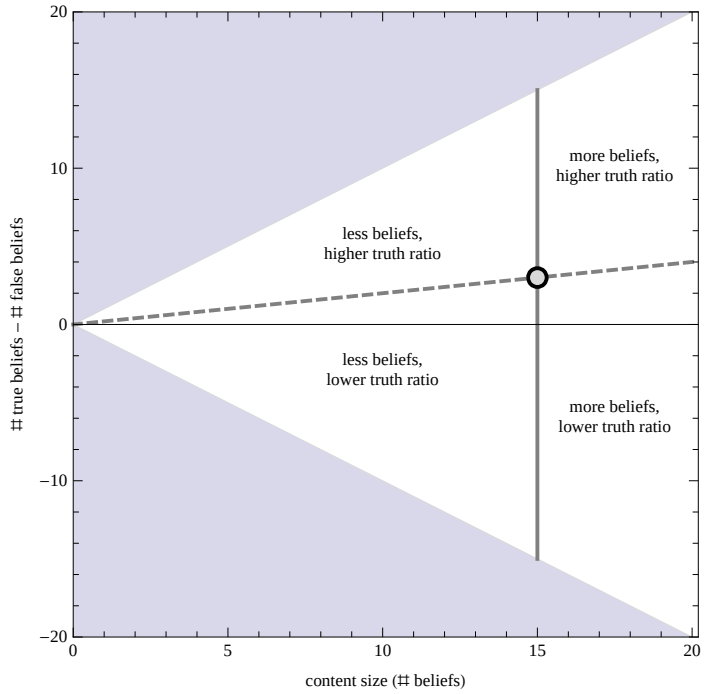


Figure 3: Division of truth-content space according to relative number of true beliefs and content size

# 6 Design and results of the multi-agent simulations

We simulate opinion dynamics on a pool of 2·20 sentences. Each individual simulation run is carried out on the background of a given finite inferential net. A randomly chosen dialectically coherent & complete position is designated as the truth **T**.[22] An evidence stream $\epsilon$ gradually reveals correct truth-value assignments (from **T**) with each step $t$. 20 agents initially hold randomly chosen dialectically coherent & complete positions[23] and revise their beliefs with each novel item of evidence ei as revealed by $\epsilon$ in accordance with the mechanism described in Section 4.

The individual opinion trajectories thus obtained will sensitively depend on various initial and boundary conditions, i.e. the structure of the finite inferential net, the correct truth-value assignments to sentences which figure in that net, the order by which correct truth-values are gradually revealed, and the initial positions the different agents hold. In order to identify the veritistic value of a belief revision strategy, we have to 'wash out' these dependencies by generating a large ensemble of opinion trajectories with varying initial and boundary conditions.

For a given $\alpha$-value we simulate, more specifically, opinion dynamics as described above on 1,000 different finite inferential nets.[24] Such an ensemble hence comprises 20,000 individual opinion trajectories, generated by equally conservative agents.

We vary $\alpha$ systematically from 0 to 1 (in steps of 0.1), and thus obtain 11 different ensembles. Our question is: Does the veritistic value of the trajectories vary systematically across the ensembles?

We consider, in the following, the simulated opinion trajectories' images in the truth-content space, in short: *truth-content trajectories*, or *trajectories* simpliciter.[25] In order to visualize the results of a single ensemble, we average over trajectories with the same initial net truth content, which

[22] We follow, more precisely, Betz (2015) in randomly choosing the truth from the set of complete position that are never rendered incoherent in the course of the debate, assuming that the "inferential density" (Betz 2012a, 44) cannot increase beyond 0.45.

[23] Where the sampling method guarantees that there is precisely one initial position for each net truth content value in $\{-20, -18, -16, \ldots, 18\}$.

[24] These finite inferential nets are chosen from previous debate simulations (Betz 2012a), they possess the same inferential density, namely $D$=0.3. Also, the findings of Betz (2015) suggest that our results will not essentially depend on this specific choice of $D$.

[25] Note that two different opinion trajectories, i.e. sequences of positions, may correspond to one and the same trajectory in the truth-content space.

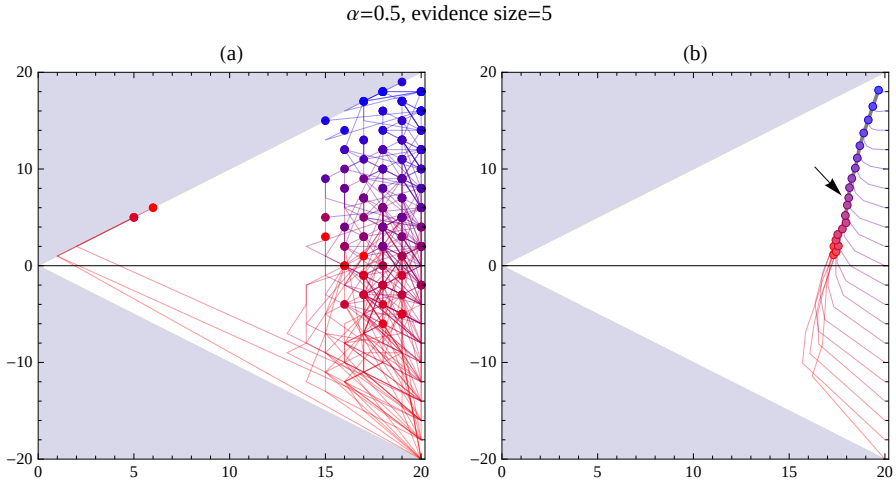


Figure 4: Illustrative individual and mean trajectories from the ensemble with $\alpha$=0.5. Plot (a): 10·20 randomly chosen individual truth-content trajectories (lines), 10 per possible initial point in the truth-content space, and corresponding points after five revision steps (disks). Plot (b): Mean trajectories that result from averaging all individual trajectories in the ensemble that depart from one and the same initial point (thin lines), corresponding mean points after five revision steps (disks), opinion frontier at step five (thick line connecting the disks); the arrow points at the 'median' point in the opinion frontier, whose trajectory departed from the point with zero net truth content.

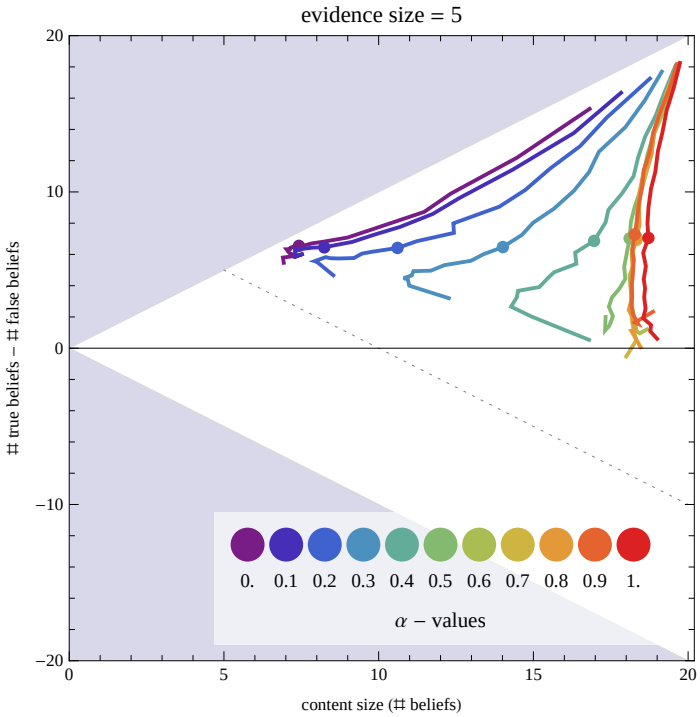


Figure 5: Opinion frontiers at $t$=5 for different $\alpha$-values. From left to right, agents are ever more conservative. In each opinion frontier, the disk marks the corresponding 'median' point (see also Figure 4). Median points divide an opinion frontier into an upper (initial net truth content >0) and a lower part (initial net truth content <0).

yields 20 *mean trajectories* per ensemble, each calculated on the basis 1,000 individual trajectories (see Figure 4). Individual trajectories are averaged point-wise: a mean trajectory's point at $t$ is the average of all the corresponding individual trajectories' points at $t$. The 20 points at step $t$ in an ensemble's *mean* trajectories are called the *opinion frontier* in the ensemble (at step $t$). An ensemble's opinion frontier (at step $t$) plots how, on aggregate, the positions of agents are distributed in that ensemble at step $t$.

Figure 5 plots the opinion frontiers (at step $t$=5) for eleven ensembles with different degrees of doxastic conservatism ($\alpha$-values). Various results can be read off Figure 5, depending on the veritistic perspective one adopts.[26]

[26]These findings are robust in regard of the specific time-step we consider, i.e., they don't depend on Figure 5 visualizing precisely time-step $t$=5.

Concerning the *net truth content*, we find:

- (A1) The net truth content of the median point in an opinion frontier doesn't depend on the degree of doxastic conservatism. (The disks are located on the same horizontal line.)
- (A2) The net truth content of positions in the upper half of an opinion frontier increases when agents become more conservative. (As we move from left to right, opinion frontiers reach out higher and higher.)
- (A3) The net truth content of positions in the lower half of an opinion frontier is reduced when agents become more conservative. (As we move from left to right, opinion frontiers reach out ever lower.)

Concerning the *number of true and false beliefs*, we find:

- (B1) An increase in doxastic conservatism drastically raises the number of false beliefs in positions in the entire opinion frontier. (Frontiers with low $\alpha$ are very close to the upper edge of the triangular truth-content space, but as we move from left to right, the opinion frontiers continually retreat from that edge; compare also Figure 1c.)
- (B2) In the same time, more conservative agents possess also more true beliefs. (As we move from left to right, opinion frontiers, including their median positions, move away from the lower edge and approach the upper corner of the triangular truth-content space; compare also Figure 1b.)

Concerning the *truth ratio and content size*, we find:

- (C1) The content size of all positions in an opinion frontier rises with increasing degree of doxastic conservatism. (As we move from left to right, content size decreases.)
- (C2) The ratio of true beliefs is reduced for nearly all positions in an opinion frontier as the degree of doxastic conservatism increases. (As we move from left to right, the opinion frontiers are tilted towards beams that correspond to ever lower truth ratios; compare also Figure 1d.)

In sum, the veritistic comparison of the two basic belief revision strategies yields highly *ambivalent* results (A2, B2 and C1 identify veritistic merits of conservatism; A3, B1 and C2 state veritistic merits of verificationism). Given that an individual agent doesn't know whether she initially holds a

position on the upper or the lower part of an opinion frontier, it seems to be a matter of preference (e.g., error aversion) which level of conservatism an individual agent should adopt in belief formation and revision.[27]

# 7 A final, social turn

Yet consider the community of (strategically homogenous) epistemic agents as a whole, which is spread over an opinion frontier (assuming sufficient plurality), rather than a single agent.[28] What about the following *social* strategy?

> (S) After sufficient evidence has been accumulated and agents have, individually and in isolation, revised their beliefs accordingly, every agent simply adopts the position of that agent in the community who holds the most beliefs.

We can safely assume that belief states are sufficiently transparent so that agents do know how many beliefs they, and their peers possess. Content size is an internal property of belief states and agents can locate themselves on the x-axis of the truth-content space. But agents cannot likewise see which of their beliefs are true, and which aren't. The truth is not accessible for the agents we model, so net truth content is an external property of belief states and agents are not able to locate themselves (nor others) on the y-axis of the truth-content space.

This said, back to the strategy (S): In communities with low $\alpha$, applying (S) has the effect that all agents will end up, on average,[29] at the upper (right) part of the opinion frontier, which is a veritistically extremely valuable position (cf. Figure 5)! Because of the virtually vertical slope of the opinion frontiers, the strategy (S) is however ineffective in communities with high $\alpha$: the agent with the maximum number of beliefs could be situated anywhere on those opinion frontiers, and could equally display above or below average net truth content.

[27] As a follow-up to this observation, one might want to assess the optimal level of conservatism ($\alpha$) *given* a specific level of error aversion (e.g., *given* a parameter $\varphi$ in a contrast measure of verisimilitude according to Cevolani, Crupi, and Festa (2011)).

[28] Since agents revise their beliefs independently of each other in our model, the opinion frontier of a strategically mixed community, i.e. a community of agents with different $\alpha$-values, is simply the weighted average of the opinion frontiers of the corresponding strategically homogenous communities, where the weights are given by the distribution of $\alpha$-values in the heterogeneous community.

[29] "On average" because the opinion frontiers are based on *mean* trajectories and ignore the variance within an ensemble.

Hence there is a strong argument for verificationism and against doxastic conservatism from a social epistemology perspective.

# 8 Conclusion

We have implemented two belief revision strategies in a bounded rationality framework: doxastic conservatism and verificationism. A veritistic assessment of these strategies, by means of multi-agent simulations, yields ambivalent results: neither do conservatively updating agents acquire true beliefs at greater pace than less conservative agents, nor are they clearly outperformed. From an individualistic point of view, it seems to be a matter of preference whether agents should be conservative or not. Still, the collective profile of beliefs (the opinion frontier) varies systematically between homogeneously conservative and homogeneously verificationist epistemic communities. Verificationist communities can use content size as a proxy to collectively improve their beliefs (namely inasmuch as every agent adopts the content-maximal position held by someone in the community); a similar social strategy is not available for homogeneously conservative communities. Individual caution in belief formation collectively pays off. Hence we have a strong argument against doxastic conservatism from a social epistemology point of view.

This paper's conclusions are tightly bound to the specific way it models conservative belief revision. It is an open question whether implementations of doxastic conservatism in alternative models yield similar results. Moreover, the model employed here can also be modified in relevant ways, which gives rise to a couple of follow-up questions, e.g.:

- We have assumed that the evidence stream only reveals correct truth-value assignments. What if the evidence is fallible and partly erroneous? Does this affect the veritistic assessment of conservatism and verificationism?

- We have assumed that agents initially hold complete positions. What if they start from incomplete positions? Is one strategy more favorable under these conditions?

- We have assumed that, for each individual agent, the degree of conservatism ($\alpha$) is fix. What if $\alpha$ may change in the course of an opinion formation process (e.g., $\alpha$ could be a function of previous opinion trajectories and argumentation strategies employed)? Do such

context-sensitive hybrid strategies outperform the simple revision procedures studied here?

- We have only investigated content conservatism. Can one model and veritistically assess other types of conservatism, for example deductivism (by comparing it with ampliative expansion), in this framework, too?

## References


Betz, Gregor. 2012a. *Debate Dynamics: How Controversy Improves Our Beliefs*. Synthese Library. Dordrecht: Springer.

———. 2012b. "On Degrees of Justification." *Erkenntnis* 77 (2): 237–72.

———. 2013a. "Justifying Inference to the Best Explanation as a Practical Meta-Syllogism on Dialectical Structures." *Synthese* 190 (16): 3553–78.

———. 2013b. "Revamping Hypothetico-Deductivism: A Dialectic Account of Confirmation." *Erkenntnis* 78 (5): 991–1009.

———. 2015. "Truth in Evidence and Truth in Arguments Without Logical Omniscience." *The British Journal for the Philosophy of Science* online first.

Cevolani, Gustavo. 2015. "Social Epistemology, Debate Dynamics, and Truth Approximation." In *Proceedings of the Fourth Conference of the European Philosophy of Science Association (Epsa13)*, edited by Uskali Mäki, Stéphanie Ruphy, Gerhard Schurz, and Ioannis Votsis. Springer.

Cevolani, Gustavo, Vincenzo Crupi, and Roberto Festa. 2011. "Verisimilitude and Belief Change for Conjunctive Theories." *Erkenntnis* 75 (2): 182–202.

Christensen, David. 1994. "Conservatism in Epistemology." *Noûs* 28 (1): 69–89.

———. 2007. "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." In *Oxford Studies in Epistemology*, edited by Tamar S. Gendler and John Hawthorne, 2:3–31. Oxford: Oxford University Press.

Foley, Richard. 1983. "Epistemic Conservatism." *Philosophical Studies* 43 (2): 165–82.

Fumerton, Richard. 2007. "Epistemic Conservatism: Theft or Honest Toil?" In *Oxford Studies in Epistemology*, edited by Tamar S. Gendler and John Hawthorne, 2:63–86. Oxford: Oxford University Press.

Goldman, Alvin I. 1999. *Knowledge in a Social World*. Oxford, New York: Oxford University Press.

Hansson, Sven Ove. 1993. "Theory Contraction and Base Contraction Unified." *The Journal of Symbolic Logic* 58 (2): 602–25.

———. 2009. "Logic of Belief Revision." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2009. http://plato.stanford.edu/archives/spr2009/entries/logic-belief-revision/.

———. 2013. "Repertoire Contraction." *Journal of Logic, Language and Information* 22 (1): 1–21.

———. 2015. "A Monoselective Presentation of AGM Revision." *Studia Logica* 103 (5): 1019–33.

———. 2016. *Choosing What to Belief*. Manuscript.

Harman, Gilbert. 1986. *Change in View: Principles of Reasoning*. Cambridge, Mass.: MIT Press.

Hegselmann, Rainer, and Ulrich Krause. 2002. "Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation." *Journal of Artificial Societies and Social Simulation* 5 (3): –.

Kuipers, Theo, and Gerhard Schurz. 2011. "Introduction and Overview: Special Issue Belief Revision Aiming at Truth Approximation." *Erkenntnis* 75 (2): 151–63.

Levi, Isaac. 1980. *The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability and Chance*. Cambridge, Mass.: MIT Press.

McCain, Kevin. 2008. "The Virtues of Epistemic Conservatism." *Synthese* 164 (2): 185–200.

Pollock, John L. 1987. "Defeasible Reasoning." *Cognitive Science* 11 (4): 481–518.

Popper, Karl R. 2002. *The Logic of Scientific Discovery*. Routledge Classics. London; New York: Routledge.

Quine, Willard Van Orman. 1951. "Two Dogmas of Empiricism." *The Philosophical Review* 60: 20–43.

———. 1992. *Pursuit of Truth*. 2nd ed. Cambridge, Mass.: Harvard University Press.

Rott, Hans. 2000. "Two Dogmas of Belief Revision." *Journal of Philosophy* 97 (9): 503–22.

———. 2001. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*. Oxford: Oxford University Press.

Rudwick, Martin J. S. 1985. *The Great Devonian Controversy: The Shaping of Scientific Knowledge Among Gentlemanly Specialists*. Science and Its Conceptual Foundations. Chicago: University of Chicago Press.

Schurz, Gerhard. 2011. "Abductive Belief Revision in Science." In *Belief Revision Meets Philosophy of Science*, edited by Erik J. Olson and Sebastian Enqvist, 21:77–104. Logic, Epistemology, and the Unity of Science. Dordrecht: Springer.

Simon, Herbert A. 1982. *Models of Bounded Rationality*. Cambridge, Mass.: MIT Press.

Spohn, Wolfgang. 2012. *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford: Oxford University Press.

Thagard, Paul. 1992. *Conceptual Revolutions*. Princeton: Princeton University Press.

Vahid, Hamid. 2004. "Varieties of Epistemic Conservatism." *Synthese* 141 (1): 97–122.