

# Evidential Reconfiguration as Bayesian Confirmation For Dark Matter

How Known Data Becomes New Evidence

Simon Allzén<sup>1,2,3</sup>

<sup>1</sup>Department of Philosophy, Stockholm University

<sup>2</sup>Institute of Physics, University of Amsterdam

<sup>3</sup>Vossius Center for the History of Humanities and Sciences

January 2026

The 1974 papers by Ostriker et al. [1974] and Einasto et al. [1974] are considered by many to be pivotal in establishing the epistemic foundations for the dark matter hypothesis. From a theory confirmation point of view, the circumstances surrounding this pivot are difficult to reconcile with common approaches to epistemic support. First, the papers did not introduce any new observations. Second, they synthesized existing data from two separate contexts to construct a hypothesis under which the joint data became evidentially relevant. Third, this synthesis was motivated in part by non-empirical reasons. The situation excludes both *temporal novelty* and *use novelty* because already known data was used in the construction of the hypothesis. Yet, the papers are widely regarded as epistemically transformative. I argue that a Bayesian can model the epistemic significance of the 1974 papers without concession. By recognizing how the papers reconfigured the existing data to bear on a missing-mass hypothesis, a novel epistemic aspect emerges. By introducing a shared halo parameterization, they made the previously disjoint data mutually constrained, thereby changing their evidential role. I develop this idea through two concepts — *evidential reconfiguration* and *structural novelty* — leveraged through Myrvold's Bayesian account of unification. The result makes Bayesianism faithful to the inferential practices in this significant part of scientific history, explains how the 1974 papers strengthened the evidential case for dark matter, and expands the Bayesian toolbox with a way to treat *novel structure* as epistemologically salient.

## 1 Introduction

The nature of  $\approx 80\%$  of the matter content in the universe remains unknown. Its presence is inferred from gravitational anomalies across a range of cosmic scales, where the observed motions and structures cannot be accounted for by ordinary matter alone. This putative mass component is known as dark matter, and it is a core posit of the standard model of cosmology,

the  $\Lambda$ CDM (concordance) model.<sup>1</sup> Dark matter is invoked to explain the motions of galaxies, the dynamics of galaxy clusters, and the overall distribution of matter in the universe. To date, it has eluded non-gravitational detection. As a result, its nature remains radically underdetermined, making dark matter one of the major open problems in contemporary physics. Yet dark matter is deeply embedded in  $\Lambda$ CDM and widely accepted by cosmologists despite lacking the confirmational profile one might expect under such underdetermination. This makes dark matter an interesting case for philosophers of science.<sup>2</sup>

One central puzzle, especially for a Bayesian, is the gap between dark matter's level of evidence and its level of acceptance. How, despite the absence of non-gravitational detection, did dark matter become so widely accepted and established as a core posit in cosmology? Textbook accounts suggest a largely linear progression of accumulating evidence. That picture makes confirmational sense only if one assumes that each observation was treated as direct evidence for an already established hypothesis. In practice, however, the relevant observations were initially often regarded as isolated, context-specific anomalies, not systematically integrated within a clearly defined framework or working hypothesis. From the perspective of theory confirmation and theory assessment, it is important to understand how and why such anomalies came to count as *evidence* for dark matter. Answering this question has the potential to reveal key insights that may generalize, providing a better understanding on how existing observational discrepancies can transform into coherent evidence for an ontological posit, and the conditions under which such changes are warranted.

This paper offers a philosophical analysis and Bayesian treatment of one such transformation in the case of dark matter. I build on the historical work by de Swart et al. [2017] and de Swart [2020], who identify two seminal papers from 1974 as "landmark papers" in the acceptance and establishment of the dark matter hypothesis. Accepting the historical thesis prompts a corresponding philosophical task of explaining *how* they could have played that role, given that neither paper contained any new empirical data. Beyond resolving a historical puzzle, the paper also aims to improve on the Bayesian problem of old evidence in confirmation theory. I contend that the distinctive epistemic contribution of the 1974 papers resides in a particular reconfiguration of pre-existing observations. Phenomena previously treated as local and disconnected could be taken to mutually support a single missing-mass hypothesis. The shift is best understood as an instance of *evidential reconfiguration*, a change in what is treated as background, what is treated as evidence, and which dependence relations are taken to be warranted within a modeling framework. As we will see, part of what motivated this reconfiguration was non-empirical in character. Cosmologists treated certain non-empirical commitments as constraints on theory assessment, a fact that matters normatively. When are such constraints epistemically defensible rather than merely heuristic? The next subsection sketches the two parts of the paper that answer that question and reconstruct the 1974 unificatory step.

<sup>1</sup> $\Lambda$  denotes dark energy, and CDM stands for cold dark matter, where 'cold' refers to the negligible free-streaming length of the dark matter component.

<sup>2</sup>See, e.g., Vanderburgh [2003, 2005, 2014b,a], Sus [2014], Jacquart [2021], Merritt [2021a,b], De Baerdemaeker and Boyd [2020], De Baerdemaeker [2021], De Baerdemaeker and Dawid [2022], Antoniou [2023, 2025], Duerr and Wolf [2023], Wolf and Read [2025], Allzén [2021, 2024], Martens [2022], Martens et al. [2022], and Vaynberg [2024].

## 1.1 Overview

From the period highlighted by de Swart et al. [2017] and de Swart [2020], I extract two philosophically significant features of the scientific reasoning at the time. These features structure the two main parts of the paper.

The first concerns theory assessment. In the mid-1970s, astronomers and cosmologists largely agreed on *what* the relevant observations were, yet they differed in how strongly those results should be treated as evidential support for missing mass. Cosmologists, I argue, were more receptive in part because they treated certain background commitments, including non-empirical ones such as a preference for a closed universe, as *constraints* on theory assessment. In a Bayesian reconstruction, this shows up not only as differences in priors, but also, and more importantly, as differences in the background assumptions that determine which connections between claims are admissible in the first place. The point is not that such assumptions provided additional evidence for missing mass, but that they shaped what counted as a pressing problem and how existing results could reasonably be integrated to address it.<sup>3</sup> The upshot is a modest permissiveness about theory assessment. Distinct evaluations of the degree of support for a hypothesis, given the same empirical data, can be *rationally permissible* when they arise from different sets of defensible constraints, non-empirical or not. A central task, therefore, is to clarify what makes such constraints normatively defensible rather than merely heuristic. I suggest that a background constraint is normatively defensible when it:

- a. is independently motivated within an established modeling context;
- b. has an established history of combining separate phenomena and models into a globally coherent framework that is empirically supported; and
- c. is treated as defeasible rather than as dogma.<sup>4</sup>

The second feature concerns unification. Since the epistemic status of unification is contested in philosophy, the question is not whether the unificatory step in the 1974 papers was *influential*, which is a sociological question, but whether it was *epistemic*. To address this, I draw on Myrvold [2003, 2017], who distinguish between unification as the postulation of a common origin and unification as an increase in mutual information between previously separate evidential claims.<sup>5</sup> The 1974 studies introduced a halo framework in which spiral galaxies are posited to have extended dark-matter halos governed by shared parameters, for example a characteristic mass discrepancy relative to the luminous component. This modeling addition did more than explain disparate observations through a common mechanism. *It reconfigured the inferential dynamics between them.* The anomalous dynamics of galaxy clusters and the flatness of galaxy rotation curves became related such that each was informative about the other under

<sup>3</sup>This is structurally similar to other accounts of indirect or meta-empirical support, where what shifts is not the first-order empirical evidence but how a successful framework constrains the space of viable hypotheses. See, for example, Dawid [2013, 2019] and Dawid et al. [2015].

<sup>4</sup>A constraint is non-normative when it functions only as a preference for a research direction without satisfying (a)–(c).

<sup>5</sup>See Castellani et al. [2025] for a recent and similar approach which proposes an expansion of the Bayesian use-case for unification beyond Myrvold's restricted version.

the halo hypothesis.<sup>6</sup> On Myrvold’s account, the epistemic source of confirmation lies in the reciprocal dependence that emerges: a mutual-information gain. Normatively, this provides a principled way to distinguish justified evidential reconfiguration, which generates genuine informational support, from unification that is merely heuristic or programmatic.

Finally, the paper’s contributions can be stated concisely. First, it shows how non-empirical commitments, for example  $\Omega \approx 1$ , the cosmological principle, and Machian considerations, can be represented in a Bayesian framework as differences in priors and background assumptions, in a way that respects cross-community variation without treating either side as biased. Second, it offers a detailed, historically grounded application of Myrvold’s mutual-information account of unification to the 1974 synthesis. The analysis clarifies how confirmation of the dark matter hypothesis can increase through the unification of cluster dynamics and galaxy rotation curves. Third, it refines the old-evidence discussion by distinguishing temporal and use novelty from *structural novelty*. Even when the underlying observations are already known and used in the construction of the hypothesis, changes in what evidential role they play imposed by structural dependencies, here characterized as *evidential reconfiguration*, can generate confirmational support in a principled Bayesian sense.

## 2 The standard historical narrative

Textbook and popular narratives of dark matter often begin with Zwicky’s work on the Coma Cluster in the 1930s, and then proceed by selecting a small number of later breakthroughs that are presented as progressively establishing the hypothesis. The canonical starting point is Zwicky’s application of the virial theorem to the Coma Cluster, which led him to conclude that the amount of visible matter, inferred from luminosities, was far too small to account for the observed galaxy velocities in the cluster. Given its luminous mass, the cluster should not have remained in the bound state it was observed to be. In narrative retrospect, this is often treated as the earliest observational basis for non-luminous mass.<sup>7</sup>

The next standard step is to emphasize the work of Rubin and Ford Jr [1970] on rotation curves in the 1970s, which resulted in the observation that the rotational velocity of matter in spiral galaxies does not decrease with distance from the galactic center, deviating from Keplerian predictions. This measurement was soon followed by observations by Roberts and Rots [1973], Bosma [1978], and others, which strengthened the case that many spiral galaxies exhibit approximately flat rotation curves. Stars far from the galactic center orbit nearly as quickly as those closer in, in tension with the decline expected from visible matter alone under simple dynamical assumptions. Textbook presentations often treat these results as decisive vindication of dark matter, despite the fact that Rubin herself was cautious about any causal interpretation of her measurements [Rubin, 2004], and that many astronomers initially explored explanations that did not require positing a new exotic mass component.

<sup>6</sup>The independence assumptions used to represent the pre-1974 situation are meant as an idealization of disciplinary practice. The claim is not that no link was *conceivable*, but that no stable, widely endorsed link was treated as evidentially significant across domains.

<sup>7</sup>At the time, Zwicky [1933, 1937] did use the term “*dunkle Materie*” for unseen mass. However, its conceptual content bears little resemblance to the current concept of dark matter (see Allzén [2024] for a historical overview).

The story is then completed by incorporating later developments such as gravitational lensing, large-scale structure formation, and the Bullet Cluster observation [Clowe et al., 2006]. These are presented as extensions of the Zwicky–Rubin sequence, progressively strengthening the hypothesis until it became a central posit of contemporary cosmology. The structure of this narrative is implicitly, and perhaps accidentally, Bayesian in character. It implies the early existence of a hypothesis followed by a linear series of evidence suitable for Bayesian conditionalizing. I do not claim that historical presentations of dark matter, typically structured in roughly this way, are illegitimate. They serve important pedagogical purposes for their audiences. They rely on hindsight to select which historical parts are worth emphasizing, namely those that connect to the currently held concept and explanatory role of dark matter. The downside is that such histories obscure an epistemically prior question. When, and by which means, did particular results come to function as evidence for missing mass rather than as local anomalies or modeling artifacts?

The mid-1970s is, as de Swart et al. [2017] and de Swart [2020] show, a particularly illuminating episode in this regard. It brings into view how the dark matter hypothesis became established without the textbook appearance of a linear accumulation of evidence. The aim here is to complement the historical claim *that* the mid-1970s mattered with an epistemic account of *why* it mattered. The empirical data did not change in 1974. What changed was which assumptions were held fixed, and what the existing results were allowed to count as evidence for.

### 3 Constraints and the mass deficit

To understand why the dark matter hypothesis gained traction in both cosmology and astronomy, it is necessary to reconstruct the cosmological background that shaped theory assessment in the decades leading up to the 1970s. In the early and mid-20th century, cosmology was often guided more by theoretical considerations than by the observation-driven practices of astronomy, and debates frequently appealed to methodological and philosophical commitments [de Swart, 2020]. A canonical example is the early acceptance of the cosmological principle, formulated by Milne [1932], which asserts that the universe appears homogeneous and isotropic on sufficiently large scales. Today, this principle has substantial empirical support, but in the 1930s it was adopted in part on philosophical grounds. Coupling the cosmological principle with general relativity led to relativistic cosmology, which by the 1960s had become the predominant framework, drawing on theoretical work by Alexander Friedmann and Georges Lemaître, and observational work by Edwin Hubble [de Swart, 2020].<sup>8</sup> By the 1950s and 1960s, two non-empirical commitments in particular structured background assumptions in cosmology. The first was the *cosmological principle*, the assumption of the large-scale homogeneity and isotropy of the universe. The second was a *Machian inheritance*, according to which local inertia is determined by the total mass of the universe. In relativistic cosmology this was often taken to favor a closed, finite universe.

Taken together, these commitments encouraged a style of reasoning in which empirical

---

<sup>8</sup>See Bondi and Gold [1948], who argued that one cannot assume *a priori* that general relativity applies globally without empirical justification, an objection to the early philosophically motivated cosmological assumptions.

results were interpreted relative to a set of cosmological models that were treated as viable in part for non-empirical reasons. de Swart [2020] documents this preference explicitly. Nearing the end of the 1960s, a closed universe remained a widely favored model for many cosmologists and relativists:

Nearing the end of the 1960s, a closed universe was still a much-preferred model for many cosmologists and relativists. “Philosophically, there might be a preference”, Wolfgang Rindler wrote in 1967, “the choice  $k = 1$  a positively curved universe might appear desirable. It implies closed space sections that would, in some sense, validate Mach’s principle according to which the totality of matter in the universe and nothing else determines the local inertial frames” (Rindler 1967: 29–30, emphasis in original). [...] The implicit preference for a closed universe was also expressed in observational studies. “One would particularly like to know whether there is enough mass to close the universe”, Princeton physicists Peebles and Partridge wrote in 1967, in a piece on estimating the mass density of the universe (Peebles and Partridge 1967: 713). [de Swart, 2020, p. 18]

The key point to stress for present purposes is that such non-empirical commitments were not considered to be salient *qua* evidence, but as constraints on theoretical modeling. They shaped which questions and problems were taken as pressing and which avenues of integration were worth pursuing in order to resolve them. Weinberg captures this in a characteristic way:

If one tentatively accepts the result that  $q_0$  is of order unity [ $\Omega \gtrsim 1$ ], then one is forced to the conclusion that a mass density of about  $2 \times 10^{-29} \text{ g/cm}^3$  must be found somewhere outside the normal galaxies. But where? [Weinberg, 1972, p. 478]

Commitments of this kind were revisable in principle, but taking them seriously made certain problems appear more pressing, as witnessed in both Peebles and Partridge’s “whether” and in Weinberg’s “where”. Peebles and Partridge ask whether there is enough mass to close the universe, while Weinberg presses the follow-up question of where that mass must be found. By the early 1970s, near-critical cosmological models functioned for some authors as a defensible target and for others as a background commitment with independent motivations. Either way, they helped create a context in which missing mass could be pursued as a candidate astronomical solution to a cosmological problem, even before any decisive new observation in its favor.

Against this background, two research groups independently arrived at similar proposals in 1974. Ostriker et al. [1974] and Einasto et al. [1974] looked to galactic astronomy for a common dynamical account of missing mass on galactic and cluster scales, with downstream implications for the cosmic mass budget. The emphasis, however, was not symmetric: Ostriker et al. [1974] foreground cosmological density and closure considerations from the outset, while Einasto et al. [1974] foreground the cluster virial mass discrepancy and treat massive coronas as a way to reduce it [Einasto et al., 1974, 310]. They did so by drawing on two kinds of phenomena that were already well known, but typically treated as separate research problems. The separation was not primarily chronological but disciplinary. The phenomena were investigated in different contexts, with different methods and uncertainties, and there was no stable, widely endorsed view that they should bear on the same hypothesis in tandem.

1. Galaxy dynamics.

Thanks to radio observations of the 21-cm hydrogen line, astronomers could measure rotation speeds of spiral galaxies out to radii well beyond the visible stellar disk. By the early 1970s, abundant rotation-curve data were available. Expectations based on Keplerian dynamics were that rotation speeds should decrease with radius once beyond most of the galaxy's light (see fig. 1, reconstructed from the original plot in Roberts and Rots [1973]). Empirically, however, many spiral galaxies showed approximately flat rotation curves. This behavior is consistent with additional gravitating mass at large radii, but it was not automatically decisive for any single ontological conclusion. In practice, these results were often treated as local anomalies, whose significance depended on auxiliary assumptions in astronomy about tracer populations, mass-to-light ratios, gas distributions, and dynamical modeling.

2. Galaxy cluster dynamics.

Here too, the data were long-standing. Since Zwicky's time, various studies (e.g., Shapiro [1971]) measured galaxy velocities and mass-to-light ratios in clusters and repeatedly found that clusters appeared to contain too little visible mass to be gravitationally bound. Again, the phenomenon suggested a mass discrepancy, but its evidential force was mediated by auxiliary assumptions (e.g., equilibrium, virialization, and the representativeness of the observed velocities). It was not uniformly treated as direct support for a single missing-mass hypothesis across the community.

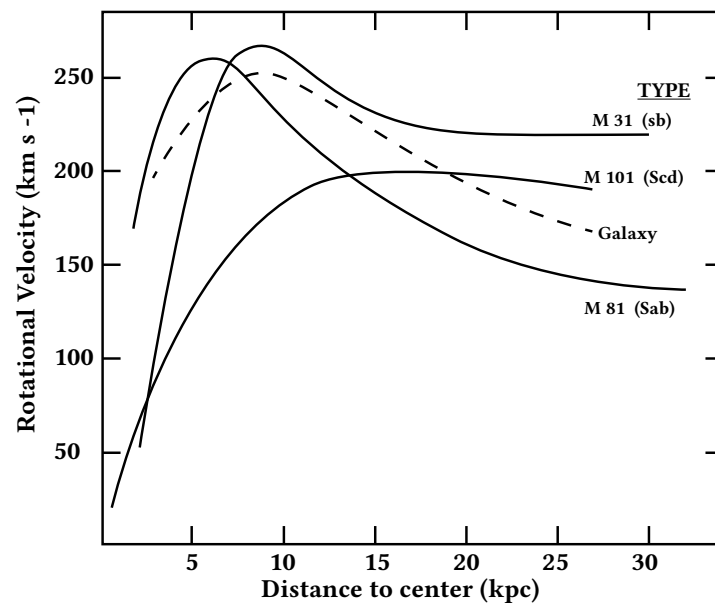
Ostriker et al. [1974] and Einasto et al. [1974] synthesized these results into a single hypothesis-driven analysis. They argued that if each galaxy is embedded in a massive halo extending beyond its visible edge, then:

- (i) galaxy rotation curves can remain approximately flat;
- (ii) a cluster of such galaxies will contain sufficient total mass, in the sum of extended halos, to account for high galaxy velocities; and
- (iii) the collective mass in galaxy halos could contribute substantially to the cosmic mass budget, potentially bringing  $\Omega$  closer to unity.

At the start of their paper, Ostriker et al. [1974] write that *"observations may be consistent with a Universe which is 'just closed' ( $\Omega = 1$ ) — a conclusion believed strongly by some (Wheeler 1973) for essentially non-experimental reasons"*. They add that there are *"reasons, increasing in number and quality, to believe that the masses of ordinary galaxies have been underestimated by a factor of 10 or more"* [Ostriker et al., 1974, L1]. They also close their paper by noting that large-scale-structure considerations — in particular, the extent of rich clusters and galaxy-position correlations interpreted via gravitational instability — *appear to indicate that  $\Omega \approx 1$*  [Ostriker et al., 1974, L4]. By contrast, Einasto et al. [1974] explicitly present massive coronas as a way to *"considerably reduce (if not remove)"* the virial mass discrepancy in clusters, while also noting an implied matter density in galaxies of about 20% of the critical cosmological density [Einasto et al., 1974, 310]. In essence, the 1974 synthesis took phenomena at different

scales, previously treated as independent, and made them jointly relevant to the introduced missing-mass hypothesis. Put differently, they offered a solution to a problem made pressing in part by non-empirical constraints on cosmological models, namely that the universe ought to be closed. Weinberg’s question in the quote above captures the spirit of the 1974 papers: if  $\Omega \geq 1$  is treated as a serious target, then missing mass becomes a necessity and not merely a speculation.

I claim that standard Bayesian conditionalization alone does not capture what changed in 1974. By introducing a shared halo framework and associated parameters, two phenomena previously treated as separate anomalies became informative about one another under a single hypothesis. In that sense, 1974 marks a transition from long-known discrepancies to evidence for dark matter. The reconfiguration matters because it changes what can be inferred from the already available empirical data. Once the halo framework is in place, rotation curves can constrain expectations about cluster dynamics, and cluster dynamics can constrain expectations about rotation curves, in ways that were not warranted when the phenomena were treated in isolation. The resulting evidential picture bears directly on the universe’s total mass budget, which as we have seen was a question many cosmologists treated as urgent in light of background constraints such as  $\Omega \approx 1$ .



**Figure 1:** Rotation curves of three spiral galaxies (solid lines) as a function of distance from the center. (The Milky Way’s expected Keplerian decline is shown for reference as the dotted line.) The flatness of the curves at large radii indicates the presence of unseen mass.

Two implications follow from this reconfiguration. First, it changes the confirmational status of a hypothesis without introducing new observations. With the halo framework in place,



rotation curves and cluster dynamics no longer function as separate anomalies. They become mutually constraining results whose expected relations under the missing-mass hypothesis can be evaluated. This is the setting in which a Myrvoldian increase in mutual information becomes available in principle. The evidential gain comes from the reciprocal dependence that emerges between the phenomena under the shared parameterization, not from the mere postulation of a common origin.

Second, the same reconfiguration helps explain an otherwise puzzling divergence in theory assessment between astronomers and cosmologists. For cosmologists, background constraints associated with  $\Omega \approx 1$  made missing mass an especially salient candidate, and they were correspondingly more willing to treat the link between the phenomena as well motivated. For many astronomers, who were not constrained by closure considerations, the dependencies introduced in 1974 would have appeared less compelling, more contingent on auxiliary assumptions, or simply ad hoc. The next sections therefore examine, first, how such constraints can be represented within a Bayesian reconstruction of theory assessment, and second, whether the 1974 synthesis can be characterized as an increase in confirmation in Myrvold's sense of mutual-information unification.

## 4 Permissible community disagreement

This section develops aspect on theory assessment announced in the overview. The 1974 synthesis papers were not received uniformly. Astronomers and cosmologists overlapped substantially in what they took the relevant observational data to be, yet they differed in how strongly that data supported a missing-mass hypothesis. I reconstruct that divergence in Bayesian terms and defend a modest normative claim: disagreement can be *rationally permissible* when it reflects different, but defensible, background constraints.

### 4.1 Evidence and background

Rubin's retrospective makes clear that many astronomers initially resisted treating extensive halos as the default conclusion from rotation-curve data:

In 1977, many astronomers hoped that dark matter might be avoided [...] there were still non-believers. One eminent astronomer said to me, 'When you observe low luminosity galaxies, you'll find Keplerian falling rotation curves.' [Rubin, 2004, 6]

And:

Kalnajs' (1983) insistence that dark matter is not required, at least for a few galaxies with spatially limited data, convinced a few astronomers that dark matter could be avoided. In retrospect, *we think it is fair to say that many astronomers hoped that Kalnajs was right; dark matter was to be avoided, if at all possible.* [Rubin, 2004, 7]  
[My emphasis]

It is not necessary to explain the divergence sociologically. Evidential support is evaluated relative to background assumptions and modeling commitments. When those differ across communities, the same observational results can reasonably be taken to bear more, or less, strongly on the same hypothesis.

For many astronomers, the salient question was whether the data, together with accepted dynamical and astrophysical auxiliaries, warranted a robust ontological inference to a new mass component on galactic scales. This foregrounded issues internal to astronomical practice, such as systematics, tracer populations, and the adequacy of particular dynamical models. In that setting, it was methodologically appropriate to explore whether revisions to auxiliaries or refinements of modeling could accommodate the anomalies without committing to pervasive unseen matter. Cosmologists, by contrast, were assessing global models and parameters, including the universe's total mass density. Within that project, background constraints associated with  $\Omega \approx 1$  and, for some, closure considerations, made the mass-budget problem urgent in a way it was not for many practicing astronomers. It is not as though the cosmological constraints were taken to be evidence for halos themselves, but they shaped which questions were treated as central, and whether it was worthwhile to integrate results from galaxy and cluster dynamics into a single evidential argument. On this reading, the split is a plausible case of permissible disagreement: distinct assessments can be rationally permissible when they arise from different, defensible background constraints, even with substantial agreement about the empirical data.

## 4.2 Cosmologists' priors

A Bayesian reconstruction makes this structure explicit. Let  $h$  denote a halo-based dark matter hypothesis (for example, that galaxies are embedded in massive halos), let  $e$  denote the relevant empirical results (rotation curves and cluster dynamics), and let  $b$  denote background assumptions that fix modeling commitments and auxiliaries. The relevant quantity is therefore  $Pr(h | e, b)$  rather than  $Pr(h | e)$ . Bayes' theorem can be written as:

$$\begin{aligned} Pr_{\text{new}}(h) &= Pr(h | e, b) = \frac{Pr(e | h, b) Pr(h | b)}{Pr(e | b)} \\ &= \frac{Pr(e | h, b) Pr(h | b)}{Pr(h | b) Pr(e | h, b) + Pr(\neg h | b) Pr(e | \neg h, b)} \end{aligned} \quad (1)$$

This representation highlights two routes by which communities can diverge while remaining epistemically responsible.

1. Differences in priors, relative to background assumptions.

Communities can assign different  $Pr(h | b)$  because their background constraints make different hypotheses antecedently plausible. In the present case, closure-related considerations and associated cosmological principles could raise the prior plausibility of additional gravitating mass. They enter through  $b$  and thereby shape which hypotheses are treated as serious candidates within the relevant modeling context.

2. Differences in likelihoods, due to auxiliaries.

Even when parties agree on the observational data, they may reasonably differ about  $Pr(e \mid h, b)$  and  $Pr(e \mid \neg h, b)$ , since these likelihoods depend on auxiliary assumptions, for example assumptions about equilibrium, tracer populations, mass-to-light ratios, or the reliability of particular measurements. In the 1970s, disagreement about such auxiliaries could make it reasonable for some astronomers to treat the support for halos as weaker or more conditional than cosmologists did when working within a mass-budget framework.

For present purposes, it helps to isolate a simplified case. Suppose that, relative to a broadly shared background  $b$  concerning mechanics and measurement practices, both communities agree that  $h$  makes  $e$  comparatively expected. That is,  $Pr(e \mid h, b)$  is high while  $Pr(e \mid \neg h, b)$  is comparatively low. The evidential force of  $e$  for  $h$  is then captured by the Bayes factor:

$$B = \frac{Pr(e \mid h, b)}{Pr(e \mid \neg h, b)} \quad (2)$$

When  $B$  is large,  $e$  strongly favors  $h$  over its negation. In that setting, if the communities' likelihood assessments are broadly aligned, divergence in  $Pr(h \mid e, b)$  will primarily reflect differences in the prior term  $Pr(h \mid b)$ .

Historically, closure considerations plausibly functioned as a non-empirical constraint for the prior [de Swart, 2020]. Again, a preference for  $\Omega = 1$  did not, by itself, supply new observational evidence for halos, but it did make additional gravitating mass a comparatively plausible solution, and it made cross-subfield integration (e.g., from galactic dynamics to the cosmic mass budget) methodologically prominent. At the same time, even in Ostriker et al. [1974] the closure target was not presented as a purely non-empirical preference. In a closing remark they note that the extent of rich clusters and galaxy-position correlations, when interpreted from the viewpoint of gravitational instability, *appear to indicate that*  $\Omega \approx 1$  [Ostriker et al., 1974, L4]. In hindsight, the Einstein–de Sitter expectation that matter alone closes the universe did not survive in its original form. Still, the broader point that the mass-budget was a serious problem, requiring cosmological modeling to include substantial non-luminous components, endured.

Normatively, this supports a moderately permissive but constrained view of non-empirical considerations in scientific reasoning. Cosmologists' increased antecedent confidence in additional mass was anchored in an established relativistic-cosmology programme, in the cosmological principle and related Machian considerations, and in theoretical arguments favoring an Einstein–de Sitter,  $\Omega = 1$ , universe, even though the best dynamical estimates at the time still placed the actual matter density at only  $\approx 10\text{--}20\%$  of the critical value. On the criteria stated in the overview, closure considerations were defensible insofar as they were motivated within an established modeling context, had previously helped organize successful global inferences, and were treated as defeasible constraints rather than as dogma.

The remaining paper will be devoted to the question of if, and how, unification in the context of the 1974 papers can be viewed as epistemically robust.

## 5 Unification

The preceding discussion helps explain how distinct communities could rationally diverge in their assessments of missing mass under different background constraints. The remaining question is how the 1974 synthesis could nonetheless be epistemically significant in the stronger sense. The 1974 papers by Ostriker et al. [1974] and Einasto et al. [1974] are seen as significant because they *unified* separate observational anomalies under a missing-mass hypothesis:

The 1974 papers synthesized the two instances of curious galaxy behaviour into a single framework, thereby coalescing the problems into a single anomaly of missing mass. [de Swart et al., 2017, 6]

Contemporaries and later commentators treated the unification as not only *epistemically significant*, but as “groundbreaking papers, each presenting a strong case for the existence of large amounts of mass in the outer parts of galaxies.” [Bertone and Hooper, 2018, 21]<sup>9</sup> Rubin’s retrospective remark is representative:

Science often advances when ideas, formerly very disparate, are united. In retrospect, it took a long time for astronomers to relate Zwicky’s dark matter to the flat rotation curves for some galaxies that were beginning to attract attention. [Rubin, 2004, 3]

Tremaine similarly characterized the unification as a watershed in galaxy dynamics and cosmology [Tremaine, 1999, 1223]. If we take seriously the proposition that the unification of flat rotation curves and the cluster mass discrepancy by dark matter in 1974 was epistemically significant, philosophy of science should be able to model, explain, or account for this significance.

The difficulty is that the epistemic value attributed to unification is not straightforwardly the value of *new* observational evidence but, as Rubin’s remark suggests, concerns how already known phenomena are connected. This sits comfortably with explanationist accounts that treat unification as an explanatory virtue (for example, by common-origin explanation). For strict Bayesian conditionalization, however, it is generally awkward since unification looks like a meta-level property of a modeling framework rather than a new data point to condition on. The task in what follows is therefore twofold: first, to show how unification can nonetheless be confirmatory in Bayesian terms (via Myrvold’s mutual-information unification); and second, to explain how such confirmation can be compatible with the fact that the relevant observational results were already known and even helped motivate the 1974 synthesis.

---

<sup>9</sup>Also, in Rubin et al., 1978, L111: “The major result of this work is the observation that rotation curves of high-luminosity spiral galaxies are flat, at nuclear distances as great as  $r = 50kpc$ . [...] These results take on added importance in conjunction with the suggestion of Einasto, Kaasik, and Saar (1974), and Ostriker, Peebles, and Yahil (1974) that galaxies contain massive halos extending to large  $r$ . Such models imply that the galaxy mass increases significantly with increasing  $r$  which in turn requires that rotational velocities remain high for large  $r$ . The observations presented here are thus a necessary but not sufficient condition for massive halos.”

### 5.1 Unification, explanation, and the Bayesian challenge

Much of the philosophical literature characterizes unification primarily as an *explanatory virtue*, and this is the sense in which unification most naturally aligns with broadly explanationist accounts of scientific inference like inference to the best explanation (IBE). The guiding thought is that if two theories accommodate the same total evidence, but one does so with a more integrated explanatory scheme, for example by positing a single cause for a variety of otherwise disparate phenomena, we are justified in believing that it is true. This explanatory view is defended as epistemically significant by, among others, Kitcher [1989], Lipton [2003], and Psillos [1999, 2009]. On this construal, the virtue is non-empirical. It concerns the organization and economy of explanation rather than the introduction of new empirical content.

Fitting this idea into Bayesian confirmation theory is difficult. On a strict Bayesian view, confirmational relations are fixed by priors and likelihoods relative to background assumptions.<sup>10</sup> If two hypotheses are empirically equivalent relative to the same  $b$ , then unification is not, *as such*, an additional evidential relation. At most, it is a heuristic guide to theory construction and pursuitworthiness. This skeptical view is expressed by for example Howson and Urbach [2006] and Hartmann and Sprenger [2011].

The choice of scientific inference to model unification may appear to be a foregone conclusion, but the issue with IBE in this context is that the 1974 synthesis was not seriously considered to settle the truth of a hypothesis implying an ontologically new form of matter, neither by astronomers or cosmologists. Moreover, standard explanationist treatments of unification typically articulate comparative epistemic virtues (better explanation, greater integration) without, by themselves, specifying a dynamics of credence change over time. For that reason, it is useful to turn to explicitly Bayesian tools for representing how unification can generate a confirmational surplus.

### 5.2 Mutual-information unification

Myrvold's account supplies a precision that pure explanatory treatments of unification typically lack. In Myrvold [2003, 2017], unification is not defined by the postulation of a common origin, but by a change in the dependence structure of the evidence. A hypothesis unifies in Myrvold's sense when it makes distinct evidential claims informative about one another.

- Let  $\mathbf{e} = \{p_1, p_2, \dots, p_n\}$  be a body of evidence.
- Let  $I(p; q \mid X)$  denote the mutual information between propositions  $p$  and  $q$  given background  $X$ .<sup>11</sup>
- If  $I(p_1; p_2 \mid b) \approx 0$  but  $I(p_1; p_2 \mid h \wedge b) > 0$ , then  $h$  renders  $p_1$  and  $p_2$  mutually informative relative to  $b$ .

Myrvold captures the unificatory contribution by defining:

<sup>10</sup>For attempts to address this *prima facie* tension between probabilistic reasoning and explanation see Henderson [2014], Weisberg [2009], Iranzo García [2008], and Douven and Schupbach [2015].

<sup>11</sup>Intuitively, mutual information measures how much learning  $p$  reduces uncertainty about  $q$ , given  $X$ .

$$U(p_1, p_2; h | b) = I(p_1; p_2 | h \wedge b) - I(p_1; p_2 | b) \quad (3)$$

In the idealized case where  $I(p_1; p_2 | b) \approx 0$ , a positive  $U$  reduces to  $I(p_1; p_2 | h \wedge b)$ . The hypothesis introduces an evidential connection that is not present given the background alone. To connect this to confirmation, it is useful to work with an information-theoretic measure of evidential support, equivalently a log Bayes factor:

$$C(h; e | b) = \log \frac{\Pr(e | h, b)}{\Pr(e | \neg h, b)} \quad (4)$$

Myrvold shows that, under appropriate assumptions, the support provided by a conjunction  $e_1 \wedge e_2$  decomposes into the support from  $e_1$ , the support from  $e_2$ , and an additional unification term:

$$C(h; e_1 \wedge e_2 | b) = C(h; e_1 | b) + C(h; e_2 | b) + U(e_1, e_2; h | b) \quad (5)$$

The normative interest of MIU is that it identifies a principled sense in which unification can yield a confirmational surplus. If  $U > 0$ , the conjunction provides more support than the sum of its conjuncts because the hypothesis makes the evidential claims mutually constraining.

### 5.3 Evidential reconfiguration in 1974

The question is whether the 1974 synthesis fits this MIU structure. Let  $h$  be the halo-based dark matter hypothesis, and let  $p_1$  and  $p_2$  denote the two relevant phenomena:

- $h$  : galaxies are embedded in extended halos with substantial unseen mass,
- $p_1$  : the galaxy cluster mass discrepancy (virial mass deficit),
- $p_2$  : flat galaxy rotation curves at large radii,
- $b$  : background assumptions (GR, ordinary matter, measurement conventions),
- $\theta$  : halo-to-luminous mass normalization parameter.

A brief clarification of notation. In any concrete halo model,  $h$  permits a substantial parameter space  $\Theta$  (profile shapes, scale radii, truncations, anisotropies, and so on). For present purposes I suppress these degrees of freedom and let  $\theta$  denote the resulting one-dimensional *normalization coordinate* that measures how much gravitating mass is associated with a galaxy relative to its luminous component. Nothing turns on whether  $\theta$  is expressed as a mass-to-light ratio (in solar units) or as a dimensionless mass-discrepancy factor — I use mass-to-light ratio because it best fits the 1974 usage. For example, Ostriker et al. [1974] define a galactic mass-to-light ratio and estimate values of order  $\sim 200$  for giant spirals:

$$f_{\text{sp}} \simeq 200 h^0 t_{10}^{-1} \quad (\text{Ostriker et al., 1974, eq. (1)})$$

Einasto et al. [1974] report that allowing for massive ‘coronas’ raises the mass–luminosity ratio to  $f \approx 100$  for spirals and  $f \gtrsim 120$  for ellipticals [Einasto et al., 1974, 310]. Nothing in the MIU argument turns on the detailed profile parameters, but it is worth seeing how the conceptual schema I developed here couples with actual parameter values from the two papers. What does matter is that learning  $p_1$  (or  $p_2$ ) materially concentrates the posterior over  $\theta$ , i.e., restricts its support to a proper subset of the  $\theta$ -range derived by  $\Theta$ . Prior to 1974,  $p_1$  and  $p_2$  were typically assessed in their own modeling contexts and with different local auxiliaries:

Clearly, in neither branch of astronomy were these observations at this point cited as positive evidence for the presence of extra matter, or falsifying evidence for any alternative hypothesis. Furthermore [...] the two problems were studied separately. [de Swart et al., 2017, 3]

It is therefore a reasonable idealization to treat them as approximately independent relative to  $b$ , so that  $I(p_1; p_2 | b) \approx 0$ .<sup>12</sup> In that setting, learning  $p_1$  does not constrain expectations about  $p_2$ , and vice versa. The 1974 unification changes the evidential situation because  $h$  introduces a shared parameterization. Under  $h \wedge b$ , both  $p_1$  and  $p_2$  depend on the same halo parameter  $\theta$ . Learning either phenomenon constrains  $\theta$ , and that constraint feeds through to expectations about the other. In Myrvold’s sense:

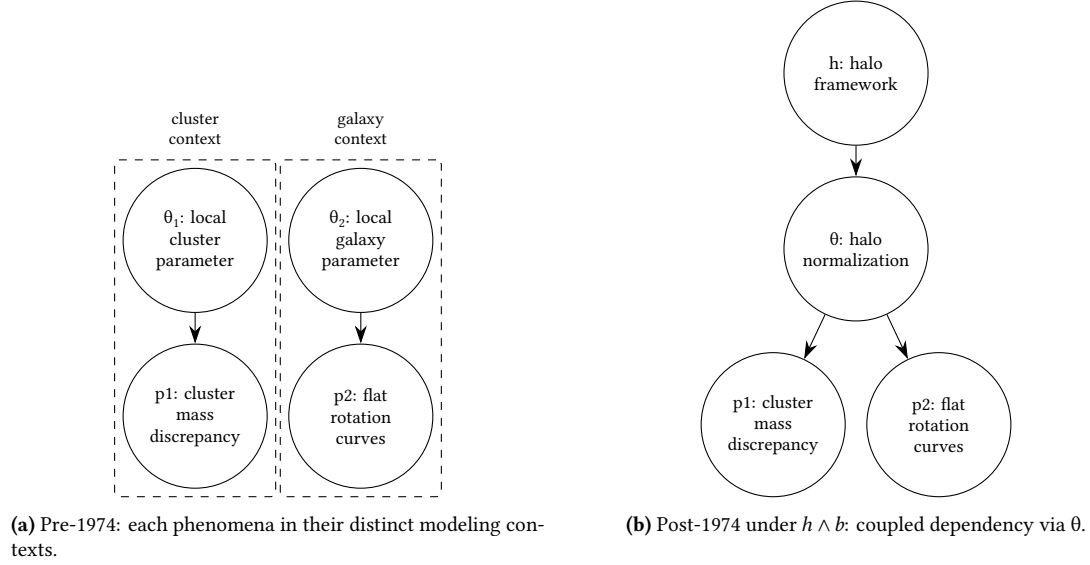
$$I(p_1; p_2 | h \wedge b) > 0 \quad \text{and therefore} \quad U(p_1, p_2; h | b) > 0 \quad (6)$$

The point is straightforward in informal terms. Under a halo framework, flat rotation curves constrain the magnitude and distribution of additional gravitating mass associated with galaxies. Those same halo parameters constrain how much mass galaxies contribute to clusters. Conversely, cluster mass discrepancies constrain the typical magnitude of galactic halos, which bears on expectations about rotation curves. What is gained is not a mere conjunction of old facts, but a defensible two-way evidential connection. What matters is not that  $p_1 \wedge p_2$  obtains, but whether their co-variation is the sort of relation one should expect under a shared, parameterized halo framework. That constraint is substantive only if learning  $p_1$  materially updates the posterior over  $\theta$  and thereby changes expectations for  $p_2$ . One way to display this is via the posterior predictive distribution:

$$Pr(p_2 | p_1 \wedge h \wedge b) = \int_{\Theta} Pr(p_2 | \theta \wedge h \wedge b) Pr(\theta | p_1 \wedge h \wedge b) d\theta \quad (7)$$

Equation (7) makes the dependence explicit. Learning  $p_1$  concentrates the posterior  $Pr(\theta | p_1 \wedge h \wedge b)$  on a restricted region of the parameter space  $\Theta$ , and expectations for  $p_2$  are obtained by averaging the model’s  $\theta$ -indexed predictions over that posterior. The modeling assumption is that, conditional on  $\theta$  (and  $b$ ),  $p_1$  bears on  $p_2$  only through its effect on  $Pr(\theta | p_1 \wedge h \wedge b)$ .

<sup>12</sup>This idealizes the pre-1974 situation. The claim is not that no link between the phenomena was conceivable, but that no stable, widely endorsed linkage was treated as evidentially significant across domains in the relevant communities.



**Figure 2:** Directed acyclic graphs illustrating evidential reconfiguration in 1974. Pre-1974:  $p_1$  and  $p_2$  are treated as separate anomalies in distinct local modeling contexts relative to background  $b$ . Post-1974: under the halo hypothesis  $h$ , both are constrained by a shared parameter  $\theta$ , which makes  $p_1$  and  $p_2$  mutually informative. Arrows represent probabilistic dependence relations in the modeling framework, not causal influence from observations to  $\theta$ .

At this point a familiar Bayesian worry arises. If  $p_1$  and  $p_2$  were already established well before 1974, then there is no straightforward conditionalization event corresponding to learning them, and it can seem that no increase in support for  $h$  is possible. The next section argues that this is a representational mistake: the epistemic novelty lies in a newly warranted dependence claim (and hence a new evidential role for  $p_1$  and  $p_2$ ), not in the arrival of new observational propositions.

## 6 Old evidence and structural novelty

The old-evidence problem is usually framed as follows [Glymour, 1981]. If an agent assigns probability 1 to an evidential proposition  $e$ , then conditionalizing on  $e$  cannot change their credences. How, then, can a hypothesis be confirmed by facts that were already established, and in the present case facts that appear to have been part of the motivation for introducing the hypothesis in the first place, even if, prior to 1974, their confirmational role had not yet been settled? It is useful to begin by isolating a false route into the problem, so it is clear what formulation of the issue we are *not* addressing. It is easy to turn Bayes' theorem into a red herring in discussions of old evidence. Consider:

$$Pr(h \mid e, b) = \frac{Pr(e \mid h, b) Pr(h \mid b)}{Pr(e \mid b)} \quad (8)$$

If  $e$  is already fully accepted, one might be tempted to set  $Pr(e \mid b) = 1$  and conclude that no



confirmational change is possible. But that is not the deeper point. The difficulty is representational in kind. When  $e$  is already part of  $b$  it is integrated in an agent's total evidence in the context of assessment, there is no further conditionalization event corresponding to 'learning'  $e$ . Here, 'old evidence' means facts whose evidential role are stable and consolidated.

For our particular case though, the 1974 period is not accurately described as anyone re-learning  $p_1$  and  $p_2$ , so there is no reason to represent it as such. The epistemic change of interest concerns the *evidential structure* and its associated dependence assumptions, that is, which propositions are treated as background, which are treated as evidence for which hypotheses, and which relations among them are taken to be warranted within an accepted modeling framework. What needs to be modeled is epistemic change of this kind, where accepted facts are reclassified and new evidential roles emerge as a consequence. This is explicitly the lesson learned from de Swart et al. [2017]:

representations of the establishment of dark matter in terms of an accumulation of evidence miss an essential part of this history: they overlook the necessary conditions that made this very accumulation possible, an accumulation that, at face value, was substantially, even if partly, a reinterpretation of existing observations. Here, theory, along with an institutional shift and expansion in astronomy, played a substantive role: a role that reminds us that simply asking "what was the evidence for missing matter" misses the point; we need to understand why certain observations were eventually conceived as 'evidence' of anything in the first place. [de Swart et al., 2017, 6]

For that reason, strategies that simulate ignorance by temporarily removing  $e$  from the background (for example via forgetful functions) are neither useful nor historically accurate here. The dark matter case concerns evidential role and dependence, not the arrival of new observational propositions.

### 6.1 Structural novelty as evidence

There are at least two concepts formulated with the purpose to resolve at least some of the challenges posed by the old evidence problem. *Temporal novelty* denotes evidence not known before a theory is formulated, and *use novelty* denotes evidence not used in constructing the theory. Earman [1992] provides a definition and confirmation criteria for both (although credits Worrall [1985, 1989] for the latter):

**Temporal novelty:** Suppose that  $T \models E$ . If  $E$  was already known to be true prior to the articulation of  $T$ , then  $E$  does not confirm  $T$ .

**Use novelty:** Suppose that  $T_1 \models E$  and  $T_2 \models E$ . If  $E$  was used in constructing  $T_1$ , but not in constructing  $T_2$ , then  $T_2$  receives more support from  $E$  than does  $T_1$ . [Earman, 1992, 114]

In the dark matter case, that galaxy rotation curves were flat, and cluster dynamics had a virial mass deficiency was *neither* temporally novel (it was known already) *nor* use novel (it was explicitly used to formulate  $h$ ). So, by these classical accounts, it should not confirm  $h$  at all —

it was both ‘old’ and ‘used’. The epistemic gain must therefore be located elsewhere. I argue that it can be found in the relations or connections that  $h$  supports among observational facts that were not previously taken to be warranted across domains.

To represent our case more accurately, it is useful to distinguish the possession of a fact from its evidential role. A community may possess an observational result and treat it as established, while still not treating it as central evidence for a given hypothesis. This better describes the pre-1974 situation. The phenomena corresponding to  $p_1$  (cluster mass discrepancies) and  $p_2$  (flat rotation curves) were known and largely uncontested *qua* observational facts, but their evidential roles were assessed within relatively local contexts of astronomical problems. What changes in 1974 is not the descriptive content of  $p_1$  and  $p_2$ , but the evidential roles they are made to play in a cosmological context, under a halo-based framework, such that they impose informational constraints on one another.

It may help to separate two aspects of an accepted observational claim. One is descriptive: the propositional content of  $p$  itself, treated as established or highly credible. The other is evidential: the role  $p$  plays in supporting or disfavoring hypotheses relative to a background  $b$ . These aspects can come apart in significant ways. In our case, the descriptive content of  $p_1$  and  $p_2$  remained stable across the relevant period, but the evidential role they played was neither stable nor settled. The epistemic event in 1974 is therefore not that the community comes to accept the conjunction  $p_1 \wedge p_2$  as a pair of descriptive facts, but that a new evidential role is ascribed to them through a structural claim about their relation. In that sense, the old-evidence slogan targets the wrong object. The epistemic significance lies in the dependence claim — the *structural novelty* — not in the bare conjunction of already accepted propositions.

Observational facts do not wear their evidential roles on their sleeves. The mere possession of a result does not determine whether, or how, it should function as evidence for a given hypothesis. What 1974 offered was a reason to treat  $p_1$  and  $p_2$  as connected under  $h \wedge b$  through a shared parameterization  $\theta$ . In the Myrvoldian framework, the point is that  $p_1$  and  $p_2$  become mutually informative under  $h \wedge b$ , even though they are (approximately) independent relative to  $b$ . From this perspective, the epistemically novel element is a structural claim, mediated by a common modeling parameter:<sup>13</sup>

$$S : I(p_1; p_2 \mid h \wedge b) > 0 \quad \text{and} \quad p_1 \perp p_2 \mid (\theta \wedge h \wedge b) \quad (9)$$

In eq. (9), the conditional-independence claim says that the evidential linkage between  $p_1$  and  $p_2$  under  $h \wedge b$  is mediated, in the probabilistic sense that  $p_1$  and  $p_2$  are screened off, by the shared halo parameter  $\theta$ . Put informally, under  $h \wedge b$ , evidence about rotation curves constrains what should be expected about cluster dynamics, and vice versa. In that sense,  $S$  functions as a predictive constraint. The claim can succeed or fail depending on whether the shared parameterization genuinely generates substantive constraints, or whether comparable dependence can be recovered without  $h$  (for example via alternative auxiliaries). This is a case of *structural novelty*. What changes is the dependence structure among accepted claims, not which first-

<sup>13</sup>I use  $\perp$  for (conditional) probabilistic independence. Formally,  $p_1 \perp p_2 \mid X$  iff  $Pr(p_1 \wedge p_2 \mid X) = Pr(p_1 \mid X) Pr(p_2 \mid X)$ , equivalently  $Pr(p_1 \mid p_2 \wedge X) = Pr(p_1 \mid X)$  (when the relevant conditional probabilities are defined). Informally, conditional on  $X$ , learning  $p_1$  does not provide any further information about  $p_2$  (and vice versa).

order propositions are known. That is also why temporal novelty and use novelty diagnoses are too blunt here.

## 6.2 When is structural novelty epistemically warranted?

Structural novelty, in the sense just introduced, is not automatically epistemically good news. The question is when a newly proposed dependence claim is substantive rather than an artifact of a flexible parameterization — that is, when it yields a genuine confirmational gain rather than merely *manufacturing* unification by stipulation. It is therefore useful to be explicit about MIU’s fallibility. In the present case there are at least three ways an MIU surplus could fail to materialize. First, learning  $p_1$  might leave  $\theta$  with too many degrees of freedom to generate any non-trivial expectation for  $p_2$ . Second, the halo framework itself might be too underspecified, allowing many halo specifications that fit  $p_1$  while leaving  $p_2$  effectively unconstrained. Third, an analogous dependence might be recoverable under  $\neg h$  by altering auxiliary assumptions in a comparably well-motivated way.

We can view fig. 2 as comparing two *inferential frameworks* for the same observational facts  $p_1$  and  $p_2$ .<sup>14</sup> The pre-1974 *baseline* in which  $p_1$  and  $p_2$  were treated in separate modeling contexts, and the post-1974 halo framework in which a shared parameterization makes the domains mutually constraining. This avoids conditioning on  $h$  in the pre-1974 setting, where no halo hypothesis with a shared parameterization was yet available. Let  $M_0$  denote this pre-1974 *disunified baseline*, lacking an articulated theory about both  $p_1$  and  $p_2$ . It represents the prevailing evidential situation in which each observational fact was assessed within its own local modeling context (with distinct auxiliaries and, if one wishes, distinct latent parameters). A natural idealization is that the joint marginal likelihood factorizes:

$$\Pr(p_1, p_2 \mid b, M_0) = \Pr(p_1 \mid b, M_0) \Pr(p_2 \mid b, M_0) \quad (10)$$

Let  $M_1$  denote the post-1974 *unified halo inferential framework* (fig. 2b), i.e., the halo hypothesis  $h$  together with a shared normalization parameter  $\theta$  that links the two domains. On  $M_1$ , the joint marginal likelihood is obtained by averaging over the shared parameter:

$$\Pr(p_1, p_2 \mid b, M_1) = \int_{\Theta} \Pr(p_1 \mid \theta, h, b) \Pr(p_2 \mid \theta, h, b) \Pr(\theta \mid h, b) d\theta \quad (11)$$

This enables assessment in standard Bayesian terms via a Bayes factor comparing the unified halo framework to the disunified baseline:

$$B_{M_1:M_0}(p_1, p_2 \mid b) = \frac{\Pr(p_1, p_2 \mid b, M_1)}{\Pr(p_1, p_2 \mid b, M_0)} \quad (12)$$

On this reading, the epistemic contribution of structure is neither mysterious nor in tension with Bayesian principles. What changes in 1974 is that  $M_1$  makes  $p_1$  and  $p_2$  *mutually constraining* by connecting both to the same parameter  $\theta$  where learning one phenomenon concentrates

<sup>14</sup>I use ‘inferential framework’ in the more colloquial sense. In Bayesian terms each framework corresponds to a probabilistic model specifying a likelihood factorization and associated parameters.

$\Pr(\theta \mid \cdot)$  and thereby alters expectations about the other (eq. (7)). By contrast, under the disunified framework  $M_0$  the evidential streams remain effectively decoupled in the sense of eq. (10). This way of stating the point simply internalizes the dependence assumptions into the framework label: where MIU is often written as a contrast between  $I(p_1; p_2 \mid b)$  and  $I(p_1; p_2 \mid h \wedge b)$ , here the relevant contrast is between the pre-1974 baseline  $M_0$  and the post-1974 halo framework  $M_1$ . In MIU terms  $M_0$  corresponds to  $I(p_1; p_2 \mid b, M_0) \approx 0$ , whereas adopting the halo framework  $M_1$  renders  $I(p_1; p_2 \mid b, M_1) > 0$ , thereby making a positive unification contribution  $U$  available without introducing new observational content. Importantly, this comparison is compatible with the historical fact that  $p_1$  and  $p_2$  helped motivate the proposal of  $M_1$  since the confirmational question is whether the unified inferential framework renders the joint pattern of the available observational facts less surprising than the disunified framework, as per eq. (12).

### 6.3 MIU and old evidence

The dark matter case does not provide a path to resolve the old-evidence problem in full generality. What it shows is that the slogan that old evidence is confirmationally inert mischaracterize cases in which epistemic change concerns evidential role and dependence rather than the arrival of new observational content. A Myrvold-consistent resolution can be summarized as follows.

- The relevant evidential object is not merely the set  $\{p_1, p_2\}$ , but the reclassified evidential situation in which  $p_1$  and  $p_2$  are treated as jointly constraining, because the halo framework supplies a shared parameterization that links them.
- The confirmational surplus is not generated by re-learning old facts, but by endorsing and exploiting the dependence claim  $S$ , namely that under  $h \wedge b$  the two phenomena become mutually informative.

In this way, MIU provides a principled route for treating the 1974 synthesis as confirmatory without invoking simulated ignorance. The epistemic change is located in a reconfiguration of evidential relations that made the linkage between cluster dynamics and rotation curves methodologically salient and epistemically assessable. The concluding section is then in a position to draw the broader lessons about evidence construction in cosmology, including the role of non-empirical constraints in structuring assessment, the rationality of community divergence, and the sense in which MIU captures the confirmational source of the 1974 pivot.

## 7 Conclusion

The 1974 papers by Ostriker et al. [1974] and Einasto et al. [1974] are recognized as pivotal in the evidential consolidation of the dark matter hypothesis. On a standard picture of Bayesian updating by conditionalization, this is puzzling. The papers did not introduce new observations, and they relied on results that were already in the scientific records. I have argued that this is only a puzzle in appearance.

The central historical point is that the 1974 halo proposal altered what hypotheses rotation curves and cluster dynamics could jointly support by linking them through a shared parameterization. That is what I have called *evidential reconfiguration*. On Myrvold’s mutual-information account of unification, the epistemic gain is the emergence of a two-way informational constraint. Evidence about one phenomenon becomes evidence about what should be expected of the other, because both imposes informational restrictions on the same halo parameter(s). In this sense, unification can yield a confirmational surplus even when the underlying observations are temporally old.

This also clarifies why astronomers and cosmologists responded differently to the same broad observational situation. The communities were not simply disagreeing about facts. They were working under different background constraints, and therefore differed both in priors and in which inferences were treated as methodologically worth pursuing. In that setting, divergent assessments of the support for missing mass are rationally permissible, provided the relevant constraints are independently motivated, empirically fruitful, and treated as defeasible.

Finally, the dark matter case sharpens how the old-evidence problem should be understood in historically realistic settings. The novelty in 1974 is not that  $p_1$  and  $p_2$  were learned again, but that a new structural claim about their relation became warranted and epistemically usable. The broader lesson is that evidence in science is not only accumulated by the aggregation of descriptive facts but also organized and structured to reveal evidential patterns. A Bayesian framework can represent this organizational dimension by tracking how background constraints and modeling choices determine which dependence relations are admissible, and therefore which bodies of results can function as evidence for which hypotheses.

## References

- S. Allén. Scientific realism and empirical confirmation: A puzzle. *Studies in History and Philosophy of Science Part A*, 90:153–159, 2021. Publisher: Elsevier.
- S. Allén. Dark Matter: Explanatory Unification and Historical Continuity, Dec. 2024. URL <http://arxiv.org/abs/2412.13404>. arXiv:2412.13404 [physics].
- A. Antoniou. Robustness and Dark-Matter Observation. *Philosophy of Science*, 90(3):629–647, 2023.
- A. Antoniou. Why did the dark matter hypothesis supersede modified gravity in the 1980s? *Studies in History and Philosophy of Science*, 112:141–152, 2025. Publisher: Elsevier.
- G. Bertone and D. Hooper. History of dark matter. *Reviews of Modern Physics*, 90(4):045002, 2018. Publisher: APS.
- H. Bondi and T. Gold. The steady-state theory of the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 108(3):252–270, 1948. Publisher: Oxford University Press Oxford, UK.
- A. Bosma. *The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types*. PhD Thesis, Rijksuniversiteit te Groningen., 1978.

- E. Castellani, R. Dardashti, and R. Dawid. Unification and surprise: On the confirmatory reach of unification, 2025. URL <https://philsci-archive.pitt.edu/26916/>.
- D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter. *The Astrophysical Journal Letters*, 648(2):L109–L113, 2006. doi: 10.1086/508162.
- R. Dawid. *String theory and the scientific method*. Cambridge University Press, 2013.
- R. Dawid. The significance of non-empirical confirmation in fundamental physics. *Why trust a theory? Epistemology of modern physics*, pages 99–119, 2019.
- R. Dawid, S. Hartmann, and J. Sprenger. The no alternatives argument. *The British Journal for the Philosophy of Science*, 2015. Publisher: The University of Chicago Press.
- S. De Baerdemaeker. Method-driven experiments and the search for dark matter. *Philosophy of Science*, 88(1):124–144, 2021. Publisher: Cambridge University Press.
- S. De Baerdemaeker and N. M. Boyd. Jump ship, shift gears, or just keep on chugging: Assessing the responses to tensions between theory and evidence in contemporary cosmology. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 72:205–216, 2020. Publisher: Elsevier.
- S. De Baerdemaeker and R. Dawid. MOND and meta-empirical theory assessment. *Synthese*, 200(5):1–28, 2022. Publisher: Springer.
- J. de Swart. *Closing in on the Cosmos: Cosmology’s Rebirth and the Rise of the Dark Matter Problem*, pages 257–284. Springer International Publishing, Cham, 2020. ISBN 978-3-030-50754-1. doi: 10.1007/978-3-030-50754-1\_8. URL [https://doi.org/10.1007/978-3-030-50754-1\\_8](https://doi.org/10.1007/978-3-030-50754-1_8).
- J. de Swart, G. Bertone, and J. van Dongen. How dark matter came to matter. *Nature Astronomy*, 1(3):1–9, 2017. Publisher: Nature Publishing Group.
- I. Douven and J. N. Schupbach. Probabilistic alternatives to bayesianism: the case of explanationism. *Frontiers in Psychology*, 6:459, 2015.
- P. M. Duerr and W. J. Wolf. Methodological reflections on the MOND/dark matter debate. *Studies in History and Philosophy of Science*, 101:1–23, 2023. ISSN 0039-3681. doi: <https://doi.org/10.1016/j.shpsa.2023.07.001>. URL <https://www.sciencedirect.com/science/article/pii/S0039368123000936>.
- J. Earman. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, ma: MIT Press, 1992.
- J. Einasto, A. Kaasik, and E. Saar. Dynamic evidence on massive coronas of galaxies. *Nature*, 250(5464):309–310, 1974. Publisher: Nature Publishing Group.
- C. Glymour. Theory and evidence. *Philosophy of Science*, 48(3), 1981.

- S. Hartmann and J. Sprenger. Bayesian epistemology. In *The Routledge companion to epistemology*, pages 609–619. Routledge, 2011.
- L. Henderson. Bayesianism and inference to the best explanation. *The British Journal for the Philosophy of Science*, 2014.
- C. Howson and P. Urbach. *Scientific reasoning: the Bayesian approach*. Open Court Publishing, 2006.
- V. Iranzo García. Bayesianism and interference to the best explanation. *Theoria*, 23(1), 2008.
- M. Jacquart.  $\Lambda$ cdm and mond: A debate about models or theory? *Studies in History and Philosophy of Science Part A*, 89:226–234, 2021. Publisher: Elsevier.
- P. Kitcher. Explanatory Unification and the Causal Structure of the World. In P. Kitcher and W. Salmon, editors, *Scientific Explanation*, pages 410–505. Minneapolis: University of Minnesota Press, 1989.
- P. Lipton. *Inference to the best explanation*. Routledge, 2003.
- N. C. Martens. Dark matter realism. *Foundations of Physics*, 52(1):16, 2022. Publisher: Springer.
- N. C. Martens, M. A. C. Sahuquillo, E. Scholz, D. Lehmkuhl, and M. Krämer. Integrating dark matter, modified gravity, and the humanities. *Studies in History and Philosophy of Science*, 91:A1–A5, 2022. Publisher: Elsevier.
- D. Merritt. Cosmological realism. *Studies in History and Philosophy of Science Part A*, 88: 193–208, 2021a. Publisher: Elsevier.
- D. Merritt. MOND and Methodology. In *Karl Popper’s Science and Philosophy*, pages 69–96. Springer, 2021b.
- E. A. Milne. World structure and the expansion of the universe. *Nature*, 130(3270):9–10, 1932.
- W. C. Myrvold. A Bayesian account of the virtue of unification. *Philosophy of Science*, 70(2): 399–423, 2003. Publisher: The University of Chicago Press.
- W. C. Myrvold. On the evidential import of unification. *Philosophy of Science*, 84(1):92–114, 2017. Publisher: University of Chicago Press Chicago, IL.
- J. P. Ostriker, P. J. E. Peebles, and A. Yahil. The Size and Mass of Galaxies, and the Mass of the Universe. *The Astrophysical Journal*, 193:L1–L4, 1974. doi: 10.1086/181617.
- S. Psillos. *Scientific realism: How science tracks truth*. Routledge, 1999.
- S. Psillos. *Knowing the structure of nature: Essays on realism and explanation*. Palgrave Macmillan, Basingstoke, 2009. ISBN 978-0-230-23466-6.
- M. Roberts and A. Rots. Comparison of rotation curves of different galaxy types. *Astronomy and Astrophysics*, 26:483–485, 1973.

- V. C. Rubin. A brief history of dark matter. In M. Livio, editor, *The Dark Universe: Matter, Energy and Gravity*, Space Telescope Science Institute Symposium Series, pages 1–13. Cambridge University Press, 2004. doi: 10.1017/CBO9780511536298.003.
- V. C. Rubin and W. K. Ford Jr. Rotation of the Andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159:379, 1970.
- V. C. Rubin, W. K. Ford, Jr., and N. Thonnard. Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa -> Sc. *The Astrophysical Journal*, 225: L107–L111, Nov. 1978. doi: 10.1086/182804.
- S. L. Shapiro. The density of matter in the form of galaxies. *The Astronomical Journal*, 76:291, 1971.
- A. Sus. Dark matter, the equivalence principle and modified gravity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 45:66–71, 2014. Publisher: Elsevier.
- S. Tremaine. Comments on Ostriker, Peebles, & Yahil (1974) “The Size and Mass of Galaxies, and the Mass of the Universe”. *The Astrophysical Journal*, 525:1223, 1999.
- W. L. Vanderburgh. The dark matter double bind: Astrophysical aspects of the evidential warrant for general relativity. *Philosophy of science*, 70(4):812–832, 2003. Publisher: Cambridge University Press.
- W. L. Vanderburgh. The methodological value of coincidences: Further remarks on dark matter and the astrophysical warrant for general relativity. *Philosophy of Science*, 72(5):1324–1335, 2005. Publisher: Cambridge University Press.
- W. L. Vanderburgh. On the interpretive role of theories of gravity and ‘ugly’ solutions to the total evidence for dark matter. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 47:62–67, 2014a. Publisher: Elsevier.
- W. L. Vanderburgh. Quantitative Parsimony, Explanatory Power and Dark Matter. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie*, 45(2):317–327, 2014b. doi: 10.1007/s10838-014-9261-9. Publisher: Springer Netherlands.
- E. Vaynberg. Realism and the detection of dark matter. *Synthese*, 204(3):82, 2024.
- S. Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley & Sons, New York, 1972.
- J. Weisberg. Locating ibe in the bayesian framework. *Synthese*, 167(1):125–143, 2009.
- W. J. Wolf and J. Read. Navigating permanent underdetermination in dark energy and inflationary cosmology, 2025. URL <https://arxiv.org/abs/2501.13521>. \_eprint: 2501.13521.
- J. Worrall. Scientific discovery and theory-confirmation. In *Change and Progress in Modern Science: Papers related to and arising from the Fourth International Conference on History and Philosophy of Science, Blacksburg, Virginia, November 1982*, pages 301–331. Springer, 1985.



- J. Worrall. Fresnel, poisson and the white spot: The role of successful. *The uses of experiment: Studies in the natural sciences*, page 135, 1989.
- F. Zwicky. How Far Do Cosmic Rays Travel? *Phys. Rev.*, 43(2):147–148, Jan. 1933. doi: 10.1103/PhysRev.43.147. Publisher: American Physical Society.
- F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *The Astrophysical Journal*, 86:217, 1937.