Evidential Reconfiguration as Bayesian Confirmation For Dark Matter

How Known Data Becomes New Evidence

Simon Allzén*[1,2,3]

[1]Department of Philosophy, Stockholm University
[2]Institute of Physics, University of Amsterdam
[3]Vossius Center for the History of Humanities and Sciences

June 3, 2026

Abstract. Dark matter is robust at the level of gravitational phenomena but underdetermined at the level of constituent nature. This paper isolates one mechanism by which such a target can nevertheless acquire stronger evidential standing. The 1974 papers of Ostriker et al. (1974) and Einasto et al. (1974) are often treated as pivotal in establishing the evidential status of the dark matter hypothesis, yet neither introduced new observations. I argue that their distinctive contribution is best characterized as *evidential reconfiguration*: a rearrangement of background assumptions, modeling commitments, and warranted dependence relations. The result makes previously domain-local anomalies — in this case galaxy rotation curves and cluster dynamics — mutually constraining under a shared halo parameterization. To explain how confirmation can increase despite relying on old evidence, I distinguish common-origin unification from Myrvold's mutual-information unification and argue that the epistemic gain is captured by the latter. The paper therefore offers a historically grounded Bayesian account of how support can increase without new data.

Contents

1 Introduction 2
1.1 Overview 3

2 The standard narrative 5

*simon.allzen@philosophy.su.se ORCID: 0000-0002-6124-8152

3 Constraints and the mass deficit 6

4 Permissible community disagreement 10
4.1 Evidence and background 10
4.2 Cosmologists' priors 11

5 Confirmatory unification 13
5.1 Common-origin unification 14
5.2 Mutual-information unification 15
5.3 Evidential reconfiguration in 1974 16

6 Old evidence without ignorance 18
6.1 Structural novelty as evidence 19
6.2 MIU and old evidence 20

7 Conclusion 20

References 22

1 Introduction

The dark matter case combines strong gravitational support with radical openness about constituent nature. One important philosophical consequence is that the route by which it acquired evidential standing cannot be read off from a simple accumulation story. Dark matter is a core posit of the standard cosmological model — the ΛCDM (concordance) model.[1] It 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, making it one of the major open problems in contemporary physics and a particularly interesting case for philosophers.[2] One central puzzle is to make sense of 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, which only makes confirmational sense if one assumes that each observation was treated as direct evidence for an already established hypothesis. In reality, however, the observations were often initially treated as local anomalies, disconnected from any specific hypothesis. It is therefore epistemically significant to understand how and why these observations

[1] Λ denotes dark energy, and CDM stands for cold dark matter, where 'cold' refers to dark matter's negligible free-streaming length.

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

came to be considered *evidence* for the dark matter hypothesis. How do long-known observational anomalies transform into coherent evidence for an ontological posit, and under what conditions is that transformation warranted?

This paper offers a philosophical analysis and Bayesian treatment of that transformation. I build on the historical work by Swart et al. (2017) and de Swart (2020), who identify two seminal papers from 1974 as pivotal in the acceptance and establishment of the dark matter hypothesis. While I concede the significance of the 1974 papers, the epistemological task to show *how* that could be, given that neither of the two papers contained any new empirical data, still remains. The paper does not attempt to explain dark matter's full later establishment. It isolates one historically pivotal step and uses it to clarify the problem of old evidence in Bayesian confirmation theory.[3] I contend that the distinctive epistemic contribution of the 1974 papers can be identified in the reconfiguration of pre-existing observations, such that previously treated local and disconnected observations could be characterized as mutually supporting a single hypothesis of missing mass. The main point is that the key epistemic 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 cross-domain dependencies between claims are taken to be warranted within a modeling framework. The normative point I defend is modest in scope, but principally non-trivial: non-empirical constraints can be epistemically valuable insofar as they constrain which dependencies it is reasonable to treat as evidentially relevant.

1.1 Overview

From the brief but significant period in the history of dark matter highlighted by Swart et al. (2017) and de Swart (2020), I extract two philosophically significant features of the reasoning involved which constitutes the two main sections of this paper.

The first concerns theory assessment, and focus on a discrepancy in how scientists in different communities (astronomy and cosmology) assessed the case for dark matter in the mid–1970s. Both communities largely agreed on *what* the relevant empirical observations were, yet cosmologists were more receptive to treating those results as evidential support for a missing-mass hypothesis than many astronomers. I argue that the divergence in theory assessment, despite a shared set of observational data, is best explained by differences in background commitments. In particular, cosmologists treated certain assumptions, sometimes non-empirical ones, (including a strong preference for a "closed" universe) as *constraints* on theory assessment. In a Bayesian reconstruction, this appears as a difference in the value of priors and, more importantly, as a difference in the background assumptions that govern which connections between claims are admissible. The point is not that these assumptions provided additional evidence for

[3] The relevant issue is not that the community *re-learned* the relevant propositions, but that their evidential roles and warranted dependence relations were reclassified and restructured in a way that impacts on the old-evidence problem.

missing mass themselves, but that they constrained and shaped what was considered a pressing problem and how existing results could reasonably be integrated to address it.[4] The result supports a modest permissivism about theory assessment where distinct evaluations of the degree of support for a hypothesis, given the same evidence, can be *rationally permissible* when they arise from different sets of defensible constraints. Again, this holds even when the parties agree on the available empirical data. A central task, therefore, is to clarify what makes such constraints normatively defensible as opposed to merely heuristic. I suggest that a background constraint is normatively defensible when it:

(i) Is independently motivated within an established modeling context

(ii) Has a track record of guiding empirically successful integration

(iii) Is treated as defeasible, not as dogma[5]

The second concerns the role of unification. The epistemic status of unification is contested in philosophy, so it is worth asking not only if, but also how, the unificatory step taken in the 1974 papers can be characterized as confirmation. To that end, I draw on (Myrvold, 2003, 2017), who distinguish between unification understood as the postulation of a common origin and unification understood as an increase in mutual information between previously separate evidential claims. 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). The introduction of this modeling step not only placed the observations in a shared category, but more significantly reconfigured the inferential dynamics between them. The apparently anomalous dynamics of galaxy clusters and the flatness of galaxy rotation curves became related such that each was informative about the other under the halo hypothesis.[6] On Myrvold's account, the epistemic source of confirmation is located in the reciprocal dependence that emerged — a mutual-information gain — even though the underlying observations themselves were not temporally novel. Normatively, this provides a principled way to distinguish warranted evidential reconfiguration (generating genuine informational support) from unification that is merely heuristic or programmatic.

The current paper's contributions can be stated briefly and in the order required by the episode itself. First, it shows how non-empirical commitments — for example,

[4]This is structurally similar to other versions of indirect or meta-empirical support in the sense that what shifts is not the evidential base itself, but how a successful framework constrains the space of viable hypotheses, for example by (Dawid, 2013, 2019; Dawid et al., 2015).

[5]A constraint is non-normative when it functions only as a preference for a research direction without satisfying (i)–(iii).

[6]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 instead that no stable, widely endorsed link was treated as evidentially significant across domains.

$\Omega \approx 1$, the cosmological principle, and Machian considerations — can be represented in a Bayesian framework as differences in priors and background assumptions. Second, it identifies *evidential reconfiguration* as the central historical mechanism: a change in evidential partitioning and warranted cross-domain dependence rather than the arrival of new observations. Third, it offers a detailed, historically grounded application of Myrvold's mutual-information account of unification to show how that reconfiguration can yield confirmational gain. Fourth, it refines the old-evidence discussion by distinguishing temporal novelty from structural novelty, and by showing how support can increase when already accepted results become mutually constraining within a single warranted framework.

2 The standard narrative

Textbook and popular narratives of dark matter typically begin with Zwicky's Coma cluster work in the 1930s, and then proceed by selecting a small number of later breakthroughs that are presented as progressively establishing the hypothesis incrementally. 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 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.[7]

The next standard step is to emphasize Rubin and Ford's work on rotation curves in the 1970s (Rubin & Ford Jr, 1970), which resulted in the observation that the rotational velocity of matter in spiral galaxies did not decrease with distance from the galactic center. This inaugural measurement was soon followed by a series of observations by Roberts & 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 the decisive vindication of 'dark matter', despite the fact that Rubin herself was cautious about any causal interpretation of her measurements (Rubin, 2004), and many astronomers initially explored explanations that did not require positing a new exotic mass component.

The cumulative story is then rounded out by incorporating later major developments like gravitational lensing measurements, large-scale structure formation, and the Bullet Cluster observation (Clowe et al., 2006). These are presented as extensions of the conceptual Zwicky–Rubin historical lineage, progressively solidifying the epistemic status

[7] 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).

of the hypothesis until it became a deeply entrenched and central tenet of contemporary cosmology. I do not intend to claim that historical presentations of dark matter — typically structured in roughly this format — are illegitimate. They serve important pedagogical purposes for their audiences, relying on hindsight to select which episodes and events are deemed significant, and therefore worth emphasizing: those that connect to the currently held concept and explanatory role of dark matter. The downside is that such histories compress the epistemically prior question of when and why particular results came to function as evidence for missing mass, rather than as local anomalies or modeling artefacts.

The mid–1970s is a particularly illuminating historical episode, identified by Swart et al. (2017) and de Swart (2020) as a pivotal moment in establishing the dark matter hypothesis. The aim is to complement the historical claim *that* this episode was significant with an epistemic account of *why* it was significant. Since the empirical data did not change in 1974, what *did* change that carried epistemic significance?

3 Constraints and the mass deficit

To understand why the dark matter hypothesis gained traction in both cosmology and astronomy, it is necessary to consider developments in cosmology leading up to the 1970s. In the early and mid-20th century, cosmology was often more guided by theory than observation-driven astronomy, and debates frequently appealed to methodological and theoretical considerations (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 partly on philosophical grounds. Coupling the cosmological principle with general relativity led to relativistic cosmology, which by the 1960s had become the predominant cosmological framework, drawing on theoretical work by Alexander Friedmann and Georges Lemaître, and observational work by Edwin Hubble (de Swart, 2020).[8] By the 1950s and 1960s, two non-empirical commitments in particular structured background assumptions in cosmology:

- The cosmological principle: homogeneity and isotropy.
- A Machian heritage: local inertia is determined by the total mass of the universe (a commitment often taken to favor a *closed*, finite universe in relativistic cosmology).

Together, these commitments contributed to a research mode in which empirical results were interpreted through the lens of cosmological models constrained in part for non-empirical reasons.

[8] See Bondi & 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.

> 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 for present purposes is not that such commitments functioned as dogma, but that they often served as boundaries, in the sense that they constrained which questions were treated as central and which lines of integration across domains were worth pursuing. Much like Peebles in the above, Steven Weinberg illustrates the tension between non-empirical commitments and empirical observations as follows:

> 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 2×10^{-29}~g/cm^3 must be found somewhere outside the normal galaxies. But where? (Weinberg, 1972, p. 478)

One can see how the *interpretative lens* through which the empirical observations are viewed act as a constraint. These constraints are revisable in principle, but taking them seriously makes certain problems appear more pressing, as evident in Peebles' "whether" and Weinberg's "where?". By the early 1970s, then, near-critical cosmological models functioned for some authors as a defensible target, and for others as a background commitment with independent non-empirical motivations. This combination created a distinctive context in which 'missing mass' could be pursued as a candidate astronomical solution to a pressing cosmological problem preceding 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 the mass needed to move Ω toward unity. They did so by drawing on two kinds of phenomena that were already well known but typically treated as independent research problems. This independence was not primarily chronological, but disciplinary: the phenomena were investigated in different contexts, with different methods and uncertainties, and there was no standing consensus that they should bear on the same hypothesis in tandem.

1. **Galaxy dynamics.**
   Thanks to radio observations (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 Figure 1, adapted from Roberts & Rots (1973)). Empirically, however, many spiral galaxies showed approximately flat rotation curves — behavior consistent with additional gravitating mass at large radii, but not automatically decisive for any single ontological conclusion. In practice, these results were often treated as local anomalies, whose significance depended on auxiliary assumptions 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, virialisation, and the representativeness of the observed velocities), and 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) combined these results into a single hypothesis-driven analysis. They argued, in effect, that if each galaxy is embedded in a massive halo extending beyond its visible edge, then:

(i) the galaxy's rotation curve 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

(iii) the collective mass in galaxy halos could contribute substantially to the cosmic mass budget, potentially bringing Ω 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, p. L1).

The central epistemic claim of this paper is that the significance of these 1974 studies lies in how they treated the anomalies as evidentially connected. By introducing a shared halo framework (and associated parameters), the two phenomena were placed within a single context in which they were connected. This is the sense in which 1974 marks a shift from isolated anomalies to evidence with a different confirmational profile where each data type begins to bear on expectations about the other under the halo hypothesis, and the resulting evidential picture bears directly on the cosmological mass-budget question that had been made pressing by the background constraints found in cosmology.

Two consequences follow from this reconceptualization. First, it changes the confirmational situation without adding new data. Once the halo framework is in place, rotation curves and cluster dynamics no longer function merely as separate anomalies, but as mutually constraining results whose expected relations under the missing-mass hypothesis can be assessed. This is the setting in which a Myrvoldian increase in mutual information becomes available in principle: the evidential force is generated by newly licensed dependencies between the phenomena, not by the mere postulation of a common origin.

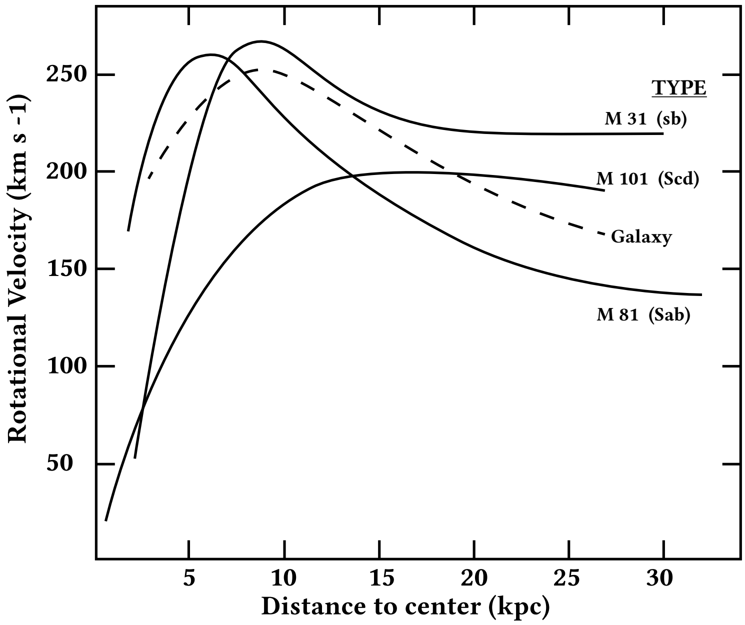


Figure 1: Rotation curves of three spiral galaxies (solid lines) as a function of distance from the center. (The Milky Way's expected Keplerian decline is shown for reference as the dotted line.) The flatness of the curves at large radii indicates the presence of unseen mass.

Second, the same reconceptualization helps explain the divergence between astronomers and cosmologists. For cosmologists, background constraints associated with $\Omega \approx 1$ made the missing-mass hypothesis an especially salient candidate, and they were correspondingly more willing to treat the link between the phenomena as well-motivated. For many astronomers, who did not treat closure considerations as constraining in the same way, the dependencies introduced in 1974 could appear less compelling, more contingent on auxiliary assumptions, or simply premature. The next sections therefore examine, first, how these constraints can be represented within a Bayesian reconstruction of theory assessment, and second, whether the 1974 synthesis is confirmatory in Myrvold's sense of mutual-information unification rather than merely a common-cause unification.

4 Permissible community disagreement

I now turn to the first of the two philosophical tasks set out above. An odd aspect of the 1974 synthesis papers is the divided reception they received in the relevant scientific communities. Astronomers and cosmologists agreed on much of the observational data as such, yet they differed in how strongly that data supported a missing-mass hypothesis. This section offers a Bayesian reconstruction of that divergence, with the modest normative aim to show how disagreement can be rationally permissible when it arises from differing, defensible background constraints, rather than from ignorance or irrationality.

4.1 Evidence and background

A feature of the historical record warrants closer attention. Astronomers and cosmologists in the 1970s (and into the early 1980s) differed in how willing they were to treat dark matter as the appropriate conclusion, despite substantial overlap in what they took the salient results to be. In her retrospective *A Brief History of Dark Matter*, Rubin recalls that many astronomers were initially reluctant to accept that the rotation-curve anomaly required unseen mass:

> 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, p. 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, p. 7)

It is tempting to cast the divergence as a difference in scientific temperament. But the philosophically relevant point is that evidential support is assessed relative to background assumptions and modeling commitments. When those differ across communities, the same empirical data can legitimately be taken to bear more — or less — strongly on the same hypothesis. For many astronomers, the central epistemic question was whether the data, together with accepted dynamical and astrophysical auxiliaries, warranted a robust inference to a new mass component on galactic scales. This made questions about systematics, tracer assumptions, and the adequacy of particular dynamical models especially salient. In that methodological setting, it was appropriate to explore whether revisions to auxiliary assumptions, refinements of modeling, or alternative dynamical strategies might accommodate the anomalies without committing to extensive unseen matter.

Cosmologists, by contrast, were often engaged in a different project, namely the assessment of global cosmological models and parameters, including the relation between observed structure and the universe's mass density. Within that project, background constraints — including a preference for a closed universe — made the mass-budget question pressing in a way it was not for many practicing astronomers. This difference does not itself supply additional evidence for missing mass. Rather, it shapes (i) which questions are treated as central, and (ii) which ways of integrating results across domains are treated as methodologically worthwhile. The resulting divergence is therefore a plausible instance of *permissible disagreement* where distinct assessments can be rationally permissible when they arise from different, defensible sets of background constraints, even when there is substantial agreement about the observational results.

4.2 Cosmologists' priors

A Bayesian reconstruction makes the above structure explicit. Let h denote a dark matter hypothesis (e.g., that galaxies are embedded in massive halos), and let e denote the relevant empirical results (e.g., rotation curves and cluster dynamics). Crucially, Bayesian confirmation is always relative to background assumptions b, which informs modeling commitments and auxiliary assumptions that mediate the connection between observation and hypothesis. The relevant quantity is therefore $Pr(h \mid e, b)$ rather than $Pr(h \mid e)$. For clarity, Bayes' theorem can be written as:

$$\begin{aligned} Pr_{\text{new}}(h) \;=\; Pr(h \mid e, b) &= \frac{Pr(e \mid h, b)\, Pr(h \mid b)}{Pr(e \mid b)} \\ &= \frac{Pr(e \mid h, b)\, Pr(h \mid b)}{Pr(h \mid b)\, Pr(e \mid h, b) + Pr(\neg h \mid b)\, Pr(e \mid \neg h, b)} \end{aligned} \tag{4.1}$$

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 $Pr(h \mid b)$ 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. Importantly, this does not mean such considerations function as evidence for h. Rather, 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 record, 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 by 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 is helpful to isolate a simplified case. Suppose that, relative to a broadly shared background b concerning mechanics and measurement practices, both communities agree that h renders the results comparatively expected, so that $Pr(e \mid h, b)$ is high, while $Pr(e \mid \neg h, b)$ is comparatively low. Then the evidential force of e for h is captured by the Bayes factor:

$$B = \frac{Pr(e \mid h, b)}{Pr(e \mid \neg h, b)} \tag{4.2}$$

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 $Pr(h \mid b)$.

This is also where the old-evidence complication becomes relevant. If e is already part of an agent's informational state, then standard conditionalization does not treat e as producing a further update, since one is already effectively working with $Pr(\cdot \mid e, b)$.[9] The present point is narrower because even when no further *formal* update is available, differences in $Pr(h \mid b)$ can still structure how agents treat a hypothesis as promising, how strongly they take it to be supported by the total evidential situation, and which lines of further work they regard as worth pursuing.

Historically, closure considerations plausibly functioned as a non-empirical constraint for the prior. A preference for $\Omega = 1$ did not supply new observational evidence for halos, but it did make additional gravitating mass a comparatively plausible solution, and it made cross-domain integration (from galactic dynamics to the cosmic mass budget) methodologically prominent. 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.

This divergence plausibly contributed to the historical reception of the dark matter hypothesis. Astronomers and cosmologists were evaluating h relative to different background constraints, and therefore differed both in prior credence and in what they

[9] One (classic) formulation of the old-evidence problem is that if $Pr(e \mid b) = 1$ for an agent, then conditionalizing on e leaves credences unchanged, so e cannot raise $Pr(h \mid b)$ in the usual way. I return to this issue in Section 6.

regarded as an appropriate evidential integration. The lesson is not that “anything goes” at the level of priors, but that theoretical and methodological commitments can legitimately enter the context of theory assessment through b, thereby shaping priors and guiding what kinds of evidential connections are treated as warranted.

Normatively, the episode supports a moderately permissive but constrained view of non-empirical considerations in scientific reasoning. Cosmologists’ increased antecedent confidence in additional mass was anchored in a successful relativistic hot Big Bang framework, 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 $\approx$ 10–20% of the critical value. In that sense, their stance can be understood as a methodologically risky, but epistemically permissible one characterized as a constraint-guided commitment within an established research framework. On the above criteria, closure considerations were defensible insofar as they were embedded in the larger context of successful relativistic cosmology and operated defeasibly as a constraint on admissible global models, as opposed to serving as a substitute for evidence.

The remaining question is whether the 1974 synthesis merely offered a common-cause proposal, or whether it generated a genuine confirmational surplus by reconceptualizing the evidential relations between the relevant results. That question is taken up in the next section through Myrvold’s distinction between common-origin unification and mutual-information unification.

5 Confirmatory unification

The 1974 papers by Ostriker et al. (1974) and Einasto et al. (1974) are considered as a turning point because they *unified* disparate results under a common halo framework. Contemporaries and later commentators treated this move as epistemically significant, not merely as a convenient re-description. 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, p. 3)

Tremaine similarly characterized the connection as a watershed in galaxy dynamics and cosmology (Tremaine, 1999, p. 1223). The philosophical question is whether the *unificatory* aspect of the 1974 move was itself confirmatory and, if so, in what sense. This is distinct from asking if unification is an explanatory virtue, since the question pursued here is whether, and how, unification can generate a *confirmational* surplus in a Bayesian sense. To answer this, it is useful to distinguish two conceptions of unification that are often run together in discussions:

(i) **common-origin unification** (COU): distinct phenomena are subsumed under a single underlying cause or mechanism.

(ii) **mutual-information unification** (MIU): a hypothesis makes previously separate results evidentially *interdependent*.

The central claim of this section is that COU, by itself, is not automatically confirmatory in a Bayesian sense. MIU, by contrast, isolates a precise source of confirmational surplus.

5.1 Common-origin unification

Much of the philosophical literature treats unification as an explanatory virtue. The guiding thought is that if two theories accommodate the same total evidence, but one does so with a more integrated explanatory scheme, scientists often regard it as better supported. This outlook is articulated in different ways by among others Kitcher (1989), Lipton (2003), and (Psillos, 1999, 2009). On this construal, the virtue is primarily structural since it concerns the organisation and economy of explanations, rather than the introduction of new empirical content.

The difficulty is that, on a strict Bayesian view, confirmational relations are fixed by priors and likelihoods relative to background assumptions. 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 guide to theory construction and pursuitworthiness. This motivates the skeptical line associated with strict Bayesianism (e.g., Howson & Urbach (2006); Hartmann & Sprenger (2011)): unification may be methodologically fruitful, but it does not automatically increase $Pr(h \mid e, b)$. This point becomes clearer once we separate two roles that are easily conflated:

- **COU as an explanatory claim.**
  A hypothesis posits a common origin for what had previously been treated as distinct phenomena. This can be intellectually satisfying and heuristically valuable.

- **COU as a confirmational claim.**
  The same hypothesis is said to receive an 'evidential boost' *because* it unifies.

The first claim does not entail the second. A common-origin proposal can be compatible with the data while remaining comparatively unconstrained; or it can be compatible because it has been tailored to fit; or it can be compatible without changing how the relevant results bear on one another. In such cases, the proposal may guide research without producing any non-trivial increment of confirmation. That is why COU, by itself, is normatively ambiguous. It may signal *theoretical promise*, but it does not yet identify a principled *increase* in evidential support. What is needed, if unification is to be confirmatory, is some change in the evidential situation that affects the likelihoods — typically by introducing constraints that make distinct results bear on one another.

The dark matter case therefore constitutes a useful test. If the 1974 unification was confirmatory in a robust sense, the source of that confirmation must be something more specific than the generic fact that a common mechanism was proposed.

5.2 Mutual-information unification

Myrvold's account supplies the missing precision. In (Myrvold, 2003, 2017), unification is not characterized by common origin, but by a change in the *dependence structure* of a body of evidence. A hypothesis unifies, in Myrvold's sense, when it makes distinct evidential claims informative about one another.

- Let $\mathbf{e} = \{p_1, p_2, \ldots, p_n\}$ be a body of evidence.
- Let $I(p; q \mid X)$ denote the mutual information between propositions p and q given background X.[10]
- If p_1 and p_2 are (approximately) independent given b, so that $I(p_1; p_2 \mid b) \approx 0$, but become informatively linked under h relative to b, so that $I(p_1; p_2 \mid h \wedge b) > 0$, then h renders the phenomena mutually informative.

Myrvold captures the unificatory contribution by defining:

$$U(p_1, p_2; h \mid b) \;=\; I(p_1; p_2 \mid h \wedge b) \;-\; I(p_1; p_2 \mid b) \tag{5.1}$$

When $I(p_1; p_2 \mid b) \approx 0$, a positive U is simply $I(p_1; p_2 \mid h \wedge b)$ meaning that the hypothesis generates an evidential connection that is not present given the background alone. To connect this to confirmation, it is helpful to use an information-theoretic measure of evidential support (equivalently, a log Bayes factor). Let:

$$C(h; e \mid b) = \log \frac{Pr(e \mid h, b)}{Pr(e \mid \neg h, b)} \tag{5.2}$$

Myrvold shows that, under appropriate assumptions, the support provided by a conjunction $e_1 \wedge e_2$ can be decomposed into the support from e_1 alone, the support from e_2 alone, and an additional term corresponding to unification:

$$C(h; e_1 \wedge e_2 \mid b) \;=\; C(h; e_1 \mid b) \;+\; C(h; e_2 \mid b) \;+\; U(e_1, e_2; h \mid b) \tag{5.3}$$

The normative interest of MIU is that it identifies a principled sense in which unification can yield an evidential surplus. If $U > 0$, the conjunction provides more support than the sum of separate supports, because the hypothesis makes the evidential claims mutually constraining. COU, by itself, does not secure any such result.

[10] Intuitively, mutual information measures how much learning p reduces uncertainty about q, given X.

5.3 Evidential reconfiguration in 1974

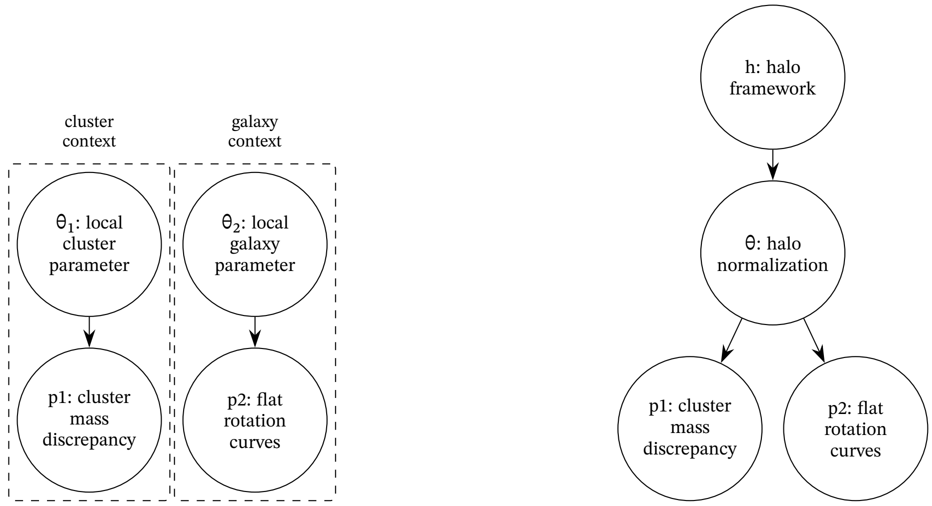
The question, then, is whether the dark matter episode fits the MIU template. For the present case, let:

h : the halo-based dark matter hypothesis (substantial unseen mass in extended halos).
p_1 : the galaxy cluster dynamics discrepancy (mass-to-light anomaly / virial mass deficit).
p_2 : flat galaxy rotation curves at large radii.
b : background assumptions (e.g., GR, ordinary matter, measurement conventions).
θ : a halo parameter (schematically: halo-to-luminous mass normalisation).

Prior to 1974, the salient point is not merely that p_1 and p_2 were already known, but that they were typically assessed in different contexts and with different local auxiliaries. For present purposes, it is therefore a reasonable idealization to model them as (approximately) independent relative to the prevailing background b: $I(p_1; p_2 \mid b) \approx 0$.[11] In that setting, learning p_1 does not materially constrain expectations about p_2, and vice versa. The 1974 move changes the epistemic situation because h does more than posit a common cause in the abstract. It introduces a *shared parameterization* (schematically represented by θ) that links the two results within a single modeling framework: under $h \wedge b$, both p_1 and p_2 depend on the same halo-normalization assumptions. As a result, learning either phenomenon constrains θ, which in turn constrains what should be expected of the other. This is exactly the MIU pattern: $I(p_1; p_2 \mid h \wedge b) > 0$, and hence $U(p_1, p_2; h \mid b) > 0$.
The point can be stated without information theory. Within a halo framework, flat rotation curves constrain the magnitude and distribution of additional gravitating mass associated with galaxies, but those same halo parameters also constrain the contribution of galaxies to the mass budget of 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 newly defensible two-way evidential connection. This also serves to clarify how the episode relates to old evidence.

If p_1 and p_2 are already in an agent's informational state, then standard conditionalization will not treat them as producing a further update. What can still be epistemically new is a newly supported claim about how they hang together under h relative to b. In that sense, the confirmational increment is associated with *structural* novelty, because what is tested is not merely that $p_1 \wedge p_2$ is true, but whether their relation is the sort of relation one should expect under a shared, parameterized halo framework. Call an epistemic change structurally novel when the novelty resides in an endorsed dependence constraint among already accepted propositions — i.e., when what is added is not new

[11] This is an idealization of 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.



(a) Pre-1974: each phenomena in their distinct modeling contexts.

(b) Post-1974 under $h \wedge b$: coupled dependency via θ.

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 θ, which makes p_1 and p_2 mutually informative. Arrows represent probabilistic dependence relations in the modeling framework, not causal influence from observations to θ.

evidence of the first-order kind, but a new, assessable claim about how accepted results should co-vary or constrain parameters under a hypothesis. In the present context, evidential reconfiguration is the historical process (reassigning propositions between background and evidence and warranting connections between domains), whereas structural novelty is the epistemic outcome when that process results in a substantive dependence claim of a certain form (for example 6.3).

The upshot is that MIU allows one to justify — in a restricted and explicitly Bayesian sense — the common judgement that the 1974 connection was confirmatory. The increase in confirmation arises from a reconception of the evidential situation that makes distinct results mutually informative by subjecting them to shared parameter constraints. Put negatively we can say that the 1974 synthesis would *fail* to be confirmatory in the MIU sense if the shared parameterization were too unconstrained to generate testable interdependence, or if a comparable connection could be recovered under $\neg h$ via alternative auxiliaries.

The next step is to make this application more exact by stating what, in the present setting, would be sufficient for $I(p_1; p_2 \mid h \wedge b) > 0$, and by clarifying which modeling assumptions are treated as part of h and which remain in b. That will also allow us to say more precisely what it would take for the 1974 synthesis *not* to be confirmatory in the MIU sense — for example, if the shared parameterization were too unconstrained to generate testable interdependence.

6 Old evidence without ignorance

The preceding section argued that the 1974 synthesis can be confirmatory in Myrvold's MIU sense, because the halo framework renders previously separate results mutually informative. A natural worry remains. MIU is formulated within Bayesian confirmation theory, and Bayesian confirmation is often said to become inert when the relevant evidence is already accepted. This is the classic *problem of old evidence*, associated with Glymour (1981): if an agent assigns probability 1 to an evidential proposition e, then conditionalizing on e cannot change their credences, since $Pr(h \mid e) = Pr(h)$. How, then, can a hypothesis be confirmed by facts that were already established — and, in the present case, facts that were explicitly mobilized in motivating the hypothesis? Two clarifications are essential.

1. In practice, *old evidence* does not necessarily mean 'already known'. It refers to evidence whose *epistemic role* has stabilized — either because it has been treated as fixed background in a context of assessment, or because it has already been absorbed into the standing conditional probability state.

2. The 1974 episode is not plausibly described as anyone learning p_1 and p_2 anew. What changes is the *evidential partition*, i.e., which propositions are treated as background, which are treated as evidence for which hypotheses, and which dependence relations are taken to be warranted within an accepted modeling framework. This is exactly the kind of shift MIU is meant to represent.

The aim of this section is therefore to move away from the idea that already known evidence is confirmationally inert, and to provide a Myrvold-consistent resolution. The epistemically novel element in 1974 is not the conjunction $p_1 \wedge p_2$, but a newly warranted *structural* claim about how p_1 and p_2 are related under a shared parameterization. A tempting formulation of the problem begins from Bayes' theorem:

$$Pr(h \mid e, b) = \frac{Pr(e \mid h, b)\, Pr(h \mid b)}{Pr(e \mid b)} \tag{6.1}$$

If e is fully accepted, one might then write $Pr(e \mid b) = 1$ and conclude that no update is possible. But this diagnosis is misleading in two ways. First, the key feature of *old evidence* is not simply that an agent assigns probability 1 to e, but that e is already part of the agent's standing conditional information state. In other words, one is already working with probabilities of the form $Pr(\,\cdot \mid e, b)$, so there is no further conditionalization event available that corresponds to 'learning' e.

Second, the old evidence problem is not solved by inserting numerals into Bayes' theorem. The representational question is how to model epistemic change when a new hypothesis is introduced *after* the relevant evidence is in place, or when accepted results are reclassified and recombined in a way that changes what they bear on. The dark matter episode is primarily of the second kind simple because it concerns evidential roles and dependence, not the arrival of new observational propositions. For that reason, strategies that simulate ignorance by temporarily removing e from the background (e.g., forgetful functions) do not describe the historical phenomenon of interest. They may be useful as formal devices in some contexts, but they mischaracterize what is significant in the 1974 episode, where the community did not lose and regain p_1 and p_2.

6.1 Structural novelty as evidence

A more faithful picture begins with a distinction between *possession* of a result and its *evidential role*. A community may possess a result (it is published, cited, treated as established) and yet not treat it as central evidence for a given hypothesis. Conversely, a result may play different roles across contexts: background in one setting, evidence in another. This distinction captures the pre-1974 situation. The phenomena corresponding to p_1 (cluster dynamics discrepancies) and p_2 (flat rotation curves or related galactic anomalies) were known, but their epistemic significance was typically restricted to local domains. They were not generally treated as jointly constraining a single parameterized hypothesis about halos in a way that mattered for cosmology. In this sense, the results were not integrated into a single cohesive evidential picture.

The 1974 synthesis can therefore be described as a kind of *re-activation*, where results that had been treated as local anomalies — or as settled background within distinct problem-settings — were reclassified as jointly relevant within a single modeling framework. What is 'new' is not the observational propositions themselves, but that they were connected. This is the intended content of *evidential reconfiguration*: a reassignment of propositions between background and evidence, together with the endorsement of new dependence relations across domains. Myrvold's MIU framework allows this change to be represented without simulating ignorance. The key point is that MIU concerns the mutual information between evidential propositions under a hypothesis:

$$U(p_1, p_2; h \mid b) = I(p_1; p_2 \mid h \wedge b) - I(p_1; p_2 \mid b) \tag{6.2}$$

Even if both p_1 and p_2 are already accepted, the *support* that a body of evidence provides for a hypothesis depends on how probable that body is under competing hypotheses. What 1974 contributes is a reason to treat p_1 and p_2 as connected under $h \wedge b$ through a shared parameterization (θ). This yields $I(p_1; p_2 \mid h \wedge b) > 0$ even when $I(p_1; p_2 \mid b) \approx 0$. From this perspective, the epistemically novel element is not the proposition $p_1 \wedge p_2$, but a structural claim about dependence, along the lines of:

$$S : I(p_1; p_2 \mid h \wedge b) > 0 \text{ via a shared parameter } \theta \tag{6.3}$$

S, together with a model-based explanation of *why* this holds, namely that both results are constrained by the same halo parameter(s) θ within the halo framework. Equivalently, we might say that under $h \wedge b$, evidence about rotation curves constrains what should be expected about cluster dynamics, and vice versa. This dependence claim S can itself be evaluated because it can succeed or fail depending on whether the shared parameterization genuinely generates cross-domain constraints, or whether comparable coherence can be recovered without h (for example, via alternative auxiliaries).

In this sense, the episode manifests a case of **structural novelty**: a change in the dependence structure among accepted evidential claims, not a change in which first-order observational propositions are known. This also clarifies why standard *temporal novelty* and *use novelty* diagnoses are too blunt for the present case. The epistemic gain is located neither in the timing of the observations nor in their role in motivating h, but in the fact that h supports *relations among results* — relations that were not previously taken to be warranted across domains.

6.2 MIU and old evidence

The dark matter case does not resolve the old evidence problem in the sense of providing a generalized Bayesian model. What it shows is that the standard claim that old evidence is inert mischaracterizes cases in which the epistemic change concerns evidential role and dependence, rather than the arrival of new observational content. A Myrvold-consistent resolution can be stated succinctly:

- The relevant evidential object is not merely the conjunction $\{p_1, p_2\}$, 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 successfully exploiting the dependence claim S — that h makes the two phenomena mutually informative.

In this way, MIU provides a principled route for treating the 1974 synthesis as confirmatory, without invoking forgetful devices or simulated ignorance. The epistemic change is located in a reconception of evidential relations that made cross-domain integration methodologically salient and epistemically assessable, precisely where the history suggests it is located. The concluding section will then be in a position to summarize what this episode shows about evidence creation in science: the legitimate 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 guiding aim of this paper has been to use a historically pivotal episode in cosmology to clarify a more general point about evidence in persistently underdetermined research programmes: observations do not acquire evidential force simply by being accumulated, but by being *made evidential* through changes in modeling frameworks, background constraints, and warranted relations of dependence. The establishment of dark matter in the mid–1970s is instructive because its decisive turning point did not consist in the arrival of new measurements. Rather, it consisted in a reconception of already familiar results that altered what could reasonably be treated as evidence for what. The analysis supports four conclusions.

First, the 1974 episode illustrates a mechanism of *evidential reconfiguration*. Prior to 1974, discrepancies in galaxy dynamics and cluster dynamics were largely treated as separate, domain-local anomalies. The halo framework introduced by Ostriker et al. (1974) and Einasto et al. (1974) did not merely place these anomalies side by side. It imposed a shared parameterisation under which each phenomenon began to constrain expectations about the other. Epistemically, what changed was the partition between background and evidence, and the cross-domain dependence relations that were treated as warranted. The resulting evidential situation thus differed from the bare conjunction $(p_1 \wedge p_2)$ because it was a structured set of claims linked by newly endorsed constraints.

Second, the episode clarifies the *legitimate role of non-empirical (meta-empirical) constraints* in theory assessment. The preference for $\Omega \approx 1$ and related cosmological principles did not function as evidence for dark matter in any straightforward sense. Their epistemic role was instead to operate as constraints which shaped what counted as a salient problem, what qualified as a serious candidate solution, and which cross-domain integrations were methodologically appropriate to pursue. In Bayesian terms, such constraints enter through the background b and thereby influence both priors $Pr(h \mid b)$ and the space of admissible likelihood judgement. The normative lesson is modest but non-trivial. Non-empirical commitments can be epistemically legitimate when they are defensible within an established framework, and when they help determine which evidential dependencies scientists ought to treat as warranted in that framework.

Third, the case provides a historically grounded instance of *rational (permissible) disagreement*. Astronomers and cosmologists could agree on much of the observational record while diverging in their assessment of a missing-mass hypothesis, because they worked under different background constraints and differed in their judgements about which cross-domain links were warranted. This need not be reduced to irrationality, bias, or mere epistemic luck. Rather, it supports a constrained permissivism: given defensible differences in background assumptions and modeling priorities, distinct posterior attitudes toward the same hypothesis can be rationally permissible even under shared evidence. The episode also indicates what a normative evaluation must address: not merely that priors differ, but whether the relevant background constraints are defensible and appropriately connected to established practice. Here 'defensible' means independently motivated within the operative framework, methodologically fruitful with some empirical track record, and taken as defeasible.

Fourth, the analysis clarifies why *MIU captures the confirmational source* of the 1974 unification. Unification understood merely as a common-origin proposal is not automatically confirmatory within Bayesian confirmation theory. Myrvold's mutual-information unification identifies a principled sense in which unification can yield confirmational surplus whereby a hypothesis receives additional support when it increases the mutual information among evidential claims, thereby making them mutually informative. The 1974 halo framework is a plausible instance of this structure. It renders p_1 and p_2 probabilistically dependent relative to $h \wedge b$, so that each result constrains expectations about the other. This explains how unification could be experienced as confirmational without appealing to new data. It also clarifies why the old evidence problem is not well approached here by simulating ignorance with 'forgetful functions' — the epistemically novel element is the structural dependence claim, not the re-learning of already accepted propositions.

Taken as a whole, these lessons complicate the textbook picture according to which dark matter became established by straightforward accumulation of evidence. The episode shows that evidence-making involves historically situated choices about modeling, background constraints, and cross-domain couplings; and it shows that Bayesian tools, suitably interpreted, can represent such choices without collapsing them into arbitrariness. The dark matter case therefore offers both a *historical correction* and a *philosophical payoff*. It provides a concrete instance in which evidential reconfiguration — rather than new data — drives a genuine change in confirmational standing. More precisely, it shows how support can increase without new observations when existing results become mutually constraining under a warranted framework.

References

Allzén, S. (2021). Scientific realism and empirical confirmation: A puzzle. *Studies in History and Philosophy of Science Part A*, *90*, 153–159.

Allzén, S. (2024). *Dark Matter: Explanatory Unification and Historical Continuity*. arXiv. https://doi.org/10.48550/arXiv.2412.13404

Antoniou, A. (2023). Robustness and Dark-Matter Observation. *Philosophy of Science*, *90*(3), 629–647.

Antoniou, A. (2025). Why did the dark matter hypothesis supersede modified gravity in the 1980s? *Studies in History and Philosophy of Science*, *112*, 141–152.

Bondi, H., & Gold, T. (1948). The steady-state theory of the expanding universe. *Monthly Notices of the Royal Astronomical Society*, *108*(3), 252–270.

Bosma, A. (1978). *The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types* [{PhD} {Thesis}]. Rijksuniversiteit te Groningen.

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C., & Zaritsky, D. (2006). A Direct Empirical Proof of the Existence of Dark Matter. *The Astrophysical Journal Letters*, *648*(2), L109–L113. https://doi.org/10.1086/508162

Dawid, R. (2013). *String theory and the scientific method*. Cambridge University Press.

Dawid, R. (2019). The significance of non-empirical confirmation in fundamental physics. *Why Trust a Theory? Epistemology of Modern Physics*, 99–119.

Dawid, R., Hartmann, S., & Sprenger, J. (2015). The no alternatives argument. *The*

British Journal for the Philosophy of Science.

De Baerdemaeker, S. (2021). Method-driven experiments and the search for dark matter. *Philosophy of Science*, *88*(1), 124–144.

De Baerdemaeker, S., & Boyd, N. M. (2020). 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.

De Baerdemaeker, S., & Dawid, R. (2022). MOND and meta-empirical theory assessment. *Synthese*, *200*(5), 1–28.

de Swart, J. (2020). Closing in on the cosmos: Cosmology's rebirth and the rise of the dark matter problem. In A. S. Blum, R. Lalli, & J. Renn (Eds.), *The renaissance of general relativity in context* (pp. 257–284). Springer International Publishing. https://doi.org/10.1007/978-3-030-50754-1_8

Duerr, P. M., & Wolf, W. J. (2023). Methodological reflections on the MOND/dark matter debate. *Studies in History and Philosophy of Science*, *101*, 1–23. https://doi.org/https://doi.org/10.1016/j.shpsa.2023.07.001

Einasto, J., Kaasik, A., & Saar, E. (1974). Dynamic evidence on massive coronas of galaxies. *Nature*, *250*(5464), 309–310.

Glymour, C. (1981). Theory and evidence. *Philosophy of Science*, *48*(3).

Hartmann, S., & Sprenger, J. (2011). Bayesian epistemology. In *The Routledge companion to epistemology* (pp. 609–619). Routledge.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach*. Open Court Publishing.

Jacquart, M. (2021). Λ CDM and MOND: A debate about models or theory? *Studies in History and Philosophy of Science Part A*, *89*, 226–234.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World. In P. Kitcher & W. Salmon (Eds.), *Scientific Explanation* (pp. 410–505). Minneapolis: University of Minnesota Press.

Lipton, P. (2003). *Inference to the best explanation*. Routledge.

Martens, N. C. (2022). Dark matter realism. *Foundations of Physics*, *52*(1), 16.

Martens, N. C., Sahuquillo, M. A. C., Scholz, E., Lehmkuhl, D., & Krämer, M. (2022). Integrating dark matter, modified gravity, and the humanities. *Studies in History and Philosophy of Science*, *91*, A1–A5.

Merritt, D. (2021a). Cosmological realism. *Studies in History and Philosophy of Science Part A*, *88*, 193–208.

Merritt, D. (2021b). MOND and Methodology. In *Karl Popper's Science and Philosophy* (pp. 69–96). Springer.

Milne, E. A. (1932). World structure and the expansion of the universe. *Nature*, *130*(3270), 9–10.

Myrvold, W. C. (2003). A Bayesian account of the virtue of unification. *Philosophy of Science*, *70*(2), 399–423.

Myrvold, W. C. (2017). On the evidential import of unification. *Philosophy of Science*, *84*(1), 92–114.

Ostriker, J. P., Peebles, P. J. E., & Yahil, A. (1974). The Size and Mass of Galaxies, and the Mass of the Universe. *The Astrophysical Journal*, *193*, L1–L4. https://doi.org/10.1086/181617

Psillos, S. (1999). *Scientific realism: How science tracks truth*. Routledge.

Psillos, S. (2009). *Knowing the structure of nature: Essays on realism and explanation*. Palgrave Macmillan.

Roberts, M., & Rots, A. (1973). Comparison of rotation curves of different galaxy types. *Astronomy and Astrophysics*, *26*, 483–485.

Rubin, V. C. (2004). A brief history of dark matter. In M. Livio (Ed.), *The Dark Universe: Matter, Energy and Gravity* (pp. 1–13). Cambridge University Press. https://doi.org/10.1017/CBO9780511536298.003

Rubin, V. C., & Ford Jr, W. K. (1970). Rotation of the Andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, *159*, 379.

Shapiro, S. L. (1971). The density of matter in the form of galaxies. *The Astronomical Journal*, *76*, 291.

Sus, A. (2014). 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.

Swart, de, Bertone, G., & Dongen, J. van. (2017). How dark matter came to matter. *Nature Astronomy*, *1*(3), 1–9.

Tremaine, S. (1999). Comments on Ostriker, Peebles, & Yahil (1974)" The Size and Mass of Galaxies, and the Mass of the Universe". *The Astrophysical Journal*, *525*, 1223.

Vanderburgh, W. L. (2003). The dark matter double bind: Astrophysical aspects of the evidential warrant for general relativity. *Philosophy of Science*, *70*(4), 812–832.

Vanderburgh, W. L. (2005). The methodological value of coincidences: Further remarks on dark matter and the astrophysical warrant for general relativity. *Philosophy of Science*, *72*(5), 1324–1335.

Vanderburgh, W. L. (2014a). 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.

Vanderburgh, W. L. (2014b). Quantitative Parsimony, Explanatory Power and Dark Matter. *Journal for General Philosophy of Science / Zeitschrift Für Allgemeine Wissenschaftstheorie*, *45*(2), 317–327. https://doi.org/10.1007/s10838-014-9261-9

Vaynberg, E. (2024). Realism and the detection of dark matter. *Synthese*, *204*(3), 82.

Weinberg, S. (1972). *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley & Sons.

Wolf, W. J., & Read, J. (2025). *Navigating permanent underdetermination in dark energy and inflationary cosmology*. https://arxiv.org/abs/2501.13521

Zwicky, F. (1933). How Far Do Cosmic Rays Travel? *Phys. Rev.*, *43*(2), 147–148. https://doi.org/10.1103/PhysRev.43.147

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. *The Astrophysical Journal*, *86*, 217.