

Consistent Standards of Evidence for Miracle Claims — 23

Phil Stilwell
Independent Scholar
ORCID: [0009-0009-3725-8682](#)
Revised August 2026

Central claim

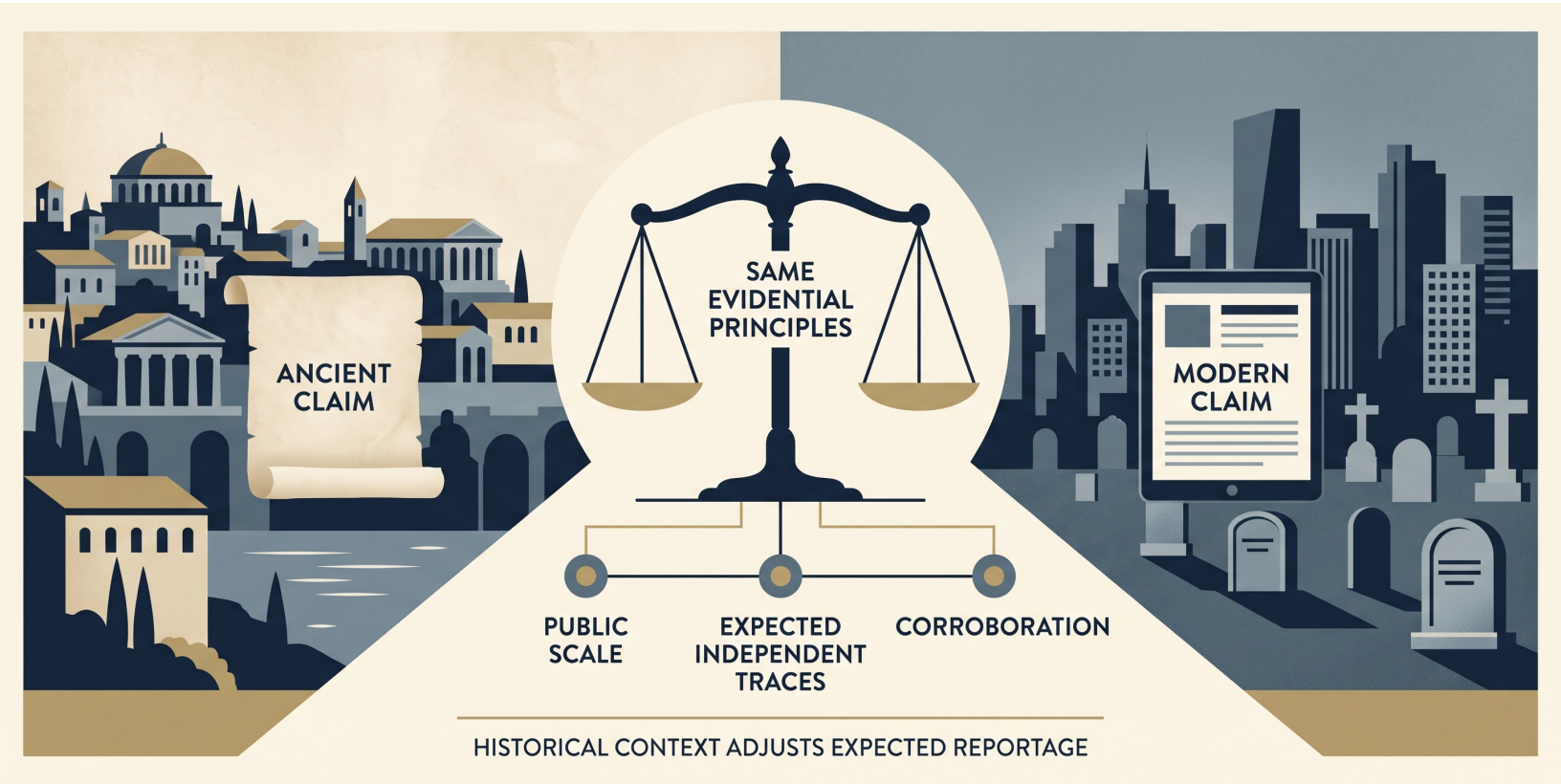
Matthew’s report that many resurrected saints entered Jerusalem and appeared to many is not established by consistent public historical standards. Ancient and modern public miracle reports should face the same question—whether the full surviving record is more expected if the event occurred than if it did not—while each era’s recording and survival conditions adjust the expected traces. Matthew’s lone extant attestation, unnamed witnesses, and lack of independently identifiable corroboration lower historical confidence; ancient record loss limits the size of that reduction, so the argument does not prove the episode false.

Preview:

Target claim	Parity rule	Observed pressure	Rational conclusion
Many resurrected saints entered Jerusalem and appeared to many, as Matthew reports.	Use the same update rule across eras; place each era’s recording and survival conditions inside it.	One extant literary attestation names no witnesses to the saints, and no independent trace is known.	The record lowers confidence and leaves the episode unestablished; ancient loss limits the reduction, so this is not a deductive disproof.

Consistency means parity of method, not an identical document quota for first-century Jerusalem and modern New York.

Visual preview:



The parchment and digital report enter the same balanced evidential rule, while their settings yield different expectations about surviving traces. The graphic states parity of method, not equal record survival or automatic disproof from silence.

Abstract

Ancient public miracle reports should not receive a weaker evidential standard merely because they are old or sacred. This paper argues that Matthew 27:52–53, where many raised saints are said to enter Jerusalem and appear to many, is not established by consistent public evidential standards. The governing rule is comparative: ask whether the full surviving record is more expected if the reported event occurred than if it did not, while placing each period's recording practices, witness access, transmission, and record survival into the background conditions. A modern cemetery-resurrection thought experiment makes the rule's parity and its historical asymmetry visible. A modern event should produce far more traces, so modern silence is stronger evidence. Yet Matthew's large public claim also raises an expectation of some independent trace. What survives is one literary attestation, no named witnesses to the saints, and no independently identifiable corroboration. That mismatch lowers historical confidence to a context-dependent degree rather than proving the episode false. The argument distinguishes testimonial warrant from source multiplication and answers the strongest replies about intended genre, lost records, oral transmission, theological background, and the wider cumulative resurrection case. Those replies can alter the interpretive framing, a prior, the expected survival of records, or the paper's scope, but they do not supply the missing independent support. The conclusion concerns this episode's public historical warrant; it does not disprove every miracle claim, the resurrection of Jesus, Christianity, or theism.

Keywords: miracle claims; evidential parity; testimony; expected traces; argument from silence; Matthew 27; Bayesian confirmation

1. The Double Standard Under Pressure

Some Christian readers and apologists distrust a modern report of a public miracle unless it comes with named witnesses, independent records, and evidence that ordinary causes cannot explain. The same readers may accept an ancient report embedded in scripture even when its witnesses are unnamed and independent traces are absent. The point is not that every Christian reasons this way. Serious defenders of the resurrection build cumulative cases from testimony, background beliefs, and explanatory comparison rather than treating antiquity itself as evidence (McGrew & McGrew, 2009; Swinburne, 2003). The target is narrower: the move that allows historical distance or sacred status to lower a standard that would remain demanding for a comparable modern claim. This paper argues that the exemption fails: historical conditions change the traces we should expect, but they do not change the rule by which those traces bear on a claim.

The governing thesis is therefore direct. Calendar age is not confirming evidence. A claim from antiquity may face a thinner expected record because fewer people could write, fewer documents were produced, and far less survives. Those differences matter, but they belong inside the evidential assessment; they do not license a different logic. The same question governs both cases: is the record we possess more expected if the event occurred than if it did not?

This paper uses *miracle claim* for a report that an event occurred through extraordinary agency or contrary to the stable course of nature. It does not assume that every miracle claim has the same prior plausibility. A *God-concept* is a specified conception of a deity. A Christian God-concept that includes both the capacity and reason to act in history may assign a miracle a different prior than a naturalistic background does. That disagreement is real, but it does not erase the need to update consistently on testimony, dependence, and missing traces.

The focal ancient example is Matthew 27:52–53. The passage says that many bodies of the saints were raised, came out of their tombs after Jesus' resurrection, entered the holy city, and appeared to many. The paper does not attempt a complete historical judgment on the resurrection of Jesus. It asks what consistent standards imply for this additional, unusually public episode.

The parity principle

Use one evidential update rule for ancient and modern reports. Adjust the inputs for literacy, publicity, record-making, transmission, and survival; do not turn age, sacred placement, or cultural familiarity into an evidential bonus. If missing traces would count against a claim when those traces were reasonably expected, they still count against an ancient claim—only with a strength proportionate to ancient conditions.

This preserves the paper’s challenge without caricaturing historical inquiry. Equal treatment does not mean demanding smartphone video from first-century Jerusalem. It means refusing to let a claim evade scrutiny simply by moving it farther into the past.

2. Consistency Is a Rule, Not a Document Quota

Evidence is comparative. Sober’s account of likelihood reasoning expresses the central idea: an observation favors one hypothesis over another when the observation is more probable on the first (Sober, 2008). A surviving report can therefore support a miracle claim if such a report is more expected when the event occurred than when it did not. Missing reports can count the other way when their absence is less expected if the event occurred.

This approach improves on two common slogans. First, “extraordinary claims require extraordinary evidence” should not function as an unexplained veto. Hume’s discussion of miracles makes the conflict between testimony and established experience vivid (Hume, 2008), but Earman rightly argues that no blanket maxim makes testimonial evidence incapable in principle of supporting a miracle (Earman, 2000). A claim that begins as very improbable can still gain rational support if the evidence has a sufficiently large likelihood advantage. The burden is demanding because the evidence must overcome the relevant background, not because a special anti-miracle rule has been added.

Second, consistent standards do not require the same quantity of documentation in every era. Consider two reports involving a dramatic event in a city. In a media-rich modern setting, a genuine event witnessed by many would be expected to generate recordings, emergency calls, location data, named testimony, follow-up reporting, and institutional records. In an ancient city, the expected set is much smaller and the chance that any document survives is lower. The parity principle uses the same rule while accepting that these background conditions alter the likelihoods.

The decisive mistake is to move from “ancient records are often sparse” to “the sparsity of this extraordinary public claim is not evidentially adverse”. That conclusion does not follow. Sparse preservation reduces the force of silence; it does not automatically erase it, and it certainly does not turn the lack of corroboration into support. The right question is event-specific: given the event’s scale, alleged witness pool, social salience, available record-making, routes of transmission, and survival chances, how surprising is the record we actually possess?

3. When Absence Becomes Evidence

An argument from silence is strong only when the missing trace was genuinely expected. Four considerations discipline the inference.

1. **Opportunity and access.** Who could have observed the event or its aftermath? A public label in a text is not enough if the relevant observers were absent or unable to identify what happened.
2. **Motivation and record-making.** Would observers have reason and means to report the event? Literate elites may ignore local events, and official writers select material according to genre and purpose.
3. **Independence and transmission.** Ten later documents copied from one stream are not ten independent

witnesses. Conversely, one surviving document may preserve more than one earlier source, but that independence must be argued rather than assumed.

4. **Survival and discovery.** Records can be lost, destroyed, never copied, or not yet found. The longer and harsher the transmission path, the weaker any inference from present absence.

These conditions prevent cheap arguments such as “a named historian did not mention it, therefore it never happened”. Unless we know that the historian had access, interest, and a work likely to preserve the notice, the silence may have little force. They also prevent the opposite error: treating every loss story as sufficient to neutralize any missing evidence. A possibility of loss matters only to the degree that loss is expected.

Testimony deserves similar care. Coady emphasizes how deeply ordinary knowledge depends on testimony; rational inquiry cannot begin by treating all reported information as worthless (Coady, 1994). But accepting testimony as a legitimate source is not the same as counting downstream repetitions as independent confirmation. A report must still be assessed for proximity, identity, incentives, internal quality, dependence, and fit with background knowledge. Sacred status may matter within a faith commitment, but it cannot serve as neutral public evidence for the very authority under dispute without circularity.

These controls yield the test used below. Matthew’s narrative counts as testimony and must not be dismissed merely because it reports a miracle. But its missing independent traces count against the episode to the extent that additional traces are more expected if a large public resurrection occurred than if the report developed without the event. Ancient loss weakens that negative contribution; it does not convert one testimonial endpoint into corroboration.

4. A Controlled Parity Test

Imagine a report claiming that, ten years ago, graves opened in a New York cemetery, many long-dead people entered the city, and they appeared to many residents. The article names no witness, links to no recording, and stands alone. The reporter is unavailable. No other outlet, family, hospital, public agency, or researcher supplies an independent trace.

The report should be rejected as historically inadequate. That judgment does not require certainty that the event was impossible. The observed combination—a massive public event plus one uncorroborated report and widespread expected silence—is far less expected if the event occurred than if the article is mistaken, embellished, satirical, or fabricated.

Now compare Matthew’s episode. The content is structurally similar at the level that matters: the dead are said to leave tombs, enter a city, and appear to many. The documentation environments are not similar. The table makes both the parity and the asymmetry visible.

Factor	Modern cemetery report	Matthew 27:52–53	Evidential role
Claimed scale	Many dead people enter a major city and appear to many	Many saints enter Jerusalem and appear to many	Both claims generate an expectation of effects beyond one narrator
Recording setting	Dense media, institutional records, and durable digital copying	Limited literacy, selective ancient genres, fragile hand-copying, and severe loss	Modern silence is much stronger; parity does not imply equal document counts
Observed attestation	One report with no named witnesses or independent trace	One extant canonical narrative; the saints and those who saw them are unnamed	A lone endpoint cannot be multiplied into independent confirmation
Background pressure	Resurrections conflict with stable biological experience	The same event type conflicts with stable biological experience; theology may alter the prior	Prior disagreement belongs in the model but does not cancel source assessment
Calibrated result	The report is historically unacceptable	Confidence is lowered, but by a smaller and less measurable amount	Same rule, context-sensitive strength, no fabricated posterior numbers

This table is not a claim that the two cases have equal posterior probability. The modern case has a far larger expected trace set, so its silence is more devastating. The comparison instead isolates the paper’s governing question: once historical context is included, does the ancient case possess positive, independent evidence strong enough to overcome both the extraordinary event type and the traces reasonably expected from the episode as narrated? For Matthew’s resurrected saints, the answer developed next is no: the surviving record lowers confidence and leaves the episode unestablished, although by a smaller and less measurable amount than in the modern case.

5. Matthew’s Resurrected Saints

Matthew’s text itself raises the expectation. This is not a private healing in an unidentified village. It describes many raised bodies, a movement into Jerusalem, and appearances to many. Yet within the surviving canonical Passion narratives, the episode is unique to Matthew. The text names neither a raised saint nor a person to whom a saint appeared. No independently identifiable extant source narrates the same event.

These observations do not demonstrate that no earlier testimony stood behind Matthew. An author can preserve true information from sources no longer recoverable. Nor does the absence of the episode from other extant writings prove that every potential observer would have written. Ancient evidence is fragmentary, and literary authors routinely omit events another author includes.

Here is the central comparison. If many formerly dead people entered Jerusalem and appeared to many, some independent social trace beyond one literary endpoint is more expected than under nonoccurrence alternatives in which the report developed without the event—for example, as a literary-theological element attached to the death and resurrection scene. What we observe instead is one endpoint with no named witnesses to the saints and no independently identifiable corroboration. The missing trace need not be a Roman government file or a surviving book by a famous historian. It could be an independent early tradition, a named line of testimony, a second narrative with evidence of independence, or some durable controversy about the people said to have returned. We presently possess none that can be identified as independent.

The evidential conclusion is therefore directional and bounded. The lone attestation does not establish the episode by ordinary public historical standards. The mismatch between the event’s narrated scale and its surviving trace lowers confidence. Because ancient recording and survival probabilities are uncertain, the magnitude of that reduction cannot be responsibly quantified from the available record.

This is also where the earlier temptation to assign exact priors and likelihood ratios must be resisted. Numbers such as a one-in-a-million prior or a one-in-a-thousand silence factor look rigorous but merely hide judgment behind decimals when no reference class or dataset warrants them. Qualitative Bayesian reasoning is enough: state the competing expectations, expose the dependence assumptions, and identify the direction and uncertainty of the update.

6. The Strongest Replies

6.1 Matthew intended literal history

Quarles argues from Matthew's vocabulary, narrative structure, scriptural allusions, and early reception that Matthew 27:51–53 was intended as a literal historical account rather than poetic fiction (Quarles, 2016). That case matters because a purely symbolic reading cannot simply be assumed. Yet intended genre and historical truth are different questions. Establishing that Matthew meant to report an occurrence identifies what his testimony asserts; it does not make that testimony independent of itself or supply the named witnesses and corroborating traces at issue. The paper's conclusion therefore does not depend on proving a symbolic genre. Even on a literal-intent reading, the surviving public evidence remains one extant narrative endpoint.

6.2 Most ancient records were lost

This reply is correct and important. It weakens the inference from silence. But it is not a universal solvent. The historical question is not whether loss was possible; it is how likely the observed trace pattern is given the event's scale, the claim that many saw the raised saints, the routes by which early Christian traditions spread, and the survival of other Passion material. Loss reduces the penalty. It does not make one attestation equivalent to corroboration, and it does not make scarcity positive evidence.

6.3 Oral transmission could preserve the event

It could. Testimony can transmit knowledge, and written evidence is not the only evidence. Yet an appeal to oral tradition must identify what it adds. If all surviving material reaches us through one textual endpoint, merely positing many earlier tellers does not establish their independence, proximity, or accuracy. Oral transmission is a possible history of the report; it is not itself additional public evidence for the event.

6.4 A Christian worldview changes the prior

It may. If a person already has strong independent reason to accept a Christian God-concept and a divine plan involving resurrection, the event begins less surprising for that person than for a naturalist. But the same update rule still applies. A higher prior cannot turn weak or dependent testimony into independent testimony. Moreover, when miracle reports are offered to establish the worldview for an inquirer, assuming that worldview in the prior risks circularity. The paper leaves the broader prior dispute open while insisting that evidence be counted honestly within it.

6.5 The wider resurrection case is cumulative

This is the strongest scope objection. Defenders such as the McGrews and Swinburne do not rest the resurrection of Jesus on Matthew's saints; they appeal to a broader set of facts, testimony, background arguments, and explanatory judgments (McGrew & McGrew, 2009; Swinburne, 2003). The present analysis does not settle that cumulative case. It isolates one public episode and one methodological temptation. If the Matthew episode contributes to a

cumulative argument, its contribution must reflect its actual evidential quality rather than borrowing the strength of the whole case and then being counted again as independent support.

6.6 Miracles cannot be judged by ordinary evidence

If this means only that a miracle is unusual, the Bayesian framework already accommodates unusual events through background probability and evidential force. If it means that no possible observation could count against a miracle report, the report has been removed from public historical assessment. It may remain an article of faith, but favorable testimony can no longer be advertised as rational public confirmation while unfavorable evidence is declared irrelevant. Evidence must be allowed to cut both ways.

Taken together, these replies affect the intended genre, the starting prior, the expected survival of records, or the scope of the paper; none supplies independent corroboration for Matthew's saints. They therefore preserve the episode's logical possibility without establishing it by public historical standards. The central conclusion survives the strongest objections: the record lowers confidence, while ancient uncertainty prevents a numerical or deductive verdict.

7. Symmetry Across Traditions

The parity principle is not a device aimed only at Christianity. A Muslim claim about the moon, a Hindu miracle narrative, a Latter-day Saint report, and a modern healing claim should face the same structure of questions: What exactly is alleged? Who had access? What independent traces were expected? What survives? How dependent are the sources? How does the full record compare under occurrence and nonoccurrence?

This cross-tradition symmetry matters because cultural familiarity can disguise a double standard. A reader may find one sacred narrative natural and a rival narrative obviously legendary even when their public evidence is structurally similar. The remedy is not equal disbelief by fiat. It is equal exposure to the same rule. A claim with stronger independent evidence should receive more confidence regardless of tradition; a claim protected only by special exemption should receive less.

The symmetry also clarifies why "ancient mystique" is not itself an argument. Distance can create an aura of seriousness, but it also removes opportunities for checking. When age reduces the available evidence, intellectual honesty should reduce confidence or increase uncertainty unless positive evidence compensates. Age alone never supplies that compensation.

8. What Would Change the Assessment?

The conclusion is defeasible. Evidence that would materially strengthen a public ancient miracle claim includes:

- multiple early reports with demonstrable independence rather than literary copying;
- named witnesses whose access and relation to the event can be assessed;
- a trace that was fixed before later theological interests could shape the report;
- corroboration from a source with different incentives or a different transmission path;
- durable public consequences specifically linked to the alleged event rather than to later belief in it; and
- a record whose difficult features are more expected if the event occurred than if a narrative developed without it.

No item is mechanically required in every case. A private event may reasonably leave one report. A public event involving many appearances carries a different trace expectation. The list shows what responsible inquiry looks for and prevents a claim from becoming immune to evidence by design.

The same standard protects believers from bad miracle reports as much as it challenges cherished ones. Requiring source independence, proportionate expectations, and symmetric evaluation filters fraud, rumor, pious embellishment, and honest error. If a genuine miracle occurred, consistent methods are the means by which it could be distinguished from the many claims that did not. Until evidence of the kinds listed above is available for Matthew's saints, the report remains unestablished by those methods.

9. Conclusion

Ancient miracle reports do not deserve a lower evidential standard because they are old, scriptural, or culturally familiar. They deserve the same comparative rule used for modern reports: assess how expected the full surviving record is if the event occurred and if it did not. Historical context changes the expected traces and therefore the strength of silence. It does not change the logic.

Applied to Matthew's resurrected saints, the result is clear but limited. The passage narrates a large public event involving many raised bodies and appearances to many, yet the surviving evidence contains one literary attestation, no named witnesses to the saints, and no independently identifiable corroboration. Ancient record loss makes the absence less damaging than it would be in modern New York. Even so, the mismatch between narrated scale and surviving trace lowers historical confidence and leaves the episode unestablished by ordinary public standards.

This conclusion does not prove that the episode is false. It does not disprove the resurrection of Jesus, every miracle claim, Christianity, or theism. It does require a change in intellectual practice: stop treating antiquity as an evidential privilege. If a method would reject a rival or modern miracle report for weak, dependent, and uncorroborated evidence, do not quietly suspend that method when the report is one's own.

A. Formal Reconstruction of the Central Thesis

Plain-language reconstruction

- P1.** Ancient and modern public miracle reports should be evaluated by the same comparative update rule, with each report's historical recording, transmission, and survival conditions included in the background information.
- P2.** Under that rule, missing traces lower confidence only when those traces were more expected if the event occurred than if it did not; the force of the reduction is proportional to that difference.
- P3.** Matthew 27:52–53 describes a large public event and appearances to many, while the surviving record offers one extant literary attestation, no named witnesses to the saints, and no independently identifiable corroboration. Even after allowance for ancient record loss, that complete record is less expected if the narrated public event occurred than under nonoccurrence alternatives in which the report developed without the event.
- P4.** Therefore, consistent evidential updating lowers historical confidence in the Matthew episode to a context-dependent degree and does not establish it as a public historical event.

Scope of the conclusion. P4 does not entail that the episode is impossible or deductively false. It does not settle the prior probability of miracles, the resurrection of Jesus, Christianity, or theism. It states the direction warranted by the surviving evidence under the parity rule.

Symbols

- *i*: an index for either the ancient or the modern case.

- H_i : the hypothesis that the reported public miracle in case i occurred.
- E_i : the complete surviving evidence in case i , including testimony and missing expected traces.
- B_i : the historical background for case i , including record-making, access, transmission, and survival conditions.
- P : the probability function used for comparative evidential assessment.
- Λ_i : the likelihood ratio supplied by E_i for H_i on background B_i .
- O : the odds function used to express prior and posterior confidence.

Symbolic counterpart

For each case i , the evidence supplies the likelihood ratio

$$\Lambda_i = \frac{P(E_i \mid H_i, B_i)}{P(E_i \mid \neg H_i, B_i)}. \quad (1)$$

The same update rule then applies in both eras:

$$O(H_i \mid E_i, B_i) = O(H_i \mid B_i) \Lambda_i. \quad (2)$$

When the full record is less expected if the event occurred,

$$P(E_i \mid H_i, B_i) < P(E_i \mid \neg H_i, B_i) \iff \Lambda_i < 1, \quad (3)$$

and therefore

$$O(H_i \mid E_i, B_i) < O(H_i \mid B_i). \quad (4)$$

Validity and support

The symbolic inference is deductively valid once the likelihood comparison is granted: a likelihood ratio below one reduces the odds of the event hypothesis. Historical parity lies in using the same equation for each i ; sensitivity to context lies in allowing E_i and B_i to differ. Here, $\neg H_i$ groups the possible ways the report could exist without the event; the argument does not require one specific nonoccurrence history to be exhaustive.

The central premise about Matthew is inductive, not deductive. Its support comes from the episode's narrated scale, the claim of many appearances, the single extant narrative endpoint, the lack of named witnesses to the saints, and the lack of independently identifiable corroboration. Uncertain ancient survival rates limit the magnitude of the update, which is why no numerical value is assigned to Λ_i . Independent early testimony or a better-supported account of the missing traces could reverse the comparison. The formalization therefore preserves the paper's bounded conclusion rather than turning an evidential criticism into a deductive disproof.

References

- Coady, C. A. J. (1994). *Testimony: A philosophical study*. Oxford University Press. <https://doi.org/10.1093/0198235518.001.0001>
- Earman, J. (2000). *Hume's abject failure: The argument against miracles*. Oxford University Press. <https://doi.org/10.1093/0195127382.001.0001>
- Hume, D. (2008). *An enquiry concerning human understanding* (P. Millican, Ed.). Oxford University Press. <https://doi.org/10.1093/owc/9780199549900.001.0001>
- McGrew, T., & McGrew, L. (2009). The argument from miracles: A cumulative case for the resurrection of Jesus of Nazareth. In W. L. Craig & J. P. Moreland (Eds.), *The blackwell companion to natural theology* (pp. 593–662). Wiley-Blackwell. <https://doi.org/10.1002/9781444308334.ch11>
- Quarles, C. L. (2016). Matthew 27:51–53: Meaning, genre, intertextuality, theology, and reception history. *Journal of the Evangelical Theological Society*, 59(2), 271–286. https://etsjets.org/wp-content/uploads/2016/06/files_JETS-PDFs_59_59-2_JETS_59-2_271-286_Quarles.pdf
- Sober, E. (2008). *Evidence and evolution: The logic behind the science*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511806285>
- Swinburne, R. (2003). *The resurrection of God incarnate*. Oxford University Press. <https://doi.org/10.1093/0199257469.001.0001>