

Historical Claims and the Discipline of Inference

Evidence, Dependence, Prior Probability, and the Methodology of Responsible Belief

Phil Stilwell

Independent Scholar

ORCID: 0009-0009-3725-8682

Revised August 2026

Central claim

Historical traces can strongly support ordinary events or community beliefs without equally supporting a miracle or its theological interpretation. A responsible public inference discounts source dependence, combines prior odds with comparative likelihoods, and adds each bridge from event to identity or meaning explicitly. This discipline does not rule out miracles, theism, or private warrant; it prevents a modest historical result from being promoted into a maximal conclusion without additional evidence.

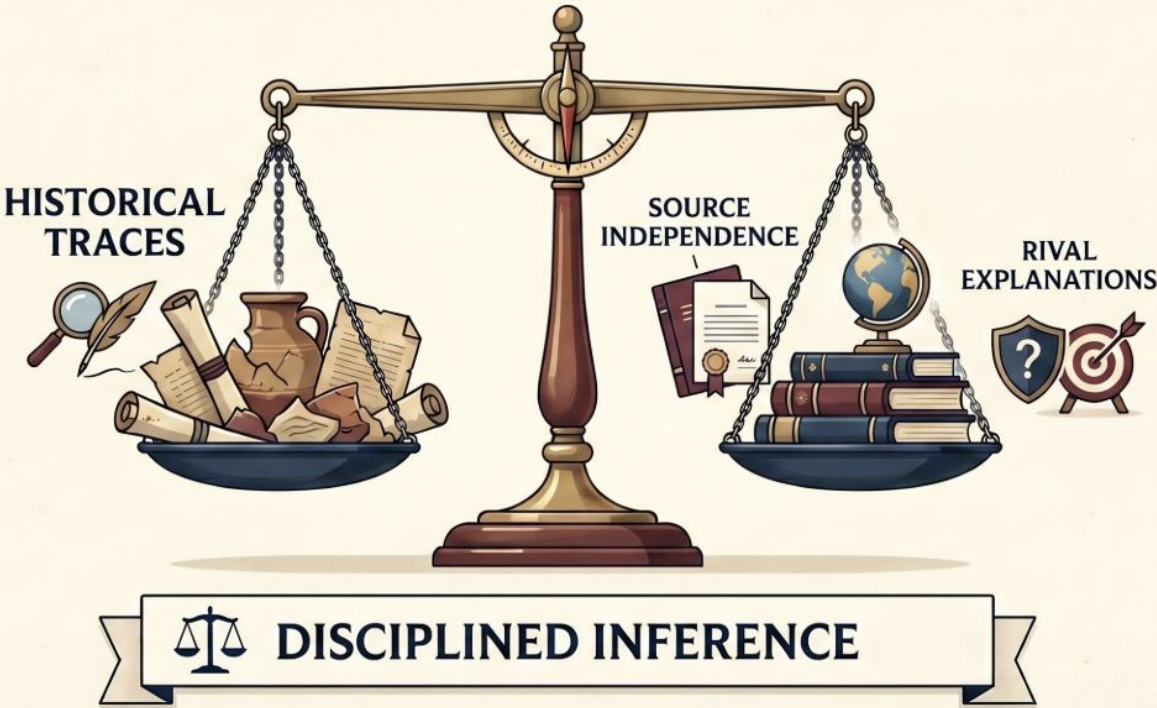
Preview: From trace to warranted conclusion

Separate the claims	Weight the evidence	Bound the conclusion
Attestation, experience, event, and theological meaning are distinct. State every bridge; do not let story sequence carry it.	Shared traditions are not independent; use prior odds and comparative likelihoods. Compare the evidence under the strongest live alternatives.	Lower layers can be strong while a miracle remains underdetermined. Stronger public confidence needs stronger discriminating evidence.

This calibrates public historical warrant claim by claim; it neither bans nor privileges religious explanations.

Visual preview:

THE DISCIPLINE OF HISTORICAL INFERENCE



The balance aligns traces with source independence, background knowledge, and rivals; it depicts disciplined comparison, not a fixed weight or verdict.

Abstract

Historical argument often moves from secure but modest facts to far more ambitious conclusions by narrative momentum. Several texts attest a belief; therefore several witnesses confirm an event. Followers report experiences; therefore the event they inferred occurred. An extraordinary event is granted; therefore a particular theological identity and meaning follow. This paper argues that each transition needs its own evidential bridge.

The proposed discipline combines source criticism with comparative probability. Reports derived from a common tradition must be represented jointly rather than multiplied as independent confirmations. Posterior odds depend on both prior odds and how much more expected the surviving evidence is under one hypothesis than under its rivals. Explanatory virtues matter, but only when they alter those parts of the comparison instead of becoming a second, unpriced scorecard. A six-level claim ladder and an evidence ledger apply the method to Christian resurrection arguments while engaging serious affirmative and cautious scholarship.

Worked examples involving syndicated news, diagnostic evidence, a missing car, and a viral bribery allegation make the method concrete outside religious debate. They show why many reports may trace to one source, why evidence can favor a claim without yet making it probable, and why establishing an event does not automatically establish its cause or meaning.

The result is calibrated rather than dismissive. Historical evidence may strongly support execution, early proclamation, and sincere experiences while providing weaker public support for an empty tomb, a supernatural bodily resurrection, or a complete theological interpretation. The argument does not prove that a resurrection did not occur, deny theism, or settle private religious warrant. It shows why minimal historical claims cannot carry maximal theological conclusions without additional independent evidence and explicit bridge premises.

Keywords: philosophy of history; historical inference; source criticism; testimony; Bayesian reasoning; prior probability; miracles; resurrection; Christian apologetics

1 The Answer Before the Method

The past does not arrive as a replay. It leaves texts, artifacts, memories, institutional records, and later interpretations. Historians can learn a great deal from those traces, but the trace is not the event and the event is not yet its meaning. Marc Bloch's practical source criticism, R. G. Collingwood's account of historical reconstruction, and modern philosophy of historiography all begin from this mediated character of historical knowledge (Bloch 1953; Collingwood 1946; Tucker 2004). Mediation does not produce skepticism. It creates work.

By *public warrant*, this paper means support that can be assessed through evidence and inferential standards available to people who do not already share one complete worldview. Public warrant does not require unanimous agreement, mathematical certainty, or worldview-free reasoning. It requires enough disclosure that another reader can identify the evidence, the alternatives, the background assumptions, and the steps connecting them.

That work becomes especially visible when an argument begins with ordinary historical claims and ends with a miracle or theological conclusion. A teacher was executed. Followers later proclaimed that he lived. Some reported experiences interpreted as appearances. Narratives developed around an empty tomb. From this material, an apologist may infer not only that an extraordinary event occurred but that the Christian deity raised Jesus, authenticated a particular identity, and conferred a specified salvific meaning.

The problem is not that such a conclusion is religious. The problem is that it contains several claims, each stronger than the one before it. Evidence can be excellent for an early proclamation and weaker for the event proclaimed. It can favor an unusual event without identifying its cause. It can identify an event and still leave its intended meaning open. Treating the whole package as one conclusion hides these differences.

The governing thesis is therefore simple: *historical confidence should be assigned claim by claim, with source dependence, prior odds, comparative likelihoods, and bridge premises made visible*. A later layer may be true. It does not inherit the warrant of an earlier layer merely because the claims appear in one story.

This thesis is directed at a common public argument, not at every sophisticated defense of resurrection. Richard Swinburne explicitly embeds the resurrection in a larger probabilistic case about a specified God-concept, incarnation, divine reasons, and evidence before and after Jesus' death (Swinburne 2003). Michael Licona develops a large historiographical comparison of resurrection hypotheses (Licona 2010). Dale Allison's wide-ranging treatment deliberately resists both apologetic and skeptical overreach (Allison 2021). These works disagree, but they agree on something the compressed popular argument often misses: the conclusion cannot responsibly be read straight off a list of minimal facts.

2 A Claim Ladder Instead of a Claim Package

Historical disputes become clearer when conclusions are placed on a ladder. The ladder is diagnostic rather than a strict chain of logical entailments: evidence may support one level unevenly, and some levels can be investigated in parallel. Still, moving from a report to a belief, from a belief to its cause, or from an event to its theological meaning ordinarily adds content and inferential distance. The added content needs stronger evidence or a defended bridge. Table 1 displays a general version tailored to resurrection debate.

Table 1: A claim ladder for extraordinary historical inference. Higher levels may be true, but they do not inherit the confidence of lower levels without added support.

Level	Claim type	Typical conclusion	Additional burden
1	Attestation	A text or community reports p	Date, provenance, transmission, genre, and wording
2	Social belief	A group sincerely accepted p	Evidence that the report reflects belief rather than later convention
3	Experience	People had experiences interpreted as p	Access to reports, diversity of cases, and alternatives for the experiences
4	Ordinary event	A publicly familiar event occurred	Source fit, background rates, material traces, and rival causal accounts
5	Extraordinary event	A rare or law-exceptional event occurred	A likelihood advantage large enough to overcome lower prior odds
6	Identity and meaning	A specified deity acted for a specified theological purpose	Independent premises about the agent, intention, authority, and interpretation

The first three levels are often confused. A document directly shows that a statement appears in the document. With further work, it can support a conclusion about what an author or community believed. It may then support a conclusion that some people had experiences they described in the statement's terms. None of these conclusions is identical to the event as interpreted.

The fourth and fifth levels differ through background expectations, not through a rule that history may consider only ordinary events. A report that a governor executed a dissident fits abundant background knowledge about governors, punishments, and mortal bodies. A report of permanent bodily return from death conflicts with a very stable generalization. The report is still evidence. It simply faces a larger comparative burden.

The sixth level prevents theological promotion. Suppose, for argument's sake, that an extraordinary return from death were established. It would not follow by historical logic alone that the agent was the Christian deity, that Jesus possessed every identity later doctrine assigns him, or that the event carried one exact theory of atonement. Those conclusions may be supported by independent theology. They are not synonyms for the event.

A bank transfer provides an ordinary analogy. A verified record may strongly establish that money moved from one account to another. It does not by itself establish whether the transfer was a gift, a debt payment, a bribe, or an error. Greater confidence that the transfer occurred does not automatically become greater confidence in one motive. Messages, timing, agreements, and subsequent actions must supply that further bridge. In the same way, confidence in an event does not simply flow into confidence about its agent and purpose.

The claim ladder does not require numerical credences. Its first purpose is grammatical and evidential: identify which proposition the evidence bears on most directly, then state every upward bridge.

3 Evidence Is a Network, Not a Stack

Lists create a visual illusion. Five reports look like five pieces of evidence. But evidential force depends on how the reports were generated. If four reproduce a fifth, the structure is one upstream stream with several downstream expressions. Repetition can reveal circulation, influence, and community acceptance without providing four independent observations of the event.

The modern news cycle makes the point familiar. A claim may appear on five websites, in three videos, and in hundreds of posts, yet every item may trace to one anonymous message. The volume is evidence that the claim spread. It is not evidence that eight reporters independently observed the alleged event. Counting webpages answers a question about circulation; tracing reporting routes answers a question about corroboration.

Source criticism asks who produced a source, when, for whom, in what genre, with what access, and under what interests or constraints (Howell and Prevenier 2001). These questions do not function as automatic defeaters. A partisan source can preserve accurate information; a late source can transmit early material; a hostile source can repeat rumor. The questions identify the route by which the source can carry evidence.

Literary dependence provides a concrete example. The Synoptic Gospels share extensive wording and structure, and proposed solutions differ about the exact relations. Mark Goodacre's accessible treatment shows why agreements among Matthew, Mark, and Luke cannot simply be counted as three independent reports (Goodacre 2001). Literary dependence does not make the shared material worthless. It changes the target of the agreement: the texts may confirm a common tradition more directly than three separate eyewitness routes.

The same discipline applies beyond copying. Reports can share a teacher, liturgy, community vocabulary, apologetic need, or earlier testimony. A later source may have independent material in one passage and dependent material in another. Dependence is therefore not a binary label attached to whole books. It is a map of pathways for particular claims.

If evidence items $E_1, \dots, E_n$ were conditionally independent given hypothesis H , their joint likelihood could be factored:

$$P(E_1, \dots, E_n \mid H) = \prod_{i=1}^n P(E_i \mid H). \quad (1)$$

In ordinary words, genuinely independent reports can multiply evidential force. But the equality is an assumption, not a reward for having several documents. When reports share a source or social process, the correct quantity is the joint probability of the whole pattern. Multiplying each report as though it arose independently counts some of the same evidential route more than once.

A toy calculation shows how large the error can become. Suppose one originating report is nine times more expected under hypothesis H than under its rival. If four later reports merely copy it, and copying is equally likely once the first report exists under either hypothesis, those copies add no comparative leverage in this simplified case. The joint likelihood ratio remains 9, not $9^5 = 59,049$. These numbers estimate no historical case; they isolate the mathematical consequence of treating copies as fresh witnesses.

Dependence does not mean deletion. Two reports sharing a source may still contain independently checked details, preserve different versions, or establish that a tradition existed by a certain date. The point is narrower and stronger: once dependence is present, the simple product rule is unavailable. The whole pattern must be evaluated jointly, with each genuinely new route identified rather than presumed.

Selection adds a second network problem. Historians usually encounter what communities copied, canonized, defended, translated, and preserved. A surviving record is shaped by those processes. This does not erase its evidence; it changes the comparison. A hypothesis about what a community would preserve can predict the surviving corpus differently from a hypothesis about what the first witnesses observed. Public historical inference should model both stages when the difference matters.

A family photo album illustrates the selection issue. An album filled with birthdays and holidays is strong evidence that those celebrations occurred. It is poor evidence that most days were celebrations, because ordinary days were less likely to be photographed and preserved. In the same way, a corpus can reliably record selected events or beliefs while giving a distorted picture of their frequency, diversity, or representativeness. Preservation must therefore be part of the explanation, not treated as a transparent window onto everything that happened.

Three practical questions expose the network:

1. Which claims have genuinely distinct upstream access routes?
2. Which agreements are more expected because of copying, shared tradition, or institutional selection?
3. What would likely have survived, disappeared, or changed under each rival hypothesis?

These questions are not hostile to testimony. They are what allow testimony to earn more than mere compatibility.

4 Prior Odds and Evidential Leverage

Historical reasoning always uses background knowledge. A source date matters because writing and transmission behave in familiar ways. Physical traces matter because materials decay in familiar ways. Testimony matters because people perceive, remember, coordinate, deceive, and sincerely misinterpret in familiar ways. A prior is simply the comparison before the evidence currently under discussion; it need not be a precise number or an ideological ban.

Bayesian notation makes the architecture visible. For a miracle hypothesis M , a live alternative A , evidence E , and background information B :

$$\frac{P(M \mid E, B)}{P(A \mid E, B)} = \frac{P(M \mid B)}{P(A \mid B)} \times \frac{P(E \mid M, B)}{P(E \mid A, B)}. \quad (2)$$

The left side is the posterior odds after considering the evidence. The first factor on the right is the prior odds. The second is the likelihood ratio: how much more expected the evidence is if the miracle occurred than if the alternative occurred. Thomas Bayes's posthumous essay helped establish the probability relation that later work developed into this familiar framework (Bayes 1763).

The formula blocks two opposite shortcuts. A critic cannot declare a miracle impossible merely because its prior is low. A sufficiently large likelihood ratio can overcome low prior odds. A defender cannot ignore the prior and say that any evidence more compatible with the miracle than with one weak rival establishes the miracle. The amount of comparative leverage matters.

Consider a purely illustrative diagnostic example. If a condition begins at prior odds of 1:1,000 and a test result is one hundred times more likely when the condition is present than when it is absent, the posterior odds become 1:10, or about nine percent. The result strongly favors the condition relative to its starting point without making the condition more likely than not. If the likelihood ratio were instead 100,000:1, the same prior odds would become 100:1. The lesson cuts both ways: a low prior is not a veto, and evidence that favors a claim can still fall short of the confidence asserted for it.

David Hume's relevant discussion is section 10, "Of Miracles," in *An Enquiry Concerning Human Understanding*, not the later *Dialogues Concerning Natural Religion* (Hume 2007). Hume's own formulation remains disputed. John Earman argues that a blanket anti-miracle conclusion cannot be secured by Hume's rhetoric and that testimony should instead be assessed with probability calculus (Earman 2000). That correction supports the present method. There is no rule that testimony can never establish a miracle. There is a demand to compare the strength and dependence of the testimony with the relevant background and alternatives.

The alternative side also needs care. "Natural explanation" is not one hypothesis with one likelihood. It is a family that may include grief experiences, visions, memory transformation, scriptural interpretation, legendary growth, deliberate invention, mistaken identification, and mixtures. A critic who vaguely invokes the whole family can hide the same flexibility criticized in a miracle account. The defensible comparison either specifies its principal alternatives or models the family as a weighted mixture and reports sensitivity to those weights.

Comparing a favored hypothesis with one weak rival is like a prosecutor disproving one suspect's alibi while ignoring the rest of the lineup. The failure of that alibi can raise the favored suspect's relative standing, but it does not establish guilt over every serious alternative. A weighted mixture avoids that mistake only if its membership and weights are defended; an elastic catch-all family is no more informative than an elastic miracle hypothesis.

Background information must also be kept distinct from the evidence being evaluated. Independent philosophical or theological evidence may legitimately raise a miracle prior. But if the background case is itself derived from the same resurrection reports, those reports cannot then be counted again as a separate likelihood boost without adjustment. A bookkeeper does not create a second expense by filing a photocopy of one receipt in another folder. Evidential organization does not multiply evidence.

No precise numerical posterior is needed for the central result. The formula identifies the missing work. A qualitative argument can say that a prior disadvantage is modest or severe, that a pattern is far more expected under one account, or that a conclusion is highly sensitive to one assumption. But when an argument supplies neither a defensible direction and rough magnitude for the prior comparison nor a comparative likelihood against serious rivals, it has not established the posterior confidence it claims. Calling the evidence "powerful" does not fill the gap.

5 Inference to the Best Explanation Without Double Counting

Historians often reason by asking which account best explains the surviving evidence. C. Behan McCullagh emphasizes scope, explanatory power, plausibility, and fit in historical description, while Peter Lipton

develops inference to the best explanation as a contrastive form of induction (McCullagh 1984; Lipton 2004). This approach is indispensable, but it requires a price list.

Consider four familiar virtues:

- **Fit:** the account makes the observed evidence less surprising;
- **Scope:** it explains several relevant features rather than one selected fact;
- **Simplicity:** it does not purchase fit through many independently adjustable assumptions;
- **Background coherence:** it connects with well-supported knowledge rather than requiring unrelated exceptions.

These virtues should not become free points added after a Bayesian comparison. Fit and scope ordinarily affect likelihoods. Simplicity and background coherence ordinarily affect priors or the size of a hypothesis family. If a critic gives a miracle a low prior because it adds divine action and then subtracts another full score for “complexity,” the cost may be counted twice. If a defender gives the miracle a high likelihood because it explains every feature and then adds a second full score for “scope,” the benefit may also be counted twice.

The mistake resembles ranking apartments by awarding points for “large rooms,” “spaciousness,” and “plenty of interior area” as though these were three independent advantages. The labels differ, but much of the underlying virtue is the same. An explanatory scorecard is useful only when each criterion marks a distinct contribution or when overlap is explicitly controlled.

Elliott Sober’s analysis of parsimony is useful because it rejects the idea that simplicity has one automatic form or justification (Sober 2015). Historical explanations should specify what is being counted: entities, causal steps, adjustable intentions, exceptional events, auxiliary assumptions, or description length. A one-sentence miracle claim can be verbally short while carrying a large background theology. A multi-step social explanation can be verbally long while drawing on familiar processes. Word count is not epistemic cost.

The best explanation must also beat the best developed rival, not an empty space. “The resurrection explains the disciples’ belief” is incomplete because grief experiences, visionary experiences, group reinforcement, scriptural reinterpretation, and later narrative development can also contribute to that belief. Conversely, “legends happen” is incomplete because it may not explain the timing, conviction, or particular form of the tradition. Each side must state what its account predicts and where it strains.

The discipline can be summarized as a fork:

1. If a hypothesis is constrained enough to make the evidence distinctly expected, it can gain comparative support and can also risk disconfirmation.
2. If it can accommodate nearly any evidential record through adjustable intentions or ad hoc auxiliaries, compatibility remains high but discriminating confirmation falls.

This standard applies symmetrically. An all-purpose social story and an all-purpose divine-intention story are equally weak as public explanations.

Agency itself is not the problem. Historians routinely infer intentions when an agent’s goals and constraints are independently known: orders, prior plans, available means, and later conduct can make one action more expected than another. The same form is available to a specified divine-agent hypothesis. The difficulty arises only when the agent is assigned whatever intention makes the observed record fit after the fact. “The agent wanted exactly this pattern” can accommodate every outcome and therefore predicts none of them strongly before inspection.

6 Case Study: Christian Resurrection Arguments

The Christian resurrection is a useful case because the literature contains maximal affirmative arguments, cautious historical judgments, and skeptical alternatives. Swinburne embeds the event within prior natural theology and a specified account of incarnation (Swinburne 2003). Licona argues that resurrection best explains a historical bedrock through explicit historiographical criteria (Licona 2010). Allison surveys early traditions, tomb narratives, appearances, comparative experiences, apologetic arguments, skeptical arguments, and the limits of historical judgment (Allison 2021). A fair method should be able to recognize the strengths of each without collapsing their conclusions.

The case is best analyzed as an evidence ledger rather than a verdict list. Table 2 states what several often-cited strands most directly support, what dependence or selection issue applies, and what additional bridge is needed for a bodily resurrection conclusion. The rows are schematic; they do not pretend to replace detailed exegesis.

Table 2: An evidence ledger for resurrection arguments. Each strand can contribute, but its direct target and inferential limits differ.

Evidence strand	Most direct support	Dependence or selection issue	Bridge still required
Early proclamation and formulaic tradition	Early communal commitment to resurrection language	Access to original claimants, transmission history, and liturgical shaping vary	From early belief to the event that best caused it
Appearance reports	Some followers or later informants reported experiences interpreted as encounters	Reports differ in date, genre, source route, and narrative development	From sincere experience to a physical transformation of the corpse
Empty-tomb narratives	A tomb tradition became part of the resurrection story	Literary relations, narrative aims, harmonization, and historical core remain disputed	From an empty or relocated body to divine resurrection rather than other causes
Conversions of Paul and James	Striking changes in allegiance associated with experiences or convictions	Later accounts mediate details; one case does not automatically generalize to another	From transformation of belief to the external event believed
Growth and persistence of the movement	Resurrection belief became socially generative and durable	Successful movements are selectively visible; conviction can have several causes	From explanatory power within a community to truth of the event claim
Theological coherence	The event fits a specified Christian narrative of incarnation and vindication	The fit depends on prior commitment to that God-concept and interpretive framework	From event to divine identity, intention, and doctrinal meaning

What the strands actually discriminate

The relevant question is not whether each strand is compatible with resurrection. It is which rivals each strand makes less plausible. Early proclamation and appearance traditions, for example, count against the claim that resurrection belief was invented from nothing generations later. They fit both a resurrection and several accounts in which early followers sincerely interpreted experiences through scriptural and communal expectations. They therefore discriminate more strongly against pure late invention than against every non-resurrection alternative.

An empty tomb, if well established, would similarly count against accounts on which the body remained known and available in its expected place. But an empty parking space offers the right analogy: it establishes that the car is not there, not whether it was stolen, towed, loaned, moved by its owner, or never parked there. Evidence about guards, location, access, prior plans, and later possession would be needed to sort those causes. The absence matters; its causal identity is a further question.

Costly conversion and durable commitment are important evidence about sincerity and impact. If a person abandons a career after what they take to be a warning dream, the costly change makes

the report of conviction harder to dismiss as casual pretense. It does not by itself establish that an external messenger produced the dream. Likewise, the transformations associated with Paul or James can strengthen the claim that something they regarded as significant occurred while leaving the external cause under comparison.

Growth and persistence have still less direct leverage on the originating event because successful movements are the ones later observers are positioned to study. The family-album problem returns: survival explains visibility. Growth can support claims about social power, institutional transmission, and experienced meaning without serving as a truth detector for every founding narrative.

The ledger yields a layered result. Early resurrection belief is not well explained by saying only that nothing happened. Something generated and stabilized the proclamation. Experiences interpreted as appearances are a serious part of the record, and a responsible account should not dismiss them through ridicule. Allison's comparative material is valuable precisely because religious experiences can be psychologically and culturally real while remaining difficult to convert into one external-event conclusion (Allison 2021).

At the same time, sincerity is not event identity. A person may accurately report having an experience while misidentifying its cause. Group reinforcement can preserve a belief without every report being fabricated. Scriptural interpretation can shape memory without reducing every memory to invention. The strongest alternatives are therefore often mixed accounts, not one-word diagnoses.

The empty tomb, if granted, would narrow some alternatives and strengthen others. It would not by itself specify resurrection. A moved body, mistaken location, incomplete knowledge, narrative development, or other unknown causes would remain competitors with different costs. Conversely, a critic who simply assumes no tomb tradition has historical value avoids rather than wins the comparison. The question is how much the evidence changes the odds among specified alternatives.

Swinburne's argument exposes the role of theological background especially clearly. A resurrection is less surprising if a specified deity exists, has reason to become incarnate, has reason to authenticate this person, and would use this event as a signature (Swinburne 2003). Those premises can raise the prior or likelihood of resurrection. But their work is independent work. A reader who does not already accept them cannot be told that the historical evidence alone supplied the full posterior.

Licona's comparative project likewise improves on a bare appeal to possibility because it states historical facts and weighs hypotheses (Licona 2010). The remaining discipline is to keep the candidate sets symmetric. A resurrection hypothesis should include enough about agency and expected evidence to be tested. Non-resurrection rivals should be developed rather than reduced to caricatures. Dependence among data points should be represented. Uncertain facts should not silently enter as fixed premises.

The case therefore supports a firm but bounded conclusion. Public historical evidence can warrant high confidence in some ordinary claims, substantial confidence in early belief and experiences, and continued dispute about the tomb. The step to divine bodily resurrection remains sensitive to priors, dependence, and the composition of alternatives. The further step to a complete Christian identity and salvific interpretation requires additional theological bridges. This is underdetermination at the upper layers, not a proof of nonoccurrence. Here *underdetermined* does not mean merely that a remote story can be imagined; it means that, after serious rivals and source relations are represented, the record has not been shown to supply the comparative advantage required for the confidence claimed.

7 Strong Objections

7.1 "Historians cannot assign meaningful numbers"

They often cannot, and they need not pretend otherwise. The posterior-odds formula can function structurally: state which prior comparison matters, which evidence should favor which hypothesis,

and where dependence prevents multiplication. Qualitative judgments such as “much more expected,” “roughly equal,” or “highly sensitive to this assumption” are preferable to invented precision.

Aviezer Tucker argues that historiography can produce probable knowledge through evidence and explanation even while many interpretations remain underdetermined (Tucker 2004). The discipline proposed here belongs to that middle ground. It makes the architecture explicit without claiming that every term can be measured.

7.2 “A low miracle prior merely assumes naturalism”

Not necessarily. The prior can reflect observed event frequencies, a larger metaphysical theory, or a mix of background commitments. A theist may assign a miracle a higher prior than a naturalist. That disagreement should be displayed rather than hidden.

Earman is right that testimony is not disqualified in advance (Earman 2000). The fair conclusion is conditional: readers with different background odds may rationally reach different posteriors from the same evidence. A public argument becomes stronger when it shows which range of prior odds its evidence could overcome.

7.3 “Natural alternatives are speculative too”

Often they are. “Hallucination,” “legend,” or “cognitive dissonance” can become labels rather than explanations. The remedy is symmetry: specify who experienced what, which source route reports it, how the tradition changes, and why the surviving pattern is expected. A weak natural account does not become strong merely because it invokes familiar causes.

Nor must a defender defeat every imaginable story. A serious rival should describe a coherent causal route, fit relevant background knowledge, and make a substantial portion of the evidence reasonably expected without unlimited adjustment. Possibility alone does not earn a place in the comparison. This standard excludes both “perhaps something unknown happened” and a miracle account defined only as “a divine agent did whatever produced these traces.”

The same point benefits the resurrection defender. If every developed alternative fits poorly while a constrained resurrection model predicts the pattern well, the likelihood ratio can favor resurrection. The critic may not answer by repeating that miracles are unusual. The prior and likelihood must both remain in view.

7.4 “Multiple dependent sources can still corroborate”

Correct. Dependence is a discount, not a deletion. A later source can preserve an earlier independent strand, confirm that a tradition existed by a certain date, or reveal how a community understood it. Even copied material can provide evidence about transmission. The error is counting each downstream copy as a fresh upstream observation.

7.5 “Several modest strands can converge into a strong case”

They can, and the method depends on that possibility. Smoke, rising heat, and a melting electrical panel can jointly provide a strong case for a fire even though all three are effects of one event. Their force comes from the joint pattern being much more expected under fire than under steam from a shower or a faulty alarm. By contrast, three websites repeating one wire-service paragraph chiefly reveal one reporting route. The difference is not the number of observations but how the entire pattern is generated under the competing accounts.

A cumulative resurrection case can therefore exceed the force of any isolated strand. Early proclamation, appearance traditions, an empty tomb, conversions, and theological background could converge powerfully if their joint occurrence were much more expected under a constrained resurrection hypothesis than

under the best weighted alternatives. The paper's objection is not to accumulation. It is to accumulation by addition when dependence, selection, and overlap require a joint comparison.

7.6 “The theological package is the hypothesis”

It can be. Swinburne's specified package is a genuine hypothesis rather than a sequence of accidental inferences (Swinburne 2003). But packaging does not remove the burden; it relocates it. The prior must price the combined content, and the likelihood must reflect the range of evidence expected under the package. A broader package can explain more, but it can also contain more adjustable assumptions.

The advantage of a package is that independently defended intentions can generate real predictions. If a specified God-concept, doctrine of incarnation, and account of public authentication jointly imply that a vindicating event should be conspicuous, early, and connected to one person, then those features can raise the package's likelihood. The evidential price is equally clear: obscurity, late development, or patterns shared widely by incompatible traditions would count against that prediction. A package becomes testable by restricting what its agent is expected to do, not merely by giving the agent enough freedom to explain whatever survives.

7.7 “Requiring a bridge for every added claim creates an endless regress”

No. The demand is not that every premise receive an infinite chain of proofs. It is that confidence not move into newly added content for free. A verified transfer plus contemporaneous messages and an agreed exchange may jointly provide strong evidence of a bribe; the inquiry need not invent another motive behind every sentence before reaching a provisional judgment. The evidential chain stops where the available support is sufficiently discriminating for the stated level of confidence.

The bridge requirement therefore prevents equivocation, not closure. If a conclusion adds an agent, intention, or doctrine, the argument must identify what supports that addition. Once it does, the new evidence can be assessed in the ordinary way. Historical conclusions remain defeasible, as do most conclusions outside mathematics, but defeasibility is not an infinite regress.

7.8 “Historical restraint becomes practical indecision”

People must often act without consensus or numerical posteriors. The method does not forbid commitment. It distinguishes personal decision from what public historical evidence alone compels. A person may accept a tradition through trust, experience, or a broader worldview while acknowledging that a narrower historical argument leaves room for reasonable disagreement. That honesty strengthens rather than weakens intellectual responsibility.

8 A Practical Audit for Extraordinary Claims

The full method can be used as nine questions.

1. **What exact claim is being assessed?** Separate attestation, belief, experience, event, identity, and meaning.
2. **What is the most direct evidence?** Identify the trace before inferring beyond it.
3. **What are the source routes?** Map copying, shared tradition, social influence, and independent access claim by claim.
4. **What selection produced the surviving record?** Consider copying, canonization, loss, institutional interest, and publication.
5. **What are the serious rivals?** Develop the strongest live alternatives, including mixed accounts.

6. **What are the prior odds?** State the relevant background assumptions and report sensitivity rather than hiding them.
7. **What is the likelihood advantage?** Ask how much more expected the evidence is under the favored claim than under its rivals.
8. **Which bridge premises remain?** Do not let an event silently establish an agent, intention, authority, or doctrine.
9. **What would lower confidence?** Name disconfirming evidence and model failures before declaring a winner.

Worked example: a viral bribery allegation

Imagine that five news sites report that a mayor accepted a bribe. Tracing the stories shows that all five repeat one anonymous social-media post containing a screenshot of a bank transfer. The five publications strongly establish that the allegation circulated; they do not supply five independent sources. If the banking record is authenticated, it may establish that money moved. The claim that the payment was a bribe is a higher rung because it adds a motive and an exchange.

The serious alternatives are not limited to “bribe” and “nothing happened.” The transfer might be repayment, a gift, a business transaction, an accounting error, or part of a corrupt exchange. Investigators should ask which account better predicts the amount, timing, secrecy, messages between the parties, later official action, and attempts at concealment. A suspiciously timed transfer followed by a matching official favor and concealed communications would create far more comparative leverage than the transfer alone.

Prior beliefs about the mayor can guide but should not decide the case. A spotless reputation is not an impossibility proof; a reputation for corruption is not a conviction. The method asks how much the new evidence should move either starting point. It therefore avoids both reflexive dismissal and narrative inflation: authenticate the trace, map the reporting route, compare motives, and grant only the conclusion those steps support.

The questions protect both sides. They block a skeptic from using prior rarity as a veto. They block an apologist from using coherence as confirmation, dependence as multiplication, or minimal facts as a complete theology. They also improve ordinary history, where nationalist legends, courtroom narratives, conspiracy claims, and institutional memories can exploit the same shortcuts.

Responsible historical belief is not timidity. It is the courage to grant exactly what the record supports and to refuse both dismissal and inflation.

9 What Would Change the Upper-Layer Result?

A bounded conclusion should state what could strengthen it, weaken it, or overturn it. The upper layers would gain substantial public warrant if several features converged: early reports with clear provenance and genuinely independent access; evidence produced before later institutional incentives hardened; material traces linked by a secure chain to the event; details that were antecedently risky rather than adjusted after the fact; and a pattern markedly less expected under the strongest developed alternatives. No single feature is mandatory. Their value lies in changing the comparative likelihood rather than merely enlarging the story.

The resurrection hypothesis would also gain if its theological background were independently supported. Evidence for a specified God-concept, a reason for incarnation, a reason to vindicate this person, and a reason to use resurrection as the sign would raise the expectation of the event under that model. This is the structure Swinburne openly defends (Swinburne 2003). The point is not that such background must come from history alone. It is that the final public conclusion should disclose how much work comes from historical traces and how much comes from prior theology.

Confidence would fall if apparently separate reports proved to share one route, if key details appeared only in later theological narratives, if source selection explained why supportive reports survived, or if a developed rival predicted the same pattern with fewer discretionary assumptions. Cross-tradition comparison matters as well. If similar testimonial structures support incompatible miracle claims, a method that credits one tradition must identify the asymmetry rather than relying on insider familiarity.

This prospective audit prevents a closed system. A conclusion that names no possible defeater is difficult to distinguish from identity protection. A conclusion that can identify better evidence and real failure conditions remains answerable to the past, even when the past cannot be observed again.

10 Conclusion

Historical knowledge is possible because traces constrain stories. It is difficult because traces arrive through sources, dependencies, selection, and interpretation. The central discipline is to preserve that constraint at every step.

The method developed here replaces a claim package with a ladder. Attestation, communal belief, reported experience, ordinary event, extraordinary event, and theological meaning are distinct targets. Multiple reports gain multiplicative force only to the extent that their evidential routes are independent. Posterior odds combine prior odds with comparative likelihoods. Explanatory virtues matter when they change those terms, not when they are added as a second unpriced scorecard.

Applied to Christian resurrection arguments, this discipline supports a calibrated conclusion. The record can strongly support ordinary claims and early proclamation, and it can take appearance experiences seriously, without thereby establishing a supernatural bodily resurrection or its complete doctrinal interpretation. A defender may close that gap with powerful comparative evidence and independent theology. The gap cannot be closed by narrative momentum.

The decisive contrast is therefore not religious versus nonreligious explanation. It is discriminating evidence versus mere compatibility. A rare claim can become well warranted when a sufficiently distinctive pattern overcomes its prior disadvantage. Dependent reports can contribute when their joint structure is modeled. A theological package can predict evidence when its agent and intentions are independently constrained. These routes remain open; each succeeds only by exposing the work that a compressed story hides.

The principal limit matters. Nothing here proves that a resurrection did not occur, denies theism, or settles private warrant. The result concerns public historical inference: a conclusion deserves no more confidence than its source structure, comparative leverage, and bridge premises earn. Readers should ask one decisive question whenever a modest fact is used to carry a maximal claim: *which step supplied the extra warrant?*

A Formal Reconstruction of the Central Thesis

The paper's conclusion rests on two related arguments. Argument X governs historical promotion from traces to an extraordinary event. Argument Y governs the further promotion from an event to a complete theological interpretation.

Argument X: Public Historical Warrant

- X1.** Public historical warrant for a miracle claim requires its posterior odds against the serious alternative set to support the level of confidence asserted.
- X2.** Those posterior odds depend on prior odds and the joint likelihood of the evidence under the miracle claim and its alternatives.
- X3.** When reports share sources, traditions, or selection processes, multiplying them as independent observations can overstate the joint likelihood ratio.
- X4.** Evidence that directly supports attestation, communal belief, or sincere experience need not have the same likelihood advantage for the external event used to interpret those reports.
- X5.** Defeating one weak rival does not establish the miracle against a serious weighted alternative set, and evidence used to build the background case cannot be counted again without adjustment.
- X6.** A list of minimal facts, considered merely as a list, specifies neither the required dependency structure nor the full comparative likelihood.
- X7. Conclusion:** Minimal historical claims may warrant lower layers strongly, but the list alone cannot establish a miracle conclusion; public warrant at that level requires the additional comparative, dependency, and background work identified above.

X7 is deliberately bounded. A future or more complete argument could supply the missing work; the conclusion is not that a miracle is impossible or that cumulative historical evidence cannot become strong.

Argument Y: Theological Promotion

- Y1.** A complete theological conclusion adds claims about the acting deity, the identity of the person raised, the agent's intention, and the event's doctrinal meaning.
- Y2.** Those claims are not identical to the bare historical claim that an extraordinary return from death occurred.
- Y3.** Evidence for the event transfers to the complete theology only through independent background premises or evidence connecting the event to that specific agent and meaning.
- Y4. Conclusion:** Even strong evidence for a resurrection event would not by itself establish the complete theological package; the added bridge must be defended separately.

Symbols

- *M*: the hypothesis that Jesus underwent a divinely caused bodily resurrection;
- *N*: a weighted mixture of live non-resurrection alternatives;
- *E*: the surviving public historical evidence;
- *B*: relevant background information and model assumptions;

- T : the fuller Christian identity-and-meaning conclusion, which entails M ;
- $P(\cdot | \cdot)$: rational credence in the proposition to the left of the bar conditional on the information to its right.

Probabilistic Structure

The first comparison asks whether E favors M over N once B is included. The second asks how support for M transmits to the fuller conclusion T . In both displays, $P(\cdot | \cdot)$ represents the conditional credence relation being evaluated.

$$\frac{P(M | E, B)}{P(N | E, B)} = \frac{P(M | B)}{P(N | B)} \times \frac{P(E | M, B)}{P(E | N, B)}. \quad (3)$$

Plain-English annotation. The odds for resurrection after considering the evidence equal the prior odds multiplied by the evidence's comparative leverage. The formula permits strong testimony to overcome a low prior; it also shows why compatibility with the evidence is not enough. Dependence and the composition of N affect the likelihood ratio.

$$P(T | E, B) = P(T | M, E, B) P(M | E, B), \quad \text{when } T \Rightarrow M. \quad (4)$$

Plain-English annotation. Because the fuller theology includes the resurrection claim, its posterior also depends on a further conditional bridge: how well the resurrection, evidence, and background support that particular identity and meaning. Establishing M does not make this bridge equal to one by definition.

Validity, Support, and Scope

Argument X is deductively valid at its stated methodological strength: if public warrant depends on posterior odds, and a bare list supplies neither the joint likelihood nor the relevant treatment of alternatives and reused background evidence, the list by itself does not establish the asserted posterior. Its application is partly empirical. Disagreement may concern source relations, which facts belong in E , how N is weighted, whether background support is independent, or the prior odds. A developed cumulative case can answer those questions and thereby exceed the force of the list alone.

Argument Y is valid because the theological conclusion contains more than the event claim. The second formula makes that added bridge explicit. A defender can support the bridge through natural theology, revelation, prophecy, or another argument, but that support is additional to the bare historical inference.

Neither argument assigns a numerical posterior, proves nonoccurrence, forbids miracle testimony, or decides whether private commitment is rational. Together they establish a rule of disciplined promotion: each increase in content requires its own increase in warrant.

References

- Allison Dale C., Jr. (2021). *The Resurrection of Jesus: Apologetics, Polemics, History*. London: T&T Clark.
- Bayes, Thomas (1763). "An Essay towards Solving a Problem in the Doctrine of Chances." In: *Philosophical Transactions of the Royal Society of London* 53, pp. 370–418. DOI: [10.1098/rstl.1763.0053](https://doi.org/10.1098/rstl.1763.0053).
- Bloch, Marc (1953). *The Historian's Craft*. Trans. by Peter Putnam. New York: Alfred A. Knopf.
- Collingwood, R. G. (1946). *The Idea of History*. Oxford: Clarendon Press.
- Earman, John (2000). *Hume's Abject Failure: The Argument against Miracles*. Oxford: Oxford University Press. DOI: [10.1093/0195127382.001.0001](https://doi.org/10.1093/0195127382.001.0001).
- Goodacre, Mark (2001). *The Synoptic Problem: A Way through the Maze*. Sheffield: Sheffield Academic Press.
- Howell, Martha C. and Walter Prevenier (2001). *From Reliable Sources: An Introduction to Historical Methods*. Ithaca, NY: Cornell University Press.
- Hume, David (2007). *An Enquiry Concerning Human Understanding and Other Writings*. Ed. by Stephen Buckle. Cambridge: Cambridge University Press. DOI: [10.1017/CB09780511808432](https://doi.org/10.1017/CB09780511808432).
- Licona, Michael R. (2010). *The Resurrection of Jesus: A New Historiographical Approach*. Downers Grove, IL: IVP Academic.
- Lipton, Peter (2004). *Inference to the Best Explanation*. 2nd ed. London: Routledge.
- McCullagh, C. Behan (1984). *Justifying Historical Descriptions*. Cambridge: Cambridge University Press.
- Sober, Elliott (2015). *Ockham's Razors: A User's Manual*. Cambridge: Cambridge University Press. DOI: [10.1017/CB09781107705937](https://doi.org/10.1017/CB09781107705937).
- Swinburne, Richard (2003). *The Resurrection of God Incarnate*. Oxford: Oxford University Press. DOI: [10.1093/0199257469.001.0001](https://doi.org/10.1093/0199257469.001.0001).
- Tucker, Aviezer (2004). *Our Knowledge of the Past: A Philosophy of Historiography*. Cambridge: Cambridge University Press. DOI: [10.1017/CB09780511498381](https://doi.org/10.1017/CB09780511498381).