

The Prophecy Marker Problem

2 Timothy 3, Vague Traits, and the Failure of Apocalyptic Measurement

Phil Stilwell

Independent Scholar

ORCID: 0009-0009-3725-8682

Revised August 2026

Central claim

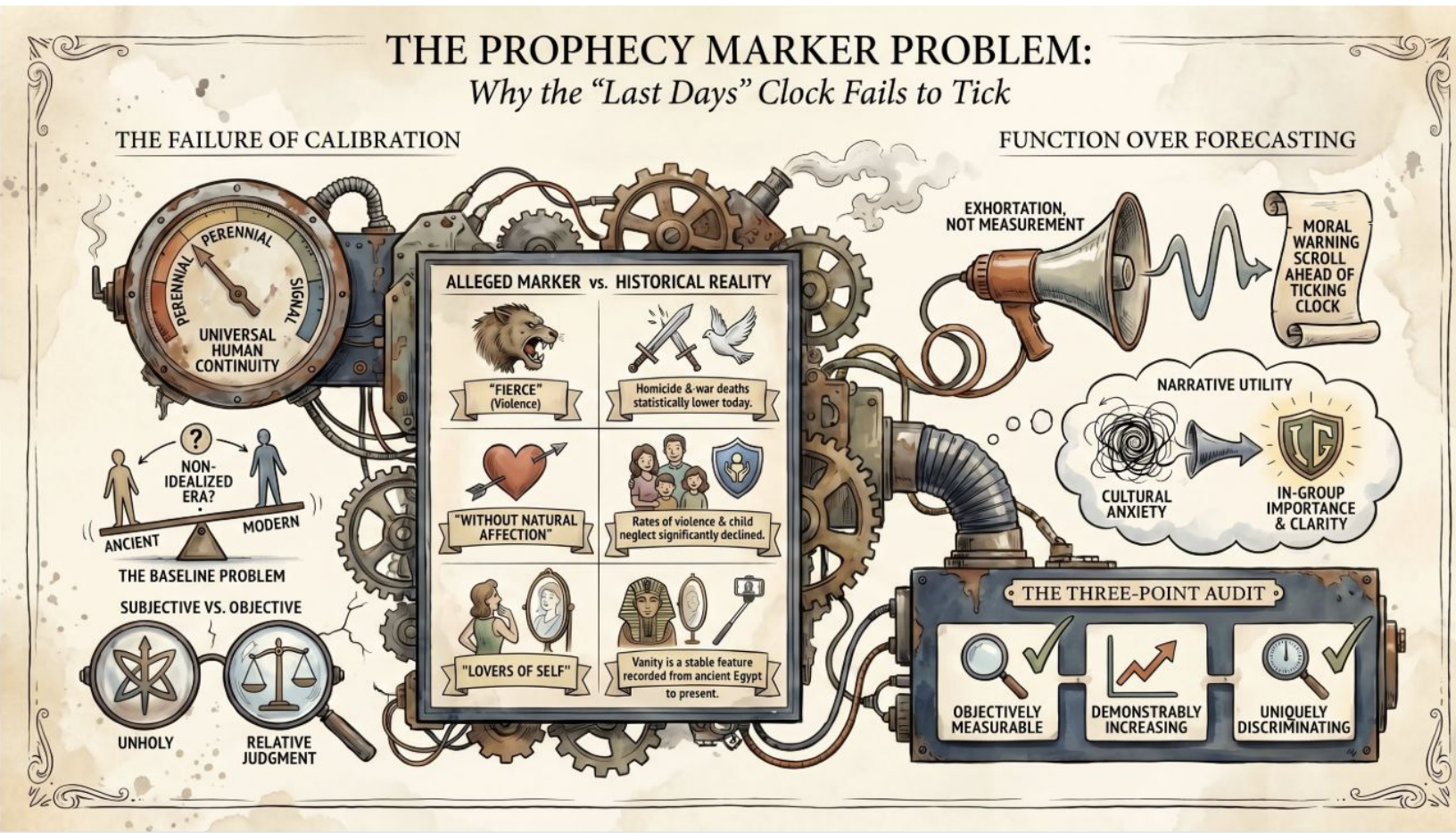
The vice list in 2 Timothy 3 may condemn real harms, but it cannot identify this generation as uniquely terminal. A clock must distinguish one time from another. These traits are undefined, uncalibrated, historically recurrent, and applied by the passage itself to people already present. Calling that a prophecy marker does not make it one.

Preview:

Popular inference	A marker must supply	What the passage supplies	Warranted verdict
People now display the listed vices; therefore, the present is uniquely or climactically the last days.	Stable definitions, a historical baseline, a demonstrated increase, and evidence that the pattern is more expected in a terminal era.	A moralized list whose categories overlap, invite judgment, recur across history, and are immediately applied to contemporaries in verses 5–9.	The list can function as warning or exhortation. Without calibration and discrimination, it cannot function as a clock.

The conclusion is deliberately bounded: it does not disprove every eschatological claim or deny the moral seriousness of cruelty, greed, betrayal, or hypocrisy. It rejects one inference—that noticing such traits now locates us near a unique temporal endpoint.

Visual preview:



The image previews the paper’s audit: an alleged time-marker needs calibration, a baseline, and discriminatory power. Its historical examples are illustrative counterchecks, not a claim that every form of violence or neglect has declined everywhere. The argument rests on the prior failure to define, compare, and distinguish the listed traits; even one credible historical counterpattern is enough to defeat the command to treat present alarm as self-evident measurement.

Abstract

The vice list in 2 Timothy 3:1–9 is often presented as a diagnostic chart for recognizing the last days. The rhetoric is familiar: observe contemporary self-love, greed, disobedience, cruelty, pleasure-seeking, or religious hypocrisy, then conclude that prophecy is visibly being fulfilled. This paper argues that the inference fails even if every moral judgment in the list is granted. A temporal marker must do more than describe conduct. It must define what is being observed, identify a comparison period, establish change, and discriminate a terminal era from ordinary human history.

The text creates an additional problem for the popular reading. Its future-sounding warning is immediately joined to present instructions: avoid “such people”; “among them” are opponents already active; and their progress will be exposed. Regardless of how one settles authorship or chronology, the discourse presents the vice-pattern as contemporary to its first audience. That leaves a dilemma. If “the last days” names a long theological period already underway, the list is not a short-range countdown. If it names a uniquely terminal phase, present counterparts weaken its ability to identify that phase.

The paper develops a four-part audit—definition, baseline, change, and discrimination—and applies it to measurable social harms, evaluative character terms, doctrinal classifications, and self-sealing descriptions of outsiders. It then answers the strongest replies, including intensification, convergence of many weak signs, and spiritual discernment. The conclusion is narrow but firm. The passage may sustain moral exhortation, communal vigilance, or theological hope. It does not, without independent comparative evidence, establish that the present generation has been prophetically singled out. A vice list is not a clock merely because an anxious age reads it while looking at itself.

Keywords: 2 Timothy 3; last days; prophecy; eschatology; apocalypticism; measurement; confirmation bias; moral panic; evidential discrimination

1 A Vice List Is Not Yet a Clock

Claims that the last days are visibly upon us often begin with an invitation that sounds empirical: “Look around.” People are selfish, boastful, disobedient, brutal, treacherous, pleasure-seeking, and religious in form but not in substance. The description resembles 2 Timothy 3, so the speaker treats resemblance as fulfillment. A moral inventory has become a calendar.

That move is too quick. There is a basic difference between saying that a trait exists, saying that it has become worse, and saying that its present level identifies a unique place in history. The first claim needs examples. The second needs a baseline and comparable observations. The third needs evidence that the pattern is substantially more expected in the target period than outside it. Popular use of the passage routinely slides across all three claims while defending only the first.

The hidden inference is easy to expose. An interpreter finds at least one contemporary example of each description. From those examples, the interpreter infers a population-level pattern; from that pattern, an increase over prior eras; and from that increase, a uniquely terminal period. Each step adds content not contained in the observation before it. A news story can establish that one cruel act occurred. It cannot by itself establish that cruelty is typical of a population, more frequent than before, or prophetically distinctive. The conclusion grows while the evidence remains an example.

The resulting argument can be stated in ordinary form:

1. If the listed traits are a clock, their observed pattern must distinguish a terminal period from relevant nonterminal periods.
2. Merely finding current examples does not establish that distinction, because the same kinds of examples occur outside the target period.
3. Therefore, current examples do not by themselves show that the clock is identifying the present.

The second premise is not a skeptical demand for certainty. It is the ordinary requirement that evidence fit the proposed conclusion better than its alternatives. A symptom that appears just as often without a disease cannot diagnose the disease merely because a patient has it. Likewise, a vice that recurs throughout recorded history cannot locate one generation without some added feature that does not recur.

This is not a demand that ancient literature contain a modern statistical appendix. It is a demand placed on a modern empirical use of that literature. Once an interpreter says that current conduct shows that *this* generation occupies a special prophetic moment, the interpreter has made a comparative claim about the world. A comparative claim cannot be rescued by devotional force. It needs a comparison.

The point is also independent of whether Christian eschatology is true. The end could be near for reasons not given in this vice list. A theological tradition might possess other arguments, other texts, or a revealed timetable. None of that would turn non-discriminating traits into discriminating evidence. Evidence must earn its relation to a conclusion; it does not inherit that relation from the importance of the conclusion.

Apocalyptic literature commonly interprets present distress within a larger cosmic drama, locating a community's suffering in a morally charged history (Collins 2016). That narrative function helps explain the emotional force of end-time language. It does not remove the measurement problem. A story may render an age meaningful without uniquely identifying that age. Paul Boyer's history of prophecy belief in modern American culture shows how repeatedly current events have been fitted to inherited prophetic schemes (Boyer 1992). Repeated fit across changing generations is not repeated confirmation. It may instead reveal how elastic the fitting procedure is.

The governing distinction

A moral diagnosis asks whether conduct is cruel, selfish, dishonest, or hypocritical. A temporal diagnosis asks whether a defined pattern distinguishes this period from relevant alternatives. The first can be true while the second is unsupported. The entire popular inference depends on concealing that difference.

2 The Passage Applies Its Warning to Its Own Present

The strongest objection to treating the list as a modern countdown comes from the passage's own movement. Verse 1 announces difficult times "in the last days", and verses 2–5 describe the people associated with them. If one stops there, the list can sound like a portrait held at a distance. But the grammar of the discourse quickly turns toward the first audience. The addressee is told to avoid "such people" in verse 5. Verses 6–8 then describe people who are already entering households, capturing adherents, opposing the truth, and imitating remembered opponents. Verse 9 predicts that *their* progress will not continue indefinitely. The warning's referents are not merely inhabitants of a remote future; the text treats counterparts as present and actionable.

Major commentaries differ on historical, theological, and authorship questions, but they necessarily confront this relation between the announced future difficulty and the opponents addressed in the letter's present setting (Johnson 2001; Towner 2006). The present paper does not require one solution to every exegetical dispute. It needs only the discourse-level fact that the list is not sealed off from the author's own social world.

That discourse-level point is stronger than a loose observation that every era contains bad people. Verse 5 does not merely add an unrelated warning after a remote prediction. The instruction to turn away from "such people" treats the preceding description as relevant to a present decision. Verses 6–8

then explain why: members of the described type are active, recognizable, and harmful within the social field of the letter. The remembered comparison with earlier opponents also places the problem in a recurring pattern rather than presenting it as behavior never before seen.

This does not require claiming that each present opponent instantiates every item in the list or that the passage denies a future culmination. Lists of this kind characterize a type; they do not function like a census in which every member must satisfy every entry. The more modest point is enough: the text's own rhetoric permits recognizable instances of the type before any uniquely modern application. Present resemblance is therefore built into the warning's use. It cannot later be offered as though resemblance itself were the surprising evidence.

Two questions must consequently be kept separate. The first is theological: what period does the phrase "last days" designate within the letter's outlook? The second is evidential: can the listed traits tell a later reader that the designated period is about to end? A broad answer to the first question does not answer the second. Saying that the first audience already lived in the last days may locate them within a theological era, but it gives no scale for measuring how near any later audience is to the era's endpoint.

That fact creates a dilemma for the countdown reading:

1. If "the last days" names a broad theological era that was already underway for the first audience, then the passage may characterize a recurring condition within that era. It does not tell a later reader that the clock is nearly at zero.
2. If "the last days" names a uniquely terminal phase near the end, then the passage's present opponents show that the listed traits also occur outside the alleged final generation. The traits therefore need an additional threshold or pattern before they can discriminate the terminal phase.

Either reading can preserve theological significance. Neither licenses the popular shortcut from "I see these vices" to "therefore my generation is the one described uniquely." The first reading makes the period too long to locate a nearby endpoint; the second requires a measurable intensification that the list itself does not specify.

This textual point matters because it blocks a common rhetorical trick. Modern interpreters often present resemblance between the passage and the present as astonishing. Yet a text that applies the same pattern to its own setting has already told us that resemblance is repeatable. The modern resemblance may be morally troubling, but it is not temporally surprising on the terms of the discourse itself.

3 What a Temporal Marker Must Do

A marker is useful only if it separates possibilities. Smoke can be evidence of fire because smoke is more likely in relevant fire conditions than in relevant non-fire conditions, although steam and dust remind us to compare alternatives. A positive laboratory result matters when the test is defined, calibrated, and known to behave differently when a condition is absent. A temporal prophecy marker needs an analogous structure, even if its evidence is qualitative rather than numerical.

Four requirements are minimal:

Table 1: The four-part audit for an alleged temporal marker.

Requirement	Question	What adequate support looks like	What failure means
Definition	What observable state counts as the trait?	A rule stable enough that different observers can classify relevant cases with reasonable agreement.	The evidence changes with the interpreter’s impressions.
Baseline	Compared with when, where, and whom?	A justified comparison class with attention to population, records, institutions, and reporting practices.	“Worse now” has no fixed meaning.
Change	Has the trait increased beyond ordinary variation?	Comparable observations showing direction, size, duration, and uncertainty.	Existence is mistaken for escalation.
Discrimination	Is the pattern more expected in the terminal period?	A reason the evidence favors the target over nonterminal periods, including recurrent crises.	The trait may condemn conduct but does not locate time.

These requirements are connected. Without a definition, there is nothing stable to count or compare. Without a baseline, an increase has no starting point. Without evidence of change, the interpreter has only presence. Without discrimination, even a real increase may support a social explanation rather than a prophetic endpoint. A recession can increase anxiety and conflict; a war can increase brutality; new media can make old conduct visible. Those are temporally specific explanations, but none by itself entails terminal history.

The audit contains two different comparisons, and confusing them weakens the argument. The first comparison is historical: is a defined trait more prevalent or intense now than in the selected baseline? The second is explanatory: even if it has increased, is that increase more expected if the period is terminal than if it is nonterminal? The historical comparison can succeed while the prophetic comparison fails. For example, a measured rise in betrayal during civil war may be real, yet civil war itself may explain the rise without implying a terminal era. Demonstrating change is therefore necessary but not sufficient.

Definition also involves more than choosing a dictionary meaning. A public claim needs a rule for classifying cases. If “brutal” means homicide in one comparison, harsh speech in another, and political disagreement in a third, the category changes whenever the evidence changes. If “lovers of self” includes every public photograph, career ambition, or request for personal autonomy, a modern society with more records and more individual choice will appear to satisfy the label automatically. The interpreter must specify the observable conduct, the population in which it is assessed, and the threshold at which the description counts as a temporal signal.

Baseline construction likewise requires a fair data-generating comparison. The numerator matters: which events count? The denominator matters: raw incidents, incidents per person, incidents per household, or incidents per year? Visibility matters: were earlier cases recorded, preserved, and accessible? Geographic scope matters: a local crisis cannot establish a global human trend. Time window matters: a brief surge may disappear within a longer series. These are not specialist obstacles placed in front of ordinary belief. They are the content hidden inside the word “worse.”

The last requirement is the controlling one. Elliott Sober emphasizes a general lesson of evidential reasoning: evidence supports one hypothesis over another when it is more expected on one than on its competitor (Sober 2008). If greed, betrayal, vanity, family conflict, and disputed piety are common

under both terminal and nonterminal conditions, observing them does little to choose between those conditions. Emotional vividness does not repair the weak comparison.

Nor must every valid marker be a precise number. Qualitative signs can discriminate when their identifying conditions are stable and when relevant alternatives do not display them equally. The problem with the vice list is not simply that it lacks percentages. The problem is that its key terms are broad, overlapping, morally evaluated, and historically recurrent. Adding decimals to unstable categories would produce false precision, not calibration.

The standard is therefore proportional to the claim. A modest pastoral claim—“we should guard against cruelty and hypocrisy”—may need no historical dataset. A sweeping temporal claim—“these traits identify our generation as uniquely near the end”—assumes comparison across societies and centuries. The evidential burden becomes heavier because the conclusion has become larger, not because moral language has been illegitimately translated into mathematics.

4 Auditing the Trait Families

The list should not be treated as though every item failed in exactly the same way. Some terms can be connected to public measures, although only through contestable choices. Others concern motives or character. Still others are doctrinal classifications that vary with the community applying them. A serious audit separates these families.

Table 2: Different traits create different evidential problems.

Trait family	Examples in common renderings	Main problem	What would be needed
Publicly observable harms	Brutality, abuse, betrayal, slander, family neglect	Records and definitions change; rates vary by place, institution, and exposure.	Defined outcomes, comparable populations, historical records, and a justified trend threshold.
Character and motive judgments	Self-love, greed, pride, recklessness, pleasure-seeking	Motives are inferred; ordinary self-regard and harmful selfishness have no supplied boundary.	A stable construct, reliable indicators, and evidence of population-level change.
Relational judgments	Disobedience, ingratitude, lack of affection, irreconcilability	Authority and obligation are disputed; observers may code dissent as vice.	A defensible account of the duty at issue and independent classification of violations.
Doctrinal judgments	Unholiness, a form of piety without power, resistance to “truth”	The classification can simply reproduce the interpreter’s theology.	Public criteria that do not assume the disputed conclusion used to label the case.

4.1 Observable harms do not arrive pre-calibrated

Violence, neglect, and some forms of betrayal can be investigated, but there is no single measure called “brutality.” Homicide rates, war deaths, domestic abuse reports, corporal punishment, torture, and online harassment capture different phenomena. Trends can move in different directions at the same time. A choice among them must be defended rather than selected because it produces the desired decline narrative.

Consider the difference between a count and a rate. A larger population can produce more violent incidents even while the risk to each person falls. A larger news system can report more distant incidents even while local exposure remains stable. A broader legal definition can increase the official count

because conduct previously ignored is finally recorded. None of these possibilities proves that brutality has declined; they show why an absolute number or a vivid feed of examples is not yet a trend. The interpreter needs a denominator and an account of how the record was produced.

Historical evidence also complicates any effortless claim that current violence is self-evidently unprecedented. Manuel Eisner's study of European homicide estimates marked long-term declines across several regions while emphasizing regional variation, changes in documentation, and the difficulty of explanation (Eisner 2003). That finding does not prove a universal moral ascent; it is not a measure of every war, assault, or unrecorded victim. Its relevance is narrower and decisive: the interpreter cannot treat modern brutality as an obvious, undifferentiated increase requiring no baseline. At least one serious long-run indicator points in the opposite direction.

Modern reporting creates another trap. More documentation can reveal harms that earlier institutions concealed. An increase in recorded cases can reflect a real increase, improved detection, expanded definitions, greater willingness to report, or several processes together. The morally welcome act of making victims visible may therefore be misread as proof that the harm itself is newly rampant. The apocalypse claim needs to separate prevalence from visibility.

A worked audit shows the burden. Suppose "brutality" is defined as intentional homicide per 100,000 people. That definition improves classification and supplies a denominator. The interpreter must then select regions and periods with records reliable enough to compare, establish the direction and durability of change, and explain why the observed pattern favors a terminal-period hypothesis over war, state failure, demographic change, weapon availability, or ordinary fluctuation. If instead brutality is broadened to include every humiliating comment, institutional injustice, or televised atrocity, the historical record becomes less comparable and the inference weaker. The more elastic the category, the easier it is to find examples and the harder it is to use those examples as a clock.

4.2 Character terms expand to fit the observer's alarm

Self-love, pride, greed, recklessness, and pleasure-seeking are not imaginary. They can explain conduct and can become serious moral faults. Yet population-level temporal claims require boundaries. How much self-regard becomes wrongful self-love? Does public self-presentation reveal more vanity than elite monuments, royal titles, heroic inscriptions, or patronage systems in earlier societies? Does consumer choice measure pleasure-seeking, material security, advertising exposure, or all three?

The distinction between access and prevalence is especially important here. Modern people leave photographs, posts, purchases, searches, and recorded speech in quantities unavailable for most earlier populations. That archive may reveal vanity, but it also reveals generosity, grief, mutual aid, and ordinary life. The existence of more observable self-presentation does not show a higher proportion of wrongful self-regard. It may show only that private behavior has become public and preserved.

The same trait can also be coded asymmetrically. A favored leader's confidence becomes courage while an opponent's confidence becomes pride. A believer's devotion becomes conviction while a dissenter's devotion to another view becomes stubborn self-love. Unless classification is possible without first knowing which side a person occupies, the category risks measuring group approval rather than a stable character trait.

Without answers, the terms behave like adjustable lenses. The interpreter can widen a category until the present fills it, then narrow the historical comparison until the past appears cleaner. That is not an innocent lack of technical detail. It permits the conclusion to choose its own evidence.

Apply the four-part audit to "lovers of self." Definition asks which observable pattern separates ordinary self-respect from harmful self-preoccupation. Baseline asks how that pattern could be compared across societies that distribute status, literacy, privacy, and public speech differently. Change asks

whether the defined pattern has risen rather than merely become easier to display and record. Discrimination asks why any rise is more expected near a terminal endpoint than under commercialization, urbanization, new communication tools, or shifts in social norms. Until those questions are answered, the label may carry moral force but no temporal location.

4.3 Relational and doctrinal terms can code dissent as fulfillment

Terms such as disobedient, ungrateful, unholy, and opposed to truth depend heavily on who is said to possess legitimate authority and correct doctrine. A reformer may be disobedient to an abusive parent, church, or state precisely because obedience would perpetuate harm. A person leaving a religious community may be described by the community as proud, ungrateful, pleasure-seeking, and unable to accept truth. The labels can be sincere and still fail as independent evidence.

The phrase about people always learning without reaching knowledge is especially vulnerable to this problem. If learning counts as successful only when it returns the interpreter's preferred doctrine, then contrary inquiry confirms the prophecy by definition. Agreement becomes evidence of truth; disagreement becomes evidence that the dissenter was foretold. A classification that cannot permit an informed counterexample is not measuring epistemic failure. It is policing an outcome.

The same issue affects claims about having a form of piety while denying its power. Competing Christian communities can direct that charge at one another. Its doctrinal force may matter within a tradition, but its public discriminatory force is weak because the classification depends on the very theological authority under dispute.

This creates a symmetry test. If Catholic, evangelical, Pentecostal, Orthodox, and restorationist communities can apply the same description to rivals, the description does not by itself identify which community has correctly recognized prophetic fulfillment. The problem is not that all doctrinal judgments are false. It is that a disputed doctrinal judgment cannot serve as independent public confirmation of the authority used to make that judgment. Otherwise, every side can certify itself and count disagreement as the predicted failure of outsiders.

5 The Baseline Cannot Be an Imagined Golden Age

Claims of decline often compare the most visible failures of the present with the remembered ideals of the past. The comparison is structurally unfair. The present is available as raw footage; the past is available as a selected story. Current scandals are documented in detail, while ordinary historical cruelty, private exploitation, and unrecorded victims are easily lost. At the same time, admired past norms are remembered more readily than the people excluded or harmed by them.

The solution is not to reverse the bias and declare every modern trend an improvement. It is to specify a baseline. Compared with the Roman imperial world, medieval Europe, the early modern period, the industrial nineteenth century, the world wars, or the late twentieth century? Across which region and social class? Using which outcome? Over what duration? The answer can differ across traits, but it cannot remain unstated.

Several choices can make a decline story appear or disappear. A short window beginning at an unusually calm point can make a later crisis look like unprecedented collapse. A window beginning during war can make the same later period look improved. A national average can hide a severe local rise, while a collection of local crises can hide a stable or falling national rate. Selecting only the conduct that an earlier society recorded can erase harms it normalized or concealed. A fair comparison must explain why its starting point, end point, population, and record system are relevant to the claimed temporal conclusion.

The basic fairness rule is simple: compare observed conduct with observed conduct, rates with rates, and similar populations under acknowledged differences in documentation. Do not compare present failures with past ideals. Do not compare the most alarming subgroup now with an entire civilization then. Do not move between absolute totals and per-person risk depending on which favors decline. These rules do not guarantee a correct history, but violating them guarantees that the claimed history has not been established.

Moral-panic research helps explain why the unstated baseline feels persuasive. Stanley Cohen's analysis describes how societies identify threatening figures, simplify them into recognizable deviance, amplify incidents, and authorize a response around the constructed threat (Cohen 2011). Not every declaration of moral decline is a moral panic, and the term should not be used as an insult. The useful point is that public attention is socially organized. What receives intense coverage can feel common because it is salient, not because a valid historical comparison has shown that it is new.

Psychological research identifies a related mechanism. People more readily notice and interpret evidence compatible with a favored belief, a family of tendencies commonly discussed under confirmation bias (Nickerson 1998). Motivation does not make reasoning wholly arbitrary, but it can influence which evidence is sought and what standards are applied (Kunda 1990). The end-time frame gives this tendency a ready script: alarming evidence is relevant; counterevidence is complacency, deception, or a temporary pause.

That asymmetry is detectable without diagnosing any individual's motives. Ask what observation would count against the claim. If worsening confirms the prophecy, improvement confirms a deceptive calm, mixed evidence confirms confusion, and missing data confirms concealment, then the interpretation has ceased to risk error. Its resilience is purchased by making the world unable to answer back.

This is why the baseline problem is not merely a missing footnote. It changes the logical status of the appeal. Without a fixed comparison, "the world is getting worse" is not one proposition waiting to be tested. It is a family of shifting comparisons from which the interpreter can select after seeing the result. A moving baseline lets the claim win by changing the question.

6 Why Many Weak Signs Do Not Automatically Become Strong

A defender may grant that no single trait is decisive but appeal to convergence: perhaps self-love, greed, betrayal, family conflict, hostility to religion, and pleasure-seeking become powerful when observed together. In principle, multiple weak indicators can combine into strong evidence. Medical diagnosis, historical inference, and scientific modeling often rely on converging lines of support. But the strength of a bundle depends on the quality and independence of its parts.

The vice-list bundle has three problems. First, the categories overlap. Pride, boasting, self-love, and arrogance may be several descriptions of the same judgment rather than independent observations. Brutality, lack of affection, and irreconcilability may arise from one conflict. Counting related labels as separate confirmations exaggerates the evidence.

Second, the same interpreter often classifies every item through the same prior narrative. A cultural change is first judged rebellious, then coded as disobedience, ingratitude, pride, love of pleasure, and hostility to truth. Five labels have not produced five observations; one reaction has produced five labels.

Dependence can also occur upstream. Several preachers, channels, or communities may repeat the same cluster of stories because they draw from the same news cycle, translation, ministry network, or viral compilation. Agreement among the reports then reflects a shared source rather than independent observation. Repetition can increase confidence psychologically while adding little new information evidentially.

Third, a conjunction helps only when it is rare or otherwise more expected under the target condition. Human societies regularly contain greed, status-seeking, family conflict, ideological accusation, cruelty, pleasure, and hypocrisy together because large societies contain many people and institutions. The relevant question is not whether examples of all items can be found. With a broad enough search, that result is nearly guaranteed. The question is whether a defined joint distribution has changed and discriminates the alleged terminal phase.

The global search space makes this problem severe. Billions of people generate an immense stream of crimes, scandals, speeches, conflicts, entertainments, and religious disputes. If an interpreter may choose one example of pride from one country, one example of brutality from another, and one example of family conflict from a third, the assembled list does not describe one population or trend. It describes the reach of the search. A prophecy can appear globally fulfilled because the evidence was gathered globally while counterexamples and denominators were omitted.

The word “simultaneous” does not solve the problem unless it is defined. Traits lasting for decades will naturally overlap. A broad term such as greed will be present somewhere at nearly every moment. When each item has a long duration and a wide geographic scope, finding them at the same time is not a rare conjunction. The interpreter needs a time window fixed in advance and evidence that the joint pattern is more concentrated than ordinary overlap would produce.

Table 3: How apparent convergence can be manufactured.

Appearance	Hidden dependence	Why the count is inflated	Corrective question
Several vice terms fit one person.	The terms partly describe the same character judgment.	Rewording one impression looks like multiple evidence.	Would different observers classify each trait independently?
Several current events fit the list.	The events were selected from a vast global stream.	Search volume guarantees matches while nonmatches disappear.	What sampling rule was fixed before looking?
Several communities report decline.	They share texts, media, and expectations.	Repeated testimony may trace one information network.	Are the reports independent and comparably measured?
Every outcome fits a warning.	The interpretation changes after the result.	No possible result lowers confidence.	What observation would count against the reading?

Convergence is therefore not magic. It cannot transform undefined, dependent, selectively gathered impressions into a calibrated historical signal. More unmeasured items can make a narrative feel dense while adding almost no discriminatory information.

Genuine convergence would look different. Each trait would have an independently usable definition; evidence would be gathered by a declared rule rather than selected after the fact; overlapping categories would be modeled as related rather than counted separately; the same population and time window would be used where possible; and the resulting joint pattern would be compared with both historical baselines and non-prophetic explanations. Until then, “many signs” is a description of the interpreter’s list, not a measurement of history.

7 Why the Reading Persists

The failure of measurement does not imply cynical manipulation. End-time readings can persist because they perform genuine social and psychological work. They place disruption inside a meaningful story,

assure a community that its moral alarm has cosmic significance, identify threats, and turn uncertainty into vigilance. These functions explain attractiveness; they do not establish truth.

Research on failed prophecy is relevant but must be used with care. The classic study by Festinger, Riecken, and Schachter proposed that some highly committed adherents respond to disconfirmation by reinterpretation and increased recruitment rather than abandonment (Festinger et al. 1956). Later reviews emphasize that responses vary with commitment, social support, prediction specificity, and available reinterpretations (Dawson 1999). It would be crude to classify every end-time believer as a member of one failed-prophecy group. The narrower lesson is methodological: a prediction can survive failure when a community has flexible ways to redescribe the result.

The vice-list reading offers exceptional flexibility. It specifies no rate, threshold, geographic scope, or deadline. A new technology can count as vanity, pleasure, deception, or loss of affection. Political conflict can count as treachery, brutality, disobedience, or slander. Religious disagreement can count as empty piety or resistance to truth. If one alleged sign declines, attention can shift to another. If a generation passes, the same list remains available to the next.

The every-outcome structure can be displayed plainly:

1. If an alleged vice worsens, the worsening is counted as prophetic confirmation.
2. If it improves, the improvement may be called temporary calm, concealment, or moral numbness and is also absorbed by the framework.
3. If the record is mixed, the interpreter selects the worsening strand or changes the scale of comparison.
4. Therefore, no ordinary outcome is allowed to lower confidence in the temporal reading.

The first three statements describe an interpretive policy; the fourth identifies its evidential cost. A view may remain meaningful under all outcomes, but it cannot then cite selected outcomes as risky predictions that succeeded. Protection from disconfirmation and public confirmation cannot both be claimed from the same flexibility.

That flexibility produces a one-way evidential channel. Events enter as confirmation, but almost nothing exits as disconfirmation. The resulting confidence may be psychologically understandable and communally reinforced. It is still epistemically defective. A method that cannot lose is not thereby undefeated.

The social functions of the reading help explain this defect without reducing belief to bad faith. A decline narrative can produce urgency, solidarity, a clear moral boundary, and a sense that the community understands events outsiders misunderstand. Those benefits are real experiences for participants. Yet usefulness and evidential discrimination are different properties. A map can comfort travelers without accurately representing the road. Explaining why a reading persists does not refute it, but it does remove persistence itself as evidence that the reading has survived rigorous tests.

8 The Strongest Replies

8.1 Reply 1: the prophecy concerns intensification, not novelty

The most serious response says that the listed traits need not be new at the end; they need only become unusually intense, widespread, or concentrated. This proposal is coherent and is compatible with the passage's present opponents. It also concedes the central criticism: recurrence is not enough. Once intensification is doing the work, the interpreter must identify what has intensified and how the increase is known.

Three additions are required. First, the trait needs a stable scale. An increase in the number of recorded incidents, the proportion of people involved, the severity of conduct, and the visibility of

conduct are different changes. Second, the comparison needs a time and population. “More than before” is empty until “before” is fixed. Third, the interpreter needs a threshold separating terminal intensification from severe but nonterminal episodes such as war, famine, institutional collapse, or authoritarian rule.

The passage provides no public threshold for these comparisons. One may add a theological premise that the decisive intensification will be spiritually recognizable, but then that premise is carrying the conclusion. The familiar traits are no longer independent evidence; they are material classified as terminal because the theological premise already says the terminal pattern has arrived. The response preserves possibility while surrendering the claim that present resemblance is self-authenticating evidence.

8.2 Reply 2: ordinary people can recognize a pattern without statistics

They sometimes can. Smoke, mass flight, empty shelves, or collapsing institutions can reveal a change before formal study is complete. Ordinary judgment is indispensable to inquiry. The issue is not whether a person may recognize anything without a spreadsheet; it is whether the relevant pattern is stable, comparable, and connected to the conclusion being drawn.

Informal recognition works best when cases are clear and alternatives are sharply different. A street that was full yesterday and empty today supplies a direct comparison. The vice list is unlike that example. It combines public actions, inferred motives, contested duties, and doctrinal verdicts. Different observers can look at the same conduct and disagree about whether it is confidence or pride, independence or disobedience, recreation or pleasure-worship, faithful piety or empty form.

Statistics are only one way of disciplining judgment. Historical documents, structured case comparisons, independent observers, and clearly stated qualitative thresholds can also provide evidence. But every adequate method must answer the same ordinary questions: which cases count, how were they selected, compared with what, and what would have looked different if the conclusion were false? Refusing numerical precision does not remove those questions. It makes careful answers to them more important.

8.3 Reply 3: many details together are too exact to be coincidence

The apparent specificity of a long list can feel impressive. If many descriptions fit at once, coincidence seems less plausible than fulfillment. That reasoning would be strong if the descriptions were narrow, independently classified, and unlikely to co-occur outside the target period. The actual list does not satisfy those conditions.

Many entries are semantically related: self-love, boasting, pride, and conceit can arise from one evaluation. Others are common features of conflict: slander, treachery, irreconcilability, and brutality often cluster whenever groups fracture. Still others are labels assigned from within a doctrinal outlook. A long list of broad and dependent descriptions is easier, not harder, to fit after the fact because it gives the interpreter more entry points.

The coincidence intuition also ignores search size. Across billions of people and continuous global reporting, examples of each broad trait are inevitable. What needs explanation is not that matches can be assembled, but whether a predeclared, independently measured pattern occurs at a rate that nonterminal history would make surprising. Without that comparison, length creates an impression of precision rather than actual precision.

8.4 Reply 4: the present application foreshadows a fuller future pattern

This reading can reconcile the text’s contemporary opponents with a future climax: the opponents are an early instance, while a final generation will display the type more fully. Nothing in the present argument proves that such a culmination is impossible. The problem is evidential. If the earlier instance

and final instance share the same listed traits, the traits alone cannot tell us which instance we are observing.

The foreshadowing response therefore needs an added discriminator: greater scale, unusual intensity, a defined global distribution, a unique conjunction, or an independent sign. Each option could in principle distinguish culmination from recurrence, but each must be specified and supported. Without that addition, “foreshadowing” merely labels earlier counterexamples as expected and later similarities as fulfillment. It prevents the present application from refuting the theology while leaving the modern countdown unsupported.

8.5 Reply 5: modern communication makes the pattern genuinely global

Modern communication makes distant conduct visible and can spread practices rapidly. A defender may argue that earlier generations had local examples, whereas the current generation displays the pattern globally and simultaneously. This is a real empirical proposal, not a mere repetition of “look around.” It still needs careful separation of transmission, observation, and prevalence.

A global media system produces global awareness even if underlying rates differ across regions or remain stable. It can also synchronize vocabulary: people learn to describe diverse events through the same prophetic list. That creates simultaneous reporting, not necessarily simultaneous intensification. To establish the stronger claim, the interpreter would need comparable measures across regions, a fixed time window, and evidence that the joint pattern exceeds what population size, connectivity, urbanization, conflict, and improved detection would predict.

Global reach may therefore be a feature of the information system rather than of the vice-pattern. The response identifies a potentially testable distinction, but it does not get the desired conclusion for free. Greater access to examples is precisely why raw example-counting becomes less reliable.

8.6 Reply 6: prophecy should not be held to laboratory standards

The objection is partly correct. Historical and theological claims cannot always be tested like chemicals under controlled conditions. Evidence can be incomplete, qualitative, and interpreted through a tradition. The four-part audit does not require laboratory control. It requires only enough stability to support the claim actually being made.

The more public and comparative the claim, the more public and comparative its support must be. Saying that a passage speaks morally to a community can be defended through textual and theological interpretation. Saying that current humanity is measurably worse and therefore uniquely near an endpoint adds claims about populations, history, and discrimination. Those additions create their own evidential responsibilities.

Lowering the standard cannot be one-sided. If impressions are sufficient for the favored prophecy, then equally sincere impressions supporting rival end-time readings, other religions, or secular collapse narratives gain the same entitlement. A method that cannot distinguish among incompatible applications cannot establish one of them. The standard is not scientism; it is symmetry.

8.7 Reply 7: spiritual discernment supplies what public measurement cannot

Spiritual discernment may be meaningful within a faith community. It cannot serve as public evidence when rival communities claim incompatible discernment and no independent test settles the difference. If the conclusion is warranted only after accepting the interpreter’s spiritual authority, then the vice list is not independently confirming that authority’s prophetic reading.

This does not demand that religious reasoning contain no tradition-specific premises. It demands honesty about where the weight lies. A believer may say, “My tradition teaches that this is the terminal pattern.” That is different from saying, “Anyone can see from current evidence that the pattern

uniquely identifies our time.” The first is an internal theological commitment. The second claims public confirmation and must expose its comparison to public criticism.

The symmetry problem returns here as well. Rival communities can report incompatible discernment with equal sincerity. If discernment authenticates itself, the method cannot adjudicate among them. If its fruits, coherence, or predictive success are offered as tests, then discernment has re-entered the field of comparison and should be assessed by the same evidential discipline.

8.8 Reply 8: the moral warning matters even if the date does not

Agreed. Cruelty, deception, exploitative greed, betrayal, and hypocrisy can deserve resistance in any era. That agreement strengthens rather than weakens the distinction. A recurring moral warning does not need to be converted into a failed measuring instrument. Its value may lie precisely in addressing every audience that can recognize the conduct among itself.

Indeed, detaching exhortation from countdown can make the warning more demanding. A community cannot treat the list mainly as a portrait of outsiders or as proof of its own prophetic insight. It must ask where its own leaders use households, status, piety, or claims to truth exploitatively. The passage’s present-directed instruction supports vigilance toward actual conduct rather than speculation about a calendar.

The temporal inference adds no moral duty that the moral judgment did not already supply. If brutality is wrong, it should be resisted whether history ends soon or continues for centuries. If the claim of imminent termination distracts from measurement, prevention, or long-term repair, the prophetic overlay can weaken the very moral response it is supposed to intensify.

9 A Better Reading and a Better Practice

The passage can be read as moral and communal exhortation rather than as a device for dating history. On that reading, its force turns inward before it turns outward. The list warns a community about recognizable failures of character and authority. The immediate instruction to avoid “such people” concerns conduct already affecting the community, not permission to scan the surrounding culture for proof of superiority.

This reading fits the practical shape of the warning. An exhortation asks what conduct should be resisted, which relationships are dangerous, and how a community should remain faithful under pressure. A forecast asks when an event will occur and which observations locate the hearer relative to it. The passage can perform the first task without successfully performing the second. Confusing the tasks does not honor the text by giving it more force; it assigns the text a measuring function that its categories are not equipped to carry.

This reading also avoids a moral hazard in modern use. When a community treats outsiders’ vices as prophetic confirmation, other people’s suffering can become useful scenery in its own cosmic story. Social breakdown, abuse, war, or alienation is no longer only a harm to prevent; it is also a sign that validates the interpreter. A non-calibrated prophecy frame can therefore reward attention to decline while weakening attention to counterevidence and repair.

A better practice asks four plain questions whenever 2 Timothy 3 is offered as a current time-marker:

1. **Definition:** What exactly is being observed, and could a person who rejects the conclusion still classify it by the same rule?
2. **Baseline:** Which past period and population provide the comparison, and why are the records comparable?

3. **Change:** What shows a durable increase rather than greater visibility, selective attention, or ordinary fluctuation?
4. **Discrimination:** Why is the pattern more expected near a unique endpoint than during recurrent social crisis or ordinary human history?

The questions can be used without specialist training. Suppose a speaker points to a series of violent reports and says that brutality proves the last days are near. Ask what kind of violence is being counted and whether the claim concerns total incidents, per-person risk, severity, or visibility. Ask which earlier population supplies the baseline and whether its cases were recorded comparably. Ask whether the trend persists across a reasonable window. Finally, ask why the trend favors a terminal phase over war, political breakdown, media expansion, or another ordinary cause. If the answers remain “it is obvious,” then the argument has not moved beyond its starting impression.

A responsible religious claim could be more modest. An interpreter might say that the passage gives a standing moral vocabulary for identifying destructive conduct, that a particular community currently faces several of those dangers, and that the community should respond with vigilance and repair. Such a claim can be evaluated through close description of the local conduct. It does not need to pretend that a local moral judgment has measured all human history.

The distinction also changes practice. If cruelty is counted mainly as a sign, the interpreter has an incentive to collect examples. If cruelty is treated mainly as a harm, the incentive is to define it, identify causes, protect victims, and test interventions. The first posture turns suffering into confirmation; the second treats suffering as a problem requiring responsibility even if no countdown follows.

If those questions cannot be answered, the responsible conclusion is not that the vices are harmless. It is that the temporal claim is unsupported. Moral urgency can remain. Prophetic certainty should not.

10 Conclusion

The popular use of 2 Timothy 3 makes a large claim from weak materials. It notices familiar human failures, applies broad evaluative labels, omits a defensible historical baseline, and declares a unique temporal diagnosis. The passage itself then complicates the maneuver by treating the listed type of people as already present to its first audience.

The cumulative case is stronger than any one criticism alone. Textually, the warning turns toward contemporaries, so resemblance is repeatable within the discourse. Conceptually, the existence of a vice is not its increase, and an increase is not its temporal discrimination. Historically, comparisons change with definitions, denominators, records, regions, and time windows. Epistemically, overlapping labels and shared sources can create an illusion of convergence. Socially, an interpretation that absorbs worsening, improvement, and mixed evidence can persist without having passed a risky test.

The result is a dilemma that the “look around” rhetoric cannot escape. If the last days already encompassed the text’s original setting, the list describes a long or recurring condition and does not locate the modern reader near the final moment. If the phrase identifies a uniquely terminal phase, the interpreter must demonstrate a special intensity or pattern that the text does not calibrate and current impressions do not establish.

Nothing in this argument proves that history has no endpoint, that every prophetic claim fails, or that the list lacks moral insight. The conclusion is more exact and more damaging to the popular inference: these traits do not mark the present as uniquely terminal merely because they can be found here. They could be found by the text’s first audience and by generations between. A clock that reads the same hour across centuries is not a clock. It is a mirror.

The requested intellectual change is therefore straightforward. Stop treating recognition as measurement. When the list is used morally, examine the conduct and act against the harm. When it is used temporally, require a definition, baseline, demonstrated change, and discriminatory comparison. If those are absent, say so plainly. The importance of an endpoint cannot compensate for evidence that does not locate it.

A Formal Reconstruction of the Central Thesis

This appendix makes the central reasoning explicit. The symbols are compact, but no symbol is allowed to do unexplained work.

A.1 Syllogistic reconstruction

1. A set of observed traits warrants a uniquely terminal period only if the set is defined, compared with a justified baseline, shown to change, and shown to discriminate that period from nonterminal alternatives.
2. The 2 Timothy 3 vice-pattern is applied to contemporaries in the discourse, and modern use of the list supplies neither a calibrated increase nor a discriminating comparison.
3. Therefore, merely observing the listed traits now does not warrant the conclusion that the present is uniquely or climactically terminal.

The conclusion follows from the two premises. The argument does not infer that the conduct is acceptable or that every eschatological claim is false; those claims are outside its premises.

A.2 Symbols

Let:

- T : “the present period is a uniquely or climactically terminal period”;
- M_i : “trait i is observed in the present”;
- D_i : “trait i has a stable definition”;
- B_i : “trait i has a justified historical baseline”;
- I_i : “trait i shows a demonstrated increase beyond ordinary variation”;
- R_i : “trait i discriminates the terminal period from relevant nonterminal periods”;
- E : the total observed bundle of alleged markers;
- C_i : the full calibration condition for trait i ;
- C_E : bundle-level calibration, including the four audit conditions and control for dependence among the traits;
- LR_i : the likelihood ratio for trait i ;
- n : the total number of listed markers being combined;
- A : “the listed type is applied to people present to the first audience”;
- H : “the first audience was followed by many later generations”;
- L : “the last days names a long era already underway”.

The subscript i simply numbers the traits. For example, M_1 might stand for observed greed and M_2 for observed brutality. The letter choices are labels, not evidence.

A.3 Presence does not entail temporal location

The popular shortcut can be represented as:

$$M_1 \wedge M_2 \wedge \cdots \wedge M_n \therefore T. \quad (1)$$

Plain-English annotation. The symbol $\wedge$ means “and”; $\therefore$ means “therefore.” The line says: several listed traits are present, so the present must be the unique terminal period. This is not a valid rule of logic. The same traits may be present in many nonterminal periods.

To make any one trait a defensible temporal marker, at least the following conjunction is needed:

$$C_i = D_i \wedge B_i \wedge I_i \wedge R_i. \quad (2)$$

Plain-English annotation. C_i means that trait i is calibrated for temporal use. Calibration requires all four parts: a definition, a baseline, an increase, and discriminatory power. If even one part is missing, the trait may still support a moral judgment, but it is not yet a reliable clock.

The crucial relation is discrimination. In likelihood form:

$$\text{LR}_i = \frac{P(M_i | T)}{P(M_i | \neg T)}. \quad (3)$$

Plain-English annotation. $P(M_i | T)$ means the probability of observing trait i if the present really is terminal. $P(M_i | \neg T)$ means the probability of observing it if the present is not terminal. The ratio compares those expectations. The symbol $\neg$ means “not.”

If a trait is common in both kinds of period, then:

$$P(M_i | T) \approx P(M_i | \neg T) \Rightarrow \text{LR}_i \approx 1. \quad (4)$$

Plain-English annotation. $\approx$ means “approximately equal”; $\Rightarrow$ means “implies.” A likelihood ratio near 1 means the observation fits the terminal and nonterminal possibilities about equally well. It therefore gives little or no reason to choose one over the other, however disturbing the trait may be.

A.4 Why a bundle does not automatically rescue the inference

Bayesian odds display what a bundle would have to accomplish:

$$\frac{P(T | E)}{P(\neg T | E)} = \frac{P(T)}{P(\neg T)} \times \frac{P(E | T)}{P(E | \neg T)}. \quad (5)$$

Plain-English annotation. The left side is the odds for a terminal period after considering the evidence. The first ratio on the right is the starting odds. The second is the evidential comparison. A long list helps only if the whole observed pattern is more expected in a terminal period than outside one. Finding many recurrent or overlapping labels does not by itself make that ratio large.

It is generally unsafe to multiply each trait’s force as though the traits were independent:

$$P(E | T) \neq \prod_{i=1}^n P(M_i | T) \quad \text{when the traits or classifications depend on one another.} \quad (6)$$

Plain-English annotation. $\prod$ means “multiply the following values.” The unequal sign warns that pride, boasting, and self-love may be overlapping descriptions, and that one interpreter’s prior judgment may generate several labels at once. Treating those labels as independent confirmations would count the same evidence more than once.

A.5 The text’s present application and the countdown dilemma

Let A mean “the listed type is applied to people present to the first audience”, and let L mean “the last days names a long era already underway.” Then the interpretive alternatives can be summarized:

$$A \wedge L \Rightarrow \neg(\text{short-range countdown from the list}), \quad (7)$$

Plain-English annotation. If the listed people are already present and the last days is a long era, then the list may describe that era broadly. It does not tell a later reader that only a short time remains.

Alternatively, include the historical fact H that many generations followed the first audience:

$$A \wedge H \wedge \neg L \Rightarrow \text{the traits recur before the unique terminal phase.} \quad (8)$$

Plain-English annotation. If the last days is not a long era but a unique final phase, and many generations followed the text’s first audience, then the text’s present examples occur before that final phase. An added threshold or pattern is therefore required to identify the phase. The passage does not provide a calibrated threshold.

The bounded conclusion is:

$$[A \wedge H \wedge \neg C_E] \Rightarrow \neg(E \text{ warrants } T). \quad (9)$$

Plain-English annotation. C_E means that the whole evidence bundle is calibrated, including dependence among its parts. Because the pattern is applied to a first audience followed by later generations and the familiar evidence bundle lacks that calibration, merely observing the bundle does not warrant the claim that our period is uniquely terminal. This conclusion rejects the temporal inference. It does not say the conduct is good, the passage has no moral value, or every possible eschatological argument is false.

A.6 Validity of the inference

The first displayed line, equation (1), is intentionally marked as an invalid shortcut: the presence of the traits does not logically entail T . The calibrated argument has a valid conditional structure. If temporal warrant requires the four audit conditions, and the alleged evidence lacks bundle-level calibration C_E , then the evidence does not supply that warrant. Likewise, the two branches of the textual dilemma are valid at the stated level: if L is true, the list does not yield a short-range countdown; if L is false while A and H are true, the recurrent traits need an added discriminator. Neither branch concludes that every possible route to T is false.

The likelihood equations do not deductively prove that the relevant ratios equal 1. They state what would have to be compared. A ratio near 1 would show little discrimination; a ratio reliably above 1 could supply

support. The body argues that broad, recurrent, dependent, and selectively classified traits have not been shown to produce the latter result. That is an evidential judgment, not a theorem derived from the notation.

The conclusion concerns warrant from this evidence, not the truth-value of T . From “ E does not warrant T ” it does not follow that T is false. The endpoint could be near for another reason, or the relevant evidence could exist but not yet have been supplied. This distinction avoids an argument from ignorance. The paper does not infer “the claim is false because the data are missing”; it infers “the familiar appeal has not earned its stated conclusion because the observations supplied do not discriminate it.”

A.7 Support for the premises

The present-application premise A is interpretive but directly grounded in the movement from verses 1–5 to the instruction concerning “such people” and the active opponents in verses 6–9. The calibration premise is conceptual: a claim that locates one period rather than another requires stable content and a relevant comparison. The claim that modern applications lack adequate calibration is defeasible and case-sensitive, but it is supported by the ordinary form of the appeal audited here—current examples assembled without fixed definitions, baselines, sampling rules, trend thresholds, or a stated disconfirming outcome.

An interpreter could challenge the second premise by supplying those missing elements. Doing so would not refute the audit; it would finally subject the temporal claim to it. The conclusion would then depend on the resulting comparison. The formal reconstruction therefore preserves a live evidential alternative while explaining why the familiar uncalibrated appeal does not yet warrant T .

The premises also differ in strength and kind. The requirement that a temporal marker discriminate is conceptual and very strong: without discrimination, the word “marker” has lost the relevant function. The claim that the list applies to contemporaries is a well-grounded textual judgment, although interpreters may disagree about how present and future instances relate. The claim that modern applications are uncalibrated is a generalization about a familiar argumentative form, not every statement any interpreter could make. It is strongest against appeals that offer selected present examples and weakest against an application that supplies independent definitions and comparative evidence. The conclusion is deliberately matched to that scope.

A.8 What a successful countercase must show

The audit is not designed so that the temporal reading must fail. A countercase can succeed by providing the elements the familiar appeal omits:

1. **Stable classification.** Define the relevant trait or joint pattern so that observers who do not already accept T can identify cases by substantially the same rule.
2. **A fair comparison.** Specify populations, regions, records, denominators, and time windows, then explain why differences in documentation do not create the trend.
3. **Demonstrated change.** Show a durable increase in prevalence, severity, or concentration rather than an increase in total population, media exposure, legal recognition, or selected examples.
4. **Prophetic discrimination.** Explain why the measured pattern is more expected if T is true than under nonterminal alternatives such as war, institutional failure, technological change, or ordinary recurrence.
5. **Independence and risk.** Avoid counting overlapping labels or shared reports as separate evidence, and state what possible result would lower confidence in the temporal interpretation.

Meeting these conditions would not automatically prove T . It would convert a rhetorical resemblance into evidence that can be evaluated. The final judgment would still depend on the quality of the measurements, the strength of rival explanations, and the size of the resulting likelihood difference. But the argument would at least possess the structure required of a temporal marker.

Failure at one condition has a specific effect. An unstable definition makes the observations movable. A missing baseline makes change unknown. A missing trend reduces the evidence to present existence. A missing discriminator leaves a real trend compatible with nonterminal history. Dependence inflates the apparent number of signs. An every-outcome interpretation prevents the result from testing anything. These are cumulative defects, but each can independently block the advertised inference.

The formal result can therefore be challenged rather than merely rejected. Supply a stable marker, a fair historical record, an increase, and a comparison on which the increase is meaningfully more expected under T . Until that work is done, the vice list may remain theologically significant and morally urgent, but the familiar present-day examples do not tell the reader where the clock stands.

Bibliography

- Boyer, Paul (1992). *When Time Shall Be No More: Prophecy Belief in Modern American Culture*. Cambridge, MA: Belknap Press of Harvard University Press.
- Cohen, Stanley (2011). *Folk Devils and Moral Panics*. 3rd ed. London: Routledge.
- Collins, John J. (2016). *The Apocalyptic Imagination: An Introduction to Jewish Apocalyptic Literature*. 3rd ed. Grand Rapids, MI: Eerdmans.
- Dawson, Lorne L. (1999). "When Prophecy Fails and Faith Persists: A Theoretical Overview." In: *Nova Religio* 3.1, pp. 60–82. DOI: [10.1525/nr.1999.3.1.60](https://doi.org/10.1525/nr.1999.3.1.60).
- Eisner, Manuel (2003). "Long-Term Historical Trends in Violent Crime." In: *Crime and Justice* 30, pp. 83–142. DOI: [10.1086/652229](https://doi.org/10.1086/652229).
- Festinger, Leon, Henry W. Riecken, and Stanley Schachter (1956). *When Prophecy Fails*. Minneapolis: University of Minnesota Press.
- Johnson, Luke Timothy (2001). *The First and Second Letters to Timothy: A New Translation with Introduction and Commentary*. Anchor Bible 35A. New Haven, CT: Yale University Press.
- Kunda, Ziva (1990). "The Case for Motivated Reasoning." In: *Psychological Bulletin* 108.3, pp. 480–498. DOI: [10.1037/0033-2909.108.3.480](https://doi.org/10.1037/0033-2909.108.3.480).
- Nickerson, Raymond S. (1998). "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." In: *Review of General Psychology* 2.2, pp. 175–220. DOI: [10.1037/1089-2680.2.2.175](https://doi.org/10.1037/1089-2680.2.2.175).
- Sober, Elliott (2008). *Evidence and Evolution: The Logic Behind the Science*. Cambridge: Cambridge University Press. DOI: [10.1017/CB09780511806285](https://doi.org/10.1017/CB09780511806285).
- Towner, Philip H. (2006). *The Letters to Timothy and Titus*. New International Commentary on the New Testament. Grand Rapids, MI: Eerdmans.