Bayes' Theorem and the Modern Historian:
Proving History Requires Improving Methods

Several examinations of the methodologies employed in the study of Jesus have consistently found those methods invalid or defective. Which fact has resulted in the proliferation of endless different conclusions as to the nature of the historical Jesus and the origins of Christianity. Attention to the logical validity of the methods we employ is essential to repairing this problem. One particular theory of human reasoning can lead the way: widely known as Bayes' Theorem, historians would benefit tremendously from understanding it and learning how to apply it in their arguments and research.

See Also: *Proving History: Bayes's Theorem and the Quest for the Historical Jesus* (Amherst, NY: Prometheus 2012)

By Richard Carrier, Ph.D.
Visiting Lecturer for the Center for Inquiry Institute
Richmond, CA, USA
April 2012

The Problem

Apart from fundamentalist Christians, all experts agree the historical Jesus is buried under some measure of myth and legend. But no attempt to discover the "true" historical Jesus has really succeeded. The latest attempt to cobble together a method for teasing out the truth involved developing a set of criteria and applying them to the evidence. But it has since been demonstrated that all those criteria, as well as the whole method of their employment, are fatally flawed.

Every expert who has specialized in examining these methods and published a study of them has come to this conclusion: Stanley Porter,[2] Dale Allison,[3] Hector Avalos,[4] and almost a dozen others[5]; even Gerd Theissen, who describes the common conclusion, that "there are no reliable criteria for separating authentic from

inauthentic Jesus tradition,"[6] attempted to resist it, but his own effort to rehabilitate the method ultimately led him to agree with that consensus, concluding that the criteria simply don't work, and some other method is needed to replace them, something that involves more general evaluations of plausibility, although he confesses he is not certain how to formally do that.[7] The latest in this series of studies is a new volume to be published this year, edited by Chris Keith and Anthony LeDonne, titled *Jesus, Criteria, and the Demise of Authenticity* (T & T Clark, 2012), and featuring such luminaries as Mark Goodacre and Morna Hooker, all coming to the same conclusion: the method of criteria is simply not logically viable. This leaves the field of Jesus studies with no valid method, and puts into question all consensus positions in the field, insofar as they have all been based, to one extent or another, on these logically invalid methods.

The consequence of this has been more than evident: every scholar using these methods "discovers" a completely different historical Jesus. As Dale Allison concludes, "these criteria have not led to any uniformity of result, or any more uniformity than would have been the case had we never heard of them," hence "the criteria themselves are seriously defective" and "cannot do what is claimed for them."[8] As Helmut Koester concluded after his own survey, "The vast variety of interpretations of the historical Jesus that the current quest has proposed is bewildering."[9] James Charlesworth concurs, concluding that "what had been perceived to be a developing consensus in the 1980s has collapsed into a chaos of opinions."[10] Several others have come to the same conclusion, demonstrating, with extensive citation of examples, the whole confusion of contradictory opinions that has resulted from applying these methods: Thomas Thompson,[11] Thomas Verenna,[12] James Crossley,[13] Mark Strauss,[14] John Poirier,[15] Mark Allen Powell,[16] and John Dominic Crossan,[17] just to name a few.

When everyone picks up the same method, applies it to the same facts, and gets a different result, we can be certain that that method is invalid and should be abandoned. Yet historians in Jesus studies don't abandon the demonstrably failed methods they purport to employ.[18] This has to end. Historians must work together to develop a method that, when applied to the same facts, always gives the same result (or always converges on the same result); a result all historians can agree must be correct (which is to say, the most probable result, as no one imagines certainty is

possible, especially in ancient history). If historians can't agree on what that method should be, then their whole enterprise is in crisis, because agreement on the fundamentals of method is the first essential requirement for any community of experts to deem itself an objective profession.

How Do We Develop a Solution?

There is a solution, and some of the critics of the method of criteria have come close to voicing it without realizing what they were trying to articulate, a method of reaching conclusions that is logically valid and allows productive criticism, debate, and revision that will approach a common conclusion, rather than a chaos of contradictory opinions. It's called Bayes' Theorem, which is a mathematical formula that models all valid empirical reasoning. To understand why it applies, and how to apply it (especially in a field, like ancient history, that seems quite incapable of mathematical analysis), it's necessary first to understand why the method of criteria *doesn't* work.

The method of criteria, which is currently the only method clearly articulated in Jesus studies, suffers three fatal flaws. The first two are failures of individual criteria: either a given criterion is invalidly applied (e.g., the evidence actually fails to fulfill the criterion, contrary to a scholar's assertion or misapprehension), or the criterion itself is invalid (e.g., the criterion depends upon a rule of inference that is inherently fallacious, contrary to a scholar's intuition), or both. To work, a criterion must be correctly applied and its logical validity established. But meeting the latter requirement always produces such restrictions on meeting the former requirement as to make any criterion largely useless in practice, especially in the study of Jesus, where the evidence is very scarce and problematic.

But the third fatal flaw lies in the entire methodology. All criteria-based methods suffer this same defect, which I call the 'Threshold Problem': At what point does meeting any number of criteria warrant the conclusion that some detail is probably historical? Is meeting one enough? Or two? Or three? Do all the criteria carry the same weight? Does every instance of meeting the same criterion carry the same weight? And what do we do when there is evidence both for and against the same conclusion? In other words, even if meeting the criteria validly increases the

likelihood of some detail being true, when does that likelihood increase to the point of being effectively certain, or at least probable? No discussions of these historicity criteria have made any headway in answering this question.

What should be immediately apparent, however, is that solving the Threshold Problem requires admitting the inherently mathematical nature of historical reasoning: some things are more probable than others; some probabilities are insufficient to warrant belief, and only much higher probabilities warrant belief; and some warrant uncertain belief while others warrant more certain belief, on a continuous scale of many different degrees of certainty or uncertainty. There is no way to avoid this. In fact, avoiding it produces a fallacy of logic formally called a false dilemma (but more familiarly known as a "black or white fallacy" or a "fallacy of either/or"), as exemplified by these assumptions: something is either certainly true or it is certainly false; an item of evidence either proves or refutes a conclusion, or it does not; a claim that meets a given criterion is true, a claim that doesn't is false.

These are all logical fallacies, because in reality something is not ether certainly true or certainly false. Because it could be "possibly false but very probably true," it could be "somewhat likely to be true but not certain," it could be "most probably false, but with a small yet tangible probability that it's true," and any number of other degrees of certainty or uncertainty; likewise any given item of evidence will not "prove" or "refute" any theory, but will only increase or decrease its probability, and by varying degrees, depending on how strongly or weakly that evidence fits or contradicts that theory. And if this is true of every premise in any argument toward a conclusion, just imagine how complicated it must be to determine how probable the conclusion actually is, when all the premises each have diverse probabilities of being true. Historians have had a tendency to bury this scandalous fact under a haze of words and assertions that cloak the fact that they don't want to answer the question.[19] Yet their conclusions have no established logical validity if they can't even answer the question of why their conclusion follows from their premises, or with what probability.

This can be illustrated with typical uses of the "criteria" currently employed in Jesus studies. I will present just a few examples. In *Proving History* I more extensively document these and many more examples, covering 31 separate criteria, and citing many other examples in the critical literature.

Using the Criterion of Dissimilarity, which in one formulation states that "if a saying attributed to Jesus is dissimilar to the views of Judaism and to the views of the early church, then it can confidently be ascribed to the historical Jesus," Mark Allan Powell argued that the historical Jesus must have actually addressed God in intimate terms as Abba ("father"), because there is no precedent for this in early Jewish or Christian practice.[20] He was wrong (there is ample precedent, so the criterion did not actually apply).[21] But even if he was right (if we had no precedent for this in early Jewish or Christian practice), the inference is still invalid. Because we simply don't have the kind of evidence surviving from antiquity to make such blanket statements about what *no* Jews or Christians were doing, at any time, much less when the Gospels were composed (whether assuming we are sure when that was, or not).[22]

This is a well-known logical fallacy called an argument from ignorance; indeed, at both levels of analysis: applying the criterion to produce a conclusion of any confidence requires just as much confidence that the practice didn't exist, which (1) is very difficult to establish (as one must thoroughly survey all relevant evidence and scholarship, no simple task, and in this case Powell simply didn't search for or find any of the articles that showed precedents existed, or any of the evidence those articles catalogue, and it might even be unreasonable to have expected him to, which should teach us all caution in making such blanket statements to begin with), and (2) this is in fact often impossible to establish, since we know for a fact there was a great deal more diversity in Jewish beliefs and practice than we presently know any specifics of, and there was a great deal going on in the early Church that we know nothing about, and since the survival of sources is so spotty, no valid conclusion can be reached about what *no* Jews or early Christians *ever* thought, said, or did.

But there is an even more pernicious fallacy here: a simple non sequitur. Even if the principle of dissimilarity could be fully verified as correctly applied (i.e. we could actually prove the item in question is most definitely novel), the conclusion "Jesus did/said that" still does not follow. For example, "if a saying *S* attributed to Jesus is dissimilar to the views of Judaism and to the views of the early church, then Jesus said *S*" assumes an invalid rule of inference: that only Jesus could innovate. But if Jesus could innovate a saying, then so could anyone, including an actual Gospel author (or any other intermediary source on the way to them). Paul, for example,

innovated a Torah-exempt Gentile mission, and if Paul could do that, so could anyone innovate anything else. We know too little about the many Christians and Jews who lived prior to the Gospels to rule any of them out as originators of any seemingly unique saying, yet we would have to rule them all out in order to isolate Jesus as the only available innovator we can credit for the innovation. So any saying *S* we think we can isolate as being unique to Jesus, may in fact be unique to Peter instead (or Paul or anyone else), who uniquely imagined or hallucinated or dreamed or invented Jesus saying *S*. There is no more reason to assume the innovation was of Jesus' own making than of Peter's (or Paul's or anyone else's)—whether consciously, for a specific purpose, or unconsciously, as a construct of what Peter (or whoever) took to be visions or revelations or divine inspirations, but were actually the product of his subconscious mind creatively responding to the problems and ambitions of his time. So how are we to tell the difference? The Criterion of Dissimilarity cannot. Therefore, it is logically invalid. It cannot do what is claimed.

A similar analysis invalidates every other criterion of historicity. The Criterion of Multiple Attestation, for example, runs into the problem of establishing whether we even have independent sources of a tradition to begin with (or whether they are all dependent on each other, e.g. Mark Goodacre's case against the existence of Q is compelling[23]), as well as the problem of determining a valid rule of inference. For even a false claim can be multiply attested in independent sources (e.g. multiple independent sources attest to the labors of Hercules), and folklorists have documented that this can occur very rapidly (there is no relevant limit to how rapidly multiple sources can transform and transmit the same false story). So mere multiple attestation is not enough. This criterion also runs into the Threshold Problem: when do we have enough independent witnesses to believe what they say? Similarly, the Criterion of Embarrassment requires establishing that some detail was in fact embarrassing to the author who records it (this cannot merely be assumed, especially for a sect that was so internally diverse and rooted in open rejection of elite norms) and that this author did not have an overriding reason to include such a detail anyway (such as to convey a lesson or shame his audience into action). This criterion also assumes a rule of inference that is demonstrably invalid, unless somehow plausibly qualified. For instance, the castration of Attis and his priests was widely regarded by the literary elite of the Roman Empire as disgusting and shameful and thus was a definite cause of

embarrassment for the cult, though the claim and the practice continued unabated. Yet no one would now argue that the God Attis must therefore have actually been castrated.[24]

The Lessons We Should Learn

In all these cases there is a common lesson: we must always ask what other reasons there might have been to invent or tell an 'embarrassing' story; or for 'independent' witnesses to repeat a false story; or for something 'innovative' to appear in the record; and so on. This is the principal failure of the method of criteria, and in fact almost all historical methods whatever: simply failing to take alternative explanations of the evidence seriously (or failing to consider them even at all).

But in all of this, the first lesson must be to be sure we have gotten the facts right (as Powell failed to do regarding use of the term Abba for God). That seems like something that needn't be said, but it is a surprisingly common failing among experts in the field, who often tend to repeat things they heard or read or assume to be true, without actually checking, much less diligently enough to be sure of the conclusion. The second lesson is that we need to finally face and understand the role of degrees of certainty and probability, rather than avoiding this complexity in our processes of inference, and instead treating everything as if it were black and white (regardless of what lip service we may give to it being the contrary). Dealing with the realities of probability theory entail a number of counter-intuitive conclusions that nevertheless must govern all sound historical reasoning.

For example, evidence that counts against a theory might not necessarily contradict that theory. Likewise, more evidence will obviously increase a theory's probability of being true, yet that means lacking that evidence must *decrease* that theory's probability of being true, by exactly as much as having that evidence would have increased it (which tells against assuming a well-founded theory must be 100% certain to be true—as that would mean there could not possibly have been, and never possibly will be, any more evidence confirming it, an obvious absurdity). Likewise, genuinely improbable things cannot simply be ruled out *tout court*, since improbable things happen all the time—in fact, it is statistically impossible for them *not* to happen, with considerable regularity. For example, an event that in any one person's

life has a probability of occurring of only one in a million, will have occurred at least a hundred times in the first century (there being a total population over that 100-year period of over 100 million people, which entails on average 100 instances of events with million-to-one odds). These and many other realizations follow from a better understanding of probability theory and correct reasoning about probability. It is therefore essential that historians acquaint themselves with this, as they should all of formal logic, in particular the study of formal and informal fallacies in reasoning.[25]

The third lesson we must learn is the one first mentioned: the need to compare any theory we aim to test or defend *with all plausible alternative explanations of the same evidence*. It is always logically invalid to argue for a conclusion by ignoring all the alternative explanations of the same evidence. And yet, perversely, this is how most historians argue. But how does one argue for a conclusion by taking all these things correctly into account? (All the complexities of probability theory and reasoning about probability, and the need to compare all plausible theories together and not just focus on one.) That is what Bayes' Theorem was developed to do.

I wrote *Proving History: Bayes's Theorem and the Quest for the Historical Jesus* specifically to explain Bayes' Theorem in terms intelligible to historians and humanities majors, who are not typically familiar with mathematics or formal logic (very little of either is actually needed to apply the theorem), and it uses the method of criteria in Jesus studies as a test case, showing why, logically, they are invalid and how we could replace them with logically valid arguments. The unwelcome problem is that when we replace them with valid methods, we discover (as I also show in *Proving History*) that far less can reliably be known about the historical Jesus than scholars of Jesus want there to be. But if that is the logically correct conclusion, then we need to accept that fact, and not try to invent methodological tricks to avoid it.

Bayes' Theorem in a Nutshell

There is an equation. In fact, several, each equivalent to the next but allowing different ways to approach a problem. But I won't frighten you by presenting one here. I provide and explain the relevant formulas in *Proving History* (and all are very simple, requiring no more than middle school math). Here I will just give you an idea of how it works in concept.

Bayesian reasoning consists of asking just three fundamental questions about the evidence and the competing theories explaining it: (1) based on what has typically turned out to be true in the past, what is the initial probability that your explanation of the evidence is true, relative to all other explanations of that same evidence (that is, given past experience, how much more likely is your theory than all those others, or how much less likely; another way to put that is, how typical, or untypical, is your kind of explanation, compared with other explanations available); (2) if your explanation is indeed true, how well does all the evidence we have fit what we would then *expect* to have (that is, how probable is it that the evidence we actually have, is the evidence we would have had); and (3) if another explanation happens to be true (and yours is not), then how well does all the evidence we have fit what we would expect to have (that is, how probable *then* is it that the evidence we actually have, is the evidence we would have had).

The answers to those three questions logically entail a conclusion regarding how probable your explanation is, among all other explanations. And that logical relationship is defined by Bayes' Theorem. This makes Bayes' Theorem a valuable method for testing competing explanations of the evidence, because it models correct human reasoning about the relationship between evidence and explanation. How it works, and how to apply it to that end, I detail extensively in *Proving History*. I also address there all the usual objections to the idea of applying mathematical reasoning to historical inference.

The most common objection is usually to insist that historical reasoning isn't mathematical. But it obviously is. As soon as you start talking about some things being more probable than other things, you are talking about math, even if you aren't aware of doing that. And yet as a historian you do this all the time. For example, saying a theory is "implausible" is just code for "improbable," and in fact not merely improbable, but usually you mean "so improbable that we can dismiss it out of hand," which is not merely improbable, but very, very improbable. This corresponds to a mathematical statement of probability. For example, when you call a claim "implausible," you surely don't mean 50% (1 chance in 2 that the claim is true) or even 10% (1 in 10 chance the claim is true), but something closer to 1% (1 in 100 chance that the claim is true) or indeed much less (since, remember, even claims that objectively have a one in a *million* chance of being true, will have turned out to be

true at least a *hundred times* in the first century). At any rate, you must mean *some* probability by this assertion, otherwise your assertion is meaningless. You cannot avoid the consequences of this. All historical thinking is inherently mathematical, and the more you avoid admitting that, the more easily you will come to illogical or invalid conclusions. It's better to bite the bullet and just learn how to deal with this fact.

The second most common objection is that mathematics implies a false sense of precision. We can never know the "exact" probabilities of anything, especially in ancient history, and using math could even give us conclusions to the tenth decimal place if we wanted, which is absurd when our uncertainties extend above even the first decimal place. But this is a common lay misunderstanding of mathematical reasoning. Although hucksters and propagandists can manipulate numbers and statistics to fool and mislead the public, that does not invalidate correct and honest uses of numbers and statistics. And honestly, mathematics is fully robust enough to model any uncertainty or ambiguity. You simply define your uncertainty mathematically. As long as you recognize that this is being done, there can be no risk of false precision.

For example, I do not know what the probability is that an airplane will crash into my house tomorrow, but I certainly know it is less than 1 in 100. The actual odds are known: insurance companies could tell you, and those odds are vastly smaller than 1 in 100. But I don't need to know that. I already know enough to be more than reasonably certain that whatever that probability is, it is *less* than 1%. Any conclusion I then derive from that premise will retain that uncertainty: any concluding probability *P* will be in fact "*P or less*," and I will be more than reasonably certain that *that* conclusion is true. We can therefore select probabilities, with which to answer those "three fundamental questions" identified earlier, that are certainly higher or lower than we know them to be (just as I know the probability of a plane crashing into my house tomorrow is certainly lower than 1%), and when we do that, we can be certain of the premises, and that entails we must be as certain of the conclusion that follows from them. That is, the conclusion that follows from *those premises*. If we don't like that conclusion or it's too vague to be of use, we can seek to push the limits of our uncertainty, and see what results. That is one of the most useful things about Bayes' Theorem.

There are many other objections people raise, and I address those in the book. I presented these two as but examples, illustrating how these kinds of objections are easily answered. But above all things, the single most valuable reason to learn to think and argue with Bayes' Theorem, is not just the fact that it ensures your conclusions are logically valid given your stated premises (which is the main reason we need to start using it), but that it exposes all your premises to the light of day. In my personal use of this theorem I have found that it allows critics to more readily identify things I overlooked or unreasonably assumed, which allows me to readily correct those mistakes and approach a more valid conclusion. It is because of this prominent feature that historians can use Bayes' Theorem to converge on a common conclusion over the course of any debate, rather than fragment into a chaos of contradictory opinions. In other words, it will not eliminate error. But it will put an expiration date on every error we make. *Proving History* explains how.

I will close with two demonstrations of what I mean.

Example 1

John Meier argues that the "criterion of embarrassment" is often flawed, and as an example he gives Jesus' cry of dereliction on the cross ("My God, my God, why have you forsaken me!?" in Mark 15:34 and Matthew 27:46), which is a quotation of Psalms 22:1. "At first glance," Meier says, "this seems a clear case of embarrassment; the unedifying groan is replaced in Luke by Christ's trustful commendation of his spirit to his Father (Luke 23:46) and in John by a cry of triumph" (John 19:30).[26] But Meier notes that we have to consider all our background knowledge (Jewish and Christian) and compare alternative explanations of the evidence; and when we do that, we can confirm that Luke and John had "later theological agendas" different from Mark's, whereas unlike them Mark shows no apologetic embarrassment at his account at all.

Meier says we can explain Mark's version by hypothesizing that it served a mytho-literary function far outweighing any embarrassment it may have later incurred: namely, the assimilation of Jesus in his death to a venerable Jewish tradition of "the suffering just man." This has established precedents in Jewish literature and therefore would not be an unusual thing to invent, whereas it would be very unusual

for an author to include a report that went against his interests but that he had no need of mentioning (just as Luke and John didn't need to mention it; indeed, they even felt free to change it, even knowing Mark's version was in circulation). Mark, we know, freely invented a lot of things (or in any event, his sources did), and thus we know he was willing and able to fabricate details (or record them). So based on past cases and our general knowledge of how authors behave (especially Jewish authors and Christian authors), Meier's explanation actually starts out more likely. In Bayesian terms, it has a higher prior probability.

When we then see how well it fits the evidence, we find Meier's explanation makes the evidence more likely, too. It would be an inherently improbable coincidence if Jesus' last utterance just "happened" to be an exact quotation from a Psalm that described an executed man in many details matching Mark's story (not impossible, just not *probable*), whereas this is exactly what we'd expect if this was a purposeful literary creation. So the evidence is more likely on Meier's hypothesis than it would be on any other hypothesis. And when we combine this fact with the prior probability, Bayes' Theorem entails the overall probability strongly favor's Meier's conclusion, and not the conclusion that the "criterion of embarrassment" would suggest. In short, Meier's hypothesis fits our background expectations better and makes the evidence more probable.

As Meier explains, "By telling the story of Jesus' passion in the words of these psalms, the narrative presented Jesus as the one who fulfilled the OT pattern of the just man afflicted and put to death by evildoers, but vindicated and raised up by God," which mythic pattern Mark realized by weaving allusions and quotations of the relevant psalms "throughout the Passion Narrative," including this cry. There is a lot more than this that supports Meier's hypothesis against the alternative,[27] and when we consider *all* of that evidence, we will find that the probability of Meier's hypothesis is not just higher than, but greatly exceeds any other. Thus Bayesian reasoning confirms he's right on this one.

Example 2

John Meier gets it wrong, however, when he attempts to use the "criterion of embarrassment" to argue for the historicity of "the baptism of the supposedly superior

and sinless Jesus by his supposed inferior, John the Baptist" (who was proclaiming "a baptism of repentance for the forgiveness of sins").[28] Meier says this must have been embarrassing to Mark, because it contradicts Christian beliefs (that Jesus was "superior" and "sinless"), and because subsequent evangelists scrambled for damage control. But this is the same double error Meier himself refuted in the case of Jesus' cry on the cross.

First, we might see subsequent evangelists were embarrassed by the story, but Mark is not. Had he been, he would already have engaged the same damage control they did. In fact, this would have been done by transmitters of the story decades before it even got to Mark (probably even before Jesus had died). All the evidence of embarrassment is thus only post-Mark. Second, Meier simply assumes Mark (and all prior Christians) believed Jesus was "superior" and "sinless" and thus would not countenance anything implying otherwise. But neither is even plausible, much less established for early Christians, or even those of Mark's time. Paul included Christ's voluntary submission and humbling as fundamental to the Gospel (Philippians 2:5–11). Christians did not imagine Jesus as then "superior" until he was exalted by God *at his resurrection* (e.g., Romans 1:4, 1 Corinthians 15:20–28). There is nothing in Mark's depiction of Jesus submitting to John that conflicts with this view. Only *later* Christians had a problem with it. So when we take our actual background knowledge into account, contrary to Meier, the evidence (Mark's story) is *not* improbable.

Mark instead portrays what Christians originally thought: that Jesus would be exalted as the superior later on. Hence he has John say exactly this (Mark 1:7–8). Likewise, the notion that Jesus was "sinless" from birth is nowhere to be found in Mark or Paul. It is clearly a later development, and thus not a concern of Mark's. To the contrary, Mark has full reason to invent Jesus' baptism by John specifically to create his sinless state, so Jesus can be adopted by God, and then live sinlessly unto death. Mark makes a point of saying John's baptism remits all sins (1:4), that Jesus submits to that baptism, and that God adopts Jesus immediately afterward (1:9–11). This is hardly a coincidence. The role of John and his baptism are explicitly stated by Mark: to prepare the way for the Lord (1:2–3). And that's just what he does. There is no embarrassment here.

So when Meier insists "it is highly unlikely that the Church went out of its way to create the cause of its own embarrassment," we can see in fact such a thing is not unlikely at all: once Christians started amplifying the divinity and sinlessness of Jesus, the story Mark had already popularized started to create a problem for them, so they had to redact that story to suit their changing theology. This proves Mark preceded those redactors and lacked their concerns, but it doesn't prove Mark's story is true. To the contrary, Mark had a clear motive to invent the story, particularly as he needed to cast someone as the predicted Elijah who would precede the messiah and "reconcile father and son" (see Mark 1:6 in light of 2 Kings 1:8; and Mark 9:11–14 in light of Malachi 4:5–6, LXX) and set up Jesus' cleansing for adoption. Why not cast in that role his most revered predecessor, John the Baptist? Having John prepare Jesus by cleansing him of sin and establishing his divine parentage, and then endorsing Jesus as his successor, is actually far too convenient for Mark. Other scholars have made these same observations, and provided many other good reasons Mark would have had to invent this story, and there are additional reasons to doubt its authenticity besides.[29] And when all of that is taken into account, the story of Jesus' baptism by John actually becomes somewhat *improbable*, and thus far from certainly true.

In Bayesian terms, precedents set in Mark and the other Gospel authors (to invent as their purposes required, and to delete or modify or explain anything they found inconvenient to their purpose) make it *more* initially likely that he would invent this story to some purpose than record it because it was true and for some unexplained reason he needed to mention it (even though almost any reason he would have to mention it if true, would be as much a reason to invent it if it wasn't). So the prior probability at least slightly favors invention. And when it comes to explaining the evidence, Meier claims that invention makes the evidence improbable (that it would be "highly unlikely" to find this in Mark if anyone had thought to invent it), but if it were true *and embarrassing* then we would expect it not even to be mentioned in Mark (just as it is never mentioned in John), and thus it's *appearance* in Mark is improbable on Meier's own theory and assumptions, whereas the wholly suspicious convenience of both the claim and the structure of its narration in Mark is far more expected if the claim is invented than if it were true.

Because we can be certain it was not *true* that God publicly acknowledged his adoption of Jesus immediately after the baptism, or that John was sent by God to

reconcile father and son, or that John actually declared Jesus not only his successor but one "mightier than I, the thong of whose sandals I am not worthy to stoop down and untie" and who would bestow on everyone a greater baptism than his. These all served Mark's purposes (or his source's) and yet all of them require inserting a famous Elijah-resembling baptizer into the framework to make those elements persuasively do their work. *Actual* history doesn't work out that conveniently (except but rarely, which means, *improbably*). But literary creations, by their very nature, *always* do.

Thus, the probability of the evidence is much greater on the hypothesis Mark was inventing, than that Mark was forced against his will to include an unnecessary true story that embarrassed him. So even if the prior probability favored neither hypothesis (although in fact it already slightly favors invention), the overall probability still favors invention. But this result might yet be altered by taking other evidence into account. For example, if we had enough *independent* reason to believe that other references linking Jesus to John (in Mark and the other Gospels) were not invented, then we might find that "invention" does not make *all* the evidence as likely as an actual historical connection would, and then we'd conclude the baptism was more probably than not a true event (even if not certainly so). Either way, it is Bayes' Theorem that gives our conclusion logical validity, and exposes our assumptions and inferences to criticism and test.

Conclusion

I more fully discuss these and many more examples in *Proving History*. The same can and should be done for any and every argument in Jesus studies. Not because giving an argument a Bayesian structure will always render unto us an infallibly true conclusion, but because it will ensure that the conclusion logically follows from our premises, and force those premises (and their precise effect on the conclusion) to be made explicit, and thus more easily examined and corrected if in any way defective (by being, for example, either false or incomplete). The method of criteria, by contrast, has nothing at all comparable to recommend it.

[1] Bernard Brandon Scott, *Finding the Historical Jesus: Rules of Evidence* (Santa Rosa, CA: Polebridge, 2008).

[2] Stanley Porter, *The Criteria for Authenticity in Historical-Jesus Research: Previous Discussion and New Proposals* (Sheffield, UK: Sheffield Academic Press, 2000).

[3] Dale Allison, "The Historians' Jesus and the Church," in Beverly Roberts Gaventa and Richard B. Hays, eds., *Seeking the Identity of Jesus: A Pilgrimage* (Grand Rapids, MI: William B. Eerdmans, 2008), pp. 79–95.

[4] Hector Avalos, *The End of Biblical Studies* (Amherst, NY: Prometheus Books, 2007), pp. 203–209.

[5] Richard Carrier, *Proving History: Bayes's Theorem and the Quest for the Historical Jesus* (Amherst, NY: Prometheus 2012), pp. 11-14, with notes and examples pp. 293-96.

[6] Gerd Theissen and Annette Merz, *The Historical Jesus: A Comprehensive Guide*, trans. John Bowden (Minneapolis: Fortress Press, 1996), p. 115.

[7] Gerd Theissen and Dagmar Winter, *The Quest for the Plausible Jesus: The Question of Criteria*, trans. M. Eugene Boring (Louisville, KY: John Knox Press, 2002).

[8] Allison, "The Historians' Jesus," p. 79.

[9] Helmut Koester, "The Historical Jesus and the Historical Situation of the Quest: An Epilogue," in Bruce Chilton and Craig Evans, eds., *Studying the Historical Jesus: Evaluations of the State of Current Research* (New York: E. J. Brill, 1998), pp. 535–45 (quoting p. 544).

[10] James Charlesworth and Petr Pokorný, eds., *Jesus Research: An International Perspective* (Grand Rapids, MI: William B. Eerdmans, 2009), p. 1.

[11] Thomas Verenna and Thomas L. Thompson, eds., '*Is This Not the Carpenter?' The Question of the Historicity of the Figure of Jesus* (Sheffield, UK: Equinox, 2012), pp. 9–10.

[12] Thomas Verenna, *Of Men and Muses: Essays on History, Literature, and Religion* (Raleigh, NC: Lulu.com, 2009), pp. 46–47.

[13] James Crossley, *Jesus in an Age of Terror: Scholarly Projects for a New American Century* (Oakville, CT: Equinox, 2008).

[14] Mark Strauss, *Four Portraits, One Jesus: An Introduction to Jesus and the Gospels* (Grand Rapids, MI: Zondervan, 2007), pp. 358–82.

[15] John Poirier, "Seeing What Is There in Spite of Ourselves: George Tyrrell, John Dominic Crossan, and Robert Frost on Faces in Deep Wells," *Journal for the Study of the Historical Jesus* 4, no. 2 (2006): 127–38

[16] Mark Allen Powell, *Jesus as a Figure in History: How Modern Historians View the Man from Galilee* (Louisville, KY: John Knox, 1998), p. 47.

[17] John Dominic Crossan, *The Historical Jesus: The Life of a Mediterranean Jewish Peasant* (New York: HarperSanFrancisco, 1992), p. xxviii.

[18] For example, Bart Ehrman's latest, *Did Jesus Exist? The Historical Argument for Jesus of Nazareth* (New York: HarperOne, 2012), is still relying on them (cf. pp. 288-93), as if unaware of all the expert literature discrediting them.

[19] This fact was noted and extensively documented by David Hackett Fischer, *Historians' Fallacies: Toward a Logic of Historical Thought* (New York: Harper & Row, 1970).

[20] Mark Allan Powell, "Sources and Criteria," in *Jesus as a Figure in History: How Modern Historians View the Man from Galilee* (Louisville, KY: John Knox, 1998), pp. 47-48.

[21] Mary Rose D'Angelo, "Abba and Father: Imperial Theology in the Contexts of Jesus and the Gospels," in Amy-Jill Levine, Dale C. Allison Jr., and John Dominic Crossan, eds., *The Historical Jesus in Context* (Princeton, NJ: Princeton University Press, 2006), pp. 64–78; and Joachim Jeremias, "Abba as an Address to God," in James Dunn and Scot McKnight, *The Historical Jesus in Recent*

Research (Winona Lake, IN: Eisenbrauns, 2005), pp. 201–206. See also, of course, Romans 8:15 and Galatians 4:6.

[22] See Carrier, *Proving History*, pp. 123-24 and 129-34, for full discussion, evidence of lost information about Judeo-Christian diversity, and citations of other scholars making this same point.

[23] Mark Goodacre, *The Case against Q: Studies in Markan Priority and the Synoptic Problem* (Harrisburg, PA: Trinity, 2002); see also his accompanying website at http://www.markgoodacre.org/Q.

[24] The argument from embarrassment, often regarded as one of the most robust and valuable criteria, is in fact one of the least valid and most unreliable criterion of them all, as I document and explain extensively in Carrier, *Proving History*, pp. 124-69, including citation of other scholars who have pointed out many of the same things.

[25] In addition to Fischer, *Historians' Fallacies* (cited above), I recommend acquiring and reading a convenient digital resource (available on kindle and various other formats): Bo Bennett, *Logically Fallacious: The Ultimate Collection of Over 300 Logical Fallacies* (eBookIt, 2012).

[26] John Meier, *A Marginal Jew: Rethinking the Historical Jesus* (New York: Doubeday 1991), vol. 1, p. 170.

[27] See Carrier, *Proving History*, pp. 131-34 and 126-28.

[28] Meier, *Marginal Jew*, pp. 168-69.

[29] See Carrier, *Proving History*, pp. 145-48 and 134-37.