Tightening the Statistical Resurrection Argument

J. A. Durieux[1]

1 Introduction

When is a report reliable? We define the *credibility* of a report as a measure of how unlikely its appearance would be if the reported fact were false. Lots of false reports reach us every day, so, intuitively, for a report to be believable, the credibility should be high[2]. McGrew & McGrew (2009) evaluates the resurrection report r, establishing a high credibility, leading to the conclusion that r (and thus the resurrection, and thus the gospel message) is reliable.

This article fills two lacunae in the argument given there, and repairs a conceptual error (making the first lacuna irrelevant in the process).

2 Prior probability

A general problem with Bayesian evaluations is the need for a prior probability. We shall call the hartley logit[3] of the prior probability the *prior*. Whereas for repeatable experiments successful repetitions can overcome any finite prior (Dawid 2018), with unrepeatable events any attempt to establish believability can be frustrated if an arbitrarily low prior may be chosen. It is therefore important to establish an objective minimum for the prior for Jesus' resurrection.

2.1 Naturalism

Here it is important to establish what kind of arguments do *not* count against the resurrection. A naturalist may claim that natural science has established some facts to such an extent that anything countering such facts (and both parties agree that Jesus' resurrection would probably counter them) has a negligible prior. That would be wrong, however, in the same way that a mathematical proof that 2+2=4 will not set the prior probability of one marble lacking from a bag in which twice two marbles had been placed to zero: something not covered by the model behind the proof might have happened.

The fat goose waddling towards the farmer on Christmas morning, expecting to be fed, may have developed a whole theory of how the daily rhythm works. This theory, however well established, does not make it more probable that on Christmas morning she will be fed as well.

2.2 Induction

The only thing that influences the rational confidence for the goose of being fed is induction. If her experience started on boxing day the previous year, she'll have reason to assume a probability of 364/365th of being fed – that is: a probability of 364/365 that the usual rules hold –, which is indeed the correct probability.

The fact that, every day, *many* observations show that the rules hold doesn't make Christmas less plausible. The one thing that counts is the thing for which the probability is estimated, in the goose case: feedings. If she were fed three times a day, she would correctly find an 1067/1068th probability of being fed again.

A more fine-grained look would actually bring the prior up, since whereas a resurrection is reasonably unlikely to be missed, we might on a daily basis miss any number of minor deviations from the naturalistic rules, so the numerator would become much less than the denominator.

In fact, most of our experiences don't *confirm* naturalism, but we *assume* it. We don't check whether a passing car may be miraculously floating a millimetre above the road – we assume it doesn't. And even with the myriad unlikely events that happen all the time ("wow, that was fast!") we assume there will be some natural explanation. No doubt there is, but such experiences don't count for our induction case, and in fact weaken it slightly.

1 E-mail: truth@b.biep.org; orcid: 0000-0003-2582-4973; web site: https://biep.org.
2 To be precise, the report should also *not* be unlikely under the alternative hypothesis. We assume that here.
3 McGrew & McGrew uses probabilities, but using logits (log odds) makes the formulas simpler both in form and in content, and the numbers more manageable. For probabilities close to 0 (or 1), the hartley logit is very close to the Napier logarithm of (one minus) the probability, i.e. the exponent of the power of ten.

2.3 The value

So what we need is the probability that the usual (naturalistic) rules hold in the case under investigation. The fact that we can, faithfully repeatably, establish that no Big Bang takes place is no disproof of it having taken place in the past.

The main claim against the resurrection, is *dead people don't rise*. That claim is an inductive one, based on a limited number of observations – taking the whole world population through the ages is certainly an overestimation. That population is thought to be in the order of 10^{11}, so an argument based on that could never bring the prior below -11.

(If we go for the claim *supernatural events don't happen*, the number of observations-with-checks we made will surely be way less than 10^{11}, given that we live less than 10^{10} seconds, and can hardly run or learn about an observation check in less than a second.)

“But wait!”, one might object: the prior for *someone* rising might be that, but in all probability that person would not have been Jesus. Even *given* a resurrection, the prior for it being Jesus would again be -11, bringing the overall prior to -22. But that objection commits the Texas sharpshooter fallacy: the claim is that *someone* rose, and that person happened to be Jesus. Admittedly there is a positive probability that, given a resurrection, it was someone else who rose[4] unreportedly, while Jesus was falsely (yet so convincingly) reported to be risen, but that probability is surely way less than $1 - 10^{-11}$.[5]

More importantly, it would defeat the induction that led to the prior of -11 in the first place: if of the world population through the ages more people may have risen, that population may not be used to establish the strength of the induction. The very possibility that resurrection does not imply a report drastically reduces the base population for the induction to those people for which a resurrection is virtually impossible – i.e. that part of the modern world population whose dead bodies have recently been correctly identified. That number, even squared, will surely be considerably smaller than 10^{11}.

If we restrict ourselves to induction showing that people rising *and convincingly reporting that fact* is low, we are back at our lower bound of 10^{-11}, yielding a prior $p(r)$ of -11. Obviously a subjective prior could be lower, but such a belief would be blind faith.

3 Countertestimonies

Given a prior $p(d)$, the next thing needed is the *credibility* $c(d)$ of a report d, i.e. the negative of the logit under the null hypothesis (i.e. the hypothesis that the report is not founded in truth) of that report reaching us. Establishing the credibility of the resurrection report is the part McGrew & McGrew focus on. Gven independence, overall credibility is the sum of the individual credibilities of the various testimonies.

Now besides the testimonies considered by McGrew & McGrew, there is also the countertestimony of the soldiers of the guard, or at least of the Jewish Council. How are we to evaluate that? Depending on how many soldiers there were, their testimony might provide quite a sizable (presubably negative) credibility.

The credibility that a testimony adds to a report is the **lop**[6] of the testimony if the report were true minus the lop of the testimony if the report were false. So credibility is positive if the testimony is more probable assuming truth of the report, and negative if it is more probable assuming falsehood.

Weirdly enough, taken in isolation, the testimony of the guard provides *positive* credibility for the resurrection. After all, the probability of them telling such a story given a resurrection is not very low (their lives might depend on it if they were Roman soldiers, and their loyalty would make them likely to comply if they were Temple guards)[7]. On the other hand, in the *absence* of a resurrection, their story is at most as likely as the probability of the tomb being empty so soon after the burial – which is quite low. So the probability is larger under truth than under falsehood, and the fact that the soldiers spread this story makes the resurrection more credible.

Now this may feel like cheating: their story makes *the empty tomb* more probable, and the probability of the resurrection, filling part of that probability space, only stretches along as a free rider. The fact that their story explicitly contradicts the resurrection might have an effect *within* that part of the probability

4 Which in itself would refute naturalism, making the truth of Christianity more likely.

5 For one thing, it is lower than the probability of a genuine resurrection report being lost. Depending on the scenario (was it God who resurrected? A statistical quantum fluke?) such being lost may have a low probability itself.

6 The **lop** is the logarithm of the probability. As the logarithm function is monotonously increasing, all qualitative statements true for (non-zero) probabilities are true for lops as well.

7 Given a resurrection, it is likely that Matthew 28:11-15 tells the truth, so we may take his version of what happened as true. If he made up the story, all we have is the fact that much later non-witnesses such as Trypho claimed that the disciples stole the body – and given a resurrection, the existence of such a made-up story would be very likely among people not believing it.

space, counteracting the overall stretching. To see whether that is the case, let us zoom in on the situation *assuming* the empty tomb.

3.1 Assuming the empty tomb

Assuming the empty tomb is not a good strategy *in general* for someone wishing to deny the resurrection. On the one hand, it reduces the induction base for the prior to only those people whose graves were found open and empty shortly after their burial. Of what fraction of them it was established that they hadn't risen? So it greatly increases the prior. On the other hand, it leaves the credibility added by the testimonies investigated by McGrew & McGrew fully intact. But since it diminishes the credibility from the guards' testimony – maybe to negative values – it is worth looking into.

Assuming the empty tomb leaves the numerator of the proportion the same, but increases the denominator: given that the probability of their testimony under a non-empty tomb is virtually zero, the denominator gets divided by the probability of the tomb being empty.

Given an empty tomb and no resurrection, their testimony may be true, or it may be false in various ways.

Now their testimony is unlikely to be true: if they slept, how would they know what had happened? Did they know the disciples well enough to recognise them in the dark? If they woke up, and were able to identify the disciples in the act of lugging an unwieldy body, why couldn't they apprehend them? And if they identified them, then why didn't they apprehend them for grave robbery fifty days later, at Pentecost – and no doubt collect a generous reward from the Jewish leaders?[8]

If their testimony is not true, what *did* happen? Did they see robbers and just assume they must have been disciples? But in that case, isn't it just as probable that they saw no-one at all, but just assumed that the open and empty grave implied a theft? After all, if we come home late at night, to find our front window broken, and our house ravaged and our money missing, we conclude to a break-in with theft – we don't need to see the burglar for that. So in that case their testimony amounts to little more than a confirmation that the grave was empty, and that no legal removal had taken place.

So even assuming an empty tomb and no resurrection, the story is unlikely to be true. But is it probable? Their testimony includes the fact that they slept on duty. It would take some exceptionally honest guards to admit that[9], so without a powerful protector they would be extremely unlikely to tell the story in that shape. And without a resurrection, there is no likely candidate for such a protector[10]. And that part of the story is also *a priori* unlikely: one guard sleeping is quite believable, but all[11]? In other words, even assuming the empty tomb the testimony of the guards would still seem to undergird the resurrection report.

But even if the testimony of the guards were to make the resurrection slightly less credible *given an empty tomb*, it will still raise its credibility overall significantly[12].

3.2 Anything else?

Finally there is the fact that no other countertestimonies exist. The Jewish authorities would have had great interest in anything that would contradict the claims of the Christians, and would have promulgated it – which would likely have been mentioned by some of the early Church fathers, or in Eusebius' Church History (cf. the gospels reporting the existence of false testimonies against Jesus). The Talmud would have been an obvious place to record details about such testimonies – albeit possibly in coded form. Books written against the Christians would have been likely to have been lost, but attempts to rebut their arguments would have had a good chance of being preserved. And given such countertestimonies, the quick spread of the Church would have been way less likely.

8 Here I gave only the arguments from the side of the guards. From the side of the disciples it is unlikely too: they were in no position to execute such a plan, and if they knew they had been seen, would they boldly have proclaimed a resurrection?

9 It would seem more likely for them to invent a story about, say, and angel opening the grave and showing it empty than to admit to something so damning for themselves.

10 The disciples could at a pinch have had a motive, but they definitely would have had the means neither to bribe the guards nor protect them while they disseminated their story. Joseph of Arimathea might have had both, but the hypothesis of a huge fraud by him is extremely implausible in itself.

11 It is not completely impossible if they were Temple guards, and had each drunk their four cups of Pesach wine, but even then their sleeping *and* admitting it seems less likely than their telling that story given a resurrection. And that would have happened in the night of the Sabbath, not the night after it.

12 In order to diminish the overall credibility, it would have to reduce the credibility given an empty tomb by more than it increases the credibility of the empty tomb, i.e. it would have to have a stronger negative effect on the resurrection than its positive effect on the emptiness of the tomb. That is very hard to achieve, given that the probability of them lying about the empty tomb (the checkable very reason for their story) is much lower than the probability of them lying about its (uncheckable) cause, and their interest in having a credible story for the latter.

While the credibility contributed by the absence of other countertestimonies than that of the guard seems hard to quantify, it seems clear that it bolsters the resurrection report, not subtract from it.

4 Credibility and believability

McGrew & McGrew establish a credibility $c(r)$ for the resurrection report of 44[13], which together with the prior of -11 gives an *intrinsic believability* $b(r) \triangleq p(r) + c(r)$ of 33.

5 Significance

It might seem that this high intrinsic believability would make the resurrection very likely. That does not follow, however. The claim that God used a credible report of an *a priori* very unlikely event to prove the truth of the gospel message defines a population R of possible highly intrinsically believable reports[14] $r_i \in R$ with $\forall_i : c(r_i) \geq c(r)$ undergirding the gospel message. Jesus' resurrection is merely an after-the-fact recognition of how God proved His Gospel claim – He could have used any of a large number of such events, and not taking that into account would be the base rate fallacy – after all, R being large, most of those reports would be false, and Bayesian testing *can* give false positives. Given $|R| \gg 0$, $\forall_i : c(r_i) \geq c(r)$ doesn't imply $c(R) \triangleq c(\exists_i : r_i) \gg 0$. What we need is a low probability of the occurrence of *any* false report with a credibility of at least 44 undergirding the gospel message, i.e. a high $c(R)$.

The way we established the prior covers all reports in R that include a resurrection, including e.g. reports of *other women finding the empty grave* or *Jesus appearing to the High Priest*, but also *John the Baptist declaring himself the Son of God, dying and resurrecting* or *Jesus reappearing alive despite His dead body having been eaten by dogs*. In fact, there is quite a bit of variation possible. The Christ might have appeared some generations earlier or later. Instead of the Israelites, God might have chosen an Indian tribe to receive what amounts m.m. to the Old Testament message, and the Saviour might have appeared there, with credible reports of His death and resurrection. Instead of appearing to His disciples after the resurrection, He might have left a message in the sky. As long as these variations include a resurrection, they are covered by our established prior. But what about possibilities other than a resurrection? Presumably, God could have used other ways to prove His point, such as Jesus *not* dying despite clearly being decapitated, or His disciples obtaining the ability to visibly change bread and wine into flesh and blood.

All these potential reports $r_i \in R$ that are at least as credible as the actual one r form the (generalised) tail of the probability distribution, and to claim significance for r it must be shown that $c(R)$ is high. But how to prove such a thing? The world is such a rich and varied place, that computing the total credibility of the generalised tail R seems impossible.

5.1 Bounding

Fortunately, while the tail may be impossible to *compute*, it is still possible to *bound* it. Let Q be the set of reports q_i with $c(q_i) \geq c(r)$ reporting *any miracle whatsoever*. Clearly, $|Q| \gg |R|$ [15], and $c(Q) \ll c(R)$. And $P(\textit{false positive}) \leq P(R) = P(\exists_i : r_i) \ll P(\exists_i : q_i) = P(Q)$. Given the fact that we have no q_i other than r at all, with high probability $P(Q)$ is not high, and $P(R) \ll P(Q)$.

Even if we allow in Q all sufficiently credible reports of not-so-coherent events (e.g. someone prophesying his death in the next week, yet living for several centuries, or levitating), which would again magnify the size by many orders of magnitude, still none has occurred. This gives us a lower bound b of

13 Corresponding to a probability of 10^{-44} that the report would exist if the resurrection hadn't taken place.

14 Actually, the relevant metric would be convincingness, which consists of believability on the one hand, and gospel-provingness on the other. An exceedingly reliable report that someone ate an ice cream would not be convincing in that sense. Here I'll take convincingness to include (but be much more specific than) convincingness that a miracle happened, so that $R \subset Q$ (for the Q to be defined later).

15 For simplicity we assume finiteness of R and of Q. A more sophisticated argument would use series of report sets R_i and Q_i, each member defined as the set of those miracles in R or Q with a complexity (unlikelihood) smaller than i. One could then take the limit for i going to infinity.

$c(R) \geq \log(|Q|/|R|)$. It should be quite feasible to show b to be at least 15 or so[16]. If that can be done, we may conclude that r is reliable.

5.2 Transfer

Another potential source of reliability is found in the fact that not only do false miracle reports of credibility 44 and beyond not exist, neither do such reports of credibility 43, or even lower. Let n be the highest credibility for which a false miracle report exists. Then a resurrection report with credibility n would already have a reliability of b (as defined in the previous section) – say, 15. And that means that the actual report, with credibility 44, has a reliability of b+44-n.

This approach is fully empirical, albeit in very different ways for b and b: whereas b can be found in a way like the one described in footnote 13, n is to be determined through historical research. A clear advantage of that is that there is no need for reliance on a subjective prior credibility.

6 References

McGrew, Timothy & McGrew, Lydia (2009). The Argument from Miracles: A Cumulative Case for the Resurrection of Jesus of Nazareth. In William Lane Craig & J. P. Moreland (eds.), *The Blackwell Companion to Natural Theology*. Blackwell. pp. 593–662.

Dawid, Richard & Hartmann, Stephan (2018). *The No Miracles Argument without the Base Rate Fallacy*. In Synthese 195 (9). pp. 4063-4079.

16 By giving a systematic injective 1:n mapping from r_1 to $q_{i1}, q_{i2}, ..$, e.g.: it could have happened in any other religion (giving maybe 3 credibility points); instead of undergirding it could reject or be irrelevant, and/or make ridiculous, and/or make disrespectful, giving another credibility point; the message preached could have been any of a huge number that is incompatible with the Gospel message; and so on. Basically any change that would clash with the gospel message would work. The mapping would not be mathematically precise, but whereas some combinations may be impossible, others members of Q will not be captured, and it would serve well for a rough estimate.