Correction note as of 8/13/26: I have recently discovered that there were two mistakes (one a mistake in modeling, the other a mistake in calculation) in the worked example 3 in the appendix. I have produced an illustrative worked example of dovetailing and dependence without these mistakes in a blog post at the following link:

https://lydiaswebpage.blogspot.com/2026/08/correction-of-worked-example-of.html

In the worked example given there, the dependence factor for the two testimonies *modulo* H is almost exactly half that of the factor given in Example 3. This is still greater than the dependence factor *modulo* ~H calculated in the previous example (varying but not dovetailing details), and I continue to argue that in the case of dovetailing details, dependence *modulo* ~H is greatly reduced if not eliminated. Obviously much depends on the witnesses' initial credibility as used in the example.

Undesigned Coincidences and Coherence for an Hypothesis

Lydia McGrew

7/18/2018

This is an accepted author manuscript version archived at my personal website after acceptance per author's agreement with *Erkenntnis*.

1. Introduction

Recent literature in the area of Bayesian coherence strongly suggests that the coherence of the contents of reports is not *per se* conducive to confirmation of the contents (Shogenji 2013, pp. 2531-34, 2542-45; Olsson 2005, appendix B, 2013, Bovens & Hartmann 2003, pp. 19-22). A promising alternative suggestion about the value of coherence is that we should focus on the coherence of the *reports* rather than solely on the degree of similarity of their *contents* and should evaluate the coherence of the reports relative to a particular hypothesis (L. McGrew 2016b). In this sense, a set of reports is coherent in a helpful way if it is "coherent for" some salient hypothesis--that is, if the reports hang together in such a way as to confirm that hypothesis.

There are a number of possible ways to parse out the notion that reports (or any set of evidence items) have "coherence for" H. For purposes of this paper, I will take it that a set of reports is coherent for H only if their conjunction constitutes a cumulative case that confirms H, even if each individual report does not *separately* confirm H. To rule out fifth wheels or negatively relevant items of evidence, I will stipulate that the set is coherent for H only if the probability of H given all items in the set is greater than the probability of H given all items except for some individual E_i. In this way, one requires that each item is, in one way or another, strengthening the case for H in the context of the others, while the conjunction as a whole also

confirms H. This is possible even if some item does not confirm H when taken by itself. (See Huemer 2007, pp. 340-342; McGrew 2003)

In what follows I analyze the phenomenon known as "undesigned coincidences" among items of testimony or reports in light of this concept of coherence for an hypothesis. In doing so I will make use of the fact that the contents of the reports are not identical but rather present the features of helpfully varied evidence (Fitelson 2001a, pp. S130-S134, see L. McGrew 2016a), while agreeing in certain respects. These features of difference and agreement are essential to a putative undesigned coincidence, as will be explained by examples.

It was noted as long ago as the 18th century that "[t]he usual character of human testimony is substantial truth under circumstantial variety" (Paley 1859, p. 336), a point that lends itself to fruitful probabilistic exploration.

2. Probabilistic tools

For my purposes in this paper, confirmation will be most conveniently measured by the likelihood ratio,

$$\frac{P(E|H)}{P(E|\sim H)}.$$

Moreover, as long as the prior probability of H is held constant (which is to be assumed in the cases I will be considering), some evidence E' confirms H more than evidence E by the *r* measure

$$\frac{P(H|E)}{P(H)}$$

just in case E' confirms H more than E by the likelihood ratio. This is also true of the difference (*d*) measure

P(H|E) - P(H)

and of any measure that meets the condition

> If P(H|E) > P(H|E') then the confirmation of H by E > the confirmation of H by E' (Fitelson 1998, p. S367).

This is true both for sets of evidence and individual items of evidence. Hence it is possible for those who prefer another measure that meets this adequacy condition to acknowledge the confirmational value of undesigned coincidences as analyzed here, including the results described below.[1]

Other probabilistic tools discussed below (the RC ratio in particular) are very naturally placed within the context of confirmation understood in terms of the likelihood (*l*) measure. While the inequality results described here hold for all of the measures described, *ceteris paribus*, one would lose some of the ease of understanding (by using these tools) the *degree* to which some alteration in type of evidence confers more or less confirmation if one used a different measure of overall confirmation. The consilience ratio for H and a set of evidence (discussed below) has been related to the *r* measure by both Wayne Myrvold (1996, 2003) and Gregory Wheeler (2009, 2012, Wheeler and Schlosshauer 2011), and Wheeler's "focused correlation" measure could in principle be used in conjunction with the *r* measure to model the points discussed here concerning the value of a mix of variety and similarity in evidence.

The individual confirmation offered by some item of evidence to H will in this discussion refer to the individual Bayes factor for some E_n, i.e.,

If E_n confirms H individually, this means that E_n has a top-heavy individual Bayes factor for H on the background evidence, without assuming any of the other items of evidence in the cumulative case under consideration.

The confirmation given to H by a cumulative case from a set of evidence $\{E_1,...,E_n\}$ is the Bayes factor for the conjunction of the items in the set:

Another important probabilistic tool for the following discussion is the consilience ratio that some set of evidence has relative to an hypothesis (T. McGrew 2003). The consilience ratio of $\{E_1,..., E_n\}$ relative to H is

The same ratio can be considered for the same set of items relative to ~H, i.e.

Each of these ratios shows how the joint probability of the evidence, conditional on the hypothesis, compares to the product of the individual conditional probabilities of the items--that is, their joint probability if they were probabilistically independent given the hypothesis. If the probability of the conjunction of the items given the hypothesis is higher than the product of their individual conditional probabilities, then the consilience ratio is top-heavy, and in that respect dependence considerations favor that hypothesis. The consilience ratio is a conditionalized version of Tomoji Shogenji's coherence ratio, as applied to a set of evidence items and an hypothesis. (See Shogenji 2013, pp. 2532-3.)

But that is not all that can be said about dependence. The consilience of the evidence conditional on H may be either higher or lower than the consilience of the evidence conditional on ~H, or they may be the same. (It is also possible to make such a comparison for some H_1 and H_2 that do not form a partition, but for my purposes here I will always compare consilience on H and on ~H.) The consilience ratios for the evidence *vis a vis* H and ~H can be "stacked" to form a ratio of ratios, i.e.

We can use this ratio as a correction factor for dependence among the items of evidence overall (see L. McGrew 2017), as shown below.

When items of evidence are independent given H and given ~H, the cumulative Bayes factor for the items of evidence is equal to the product of their individual Bayes factors,

This product can make for a strong cumulative case with multiple, independent items of evidence, such as witness testimonies (Earman 2000, pp. 55ff, Holder, p. 53, L. McGrew and T. McGrew 2009, p. 618), but the independence consideration is important, and sometimes a correction is needed. For example, if a series of witnesses all testify to the same content, but we have reason to suspect that they are colluding, the correction factor above should be bottom-heavy, showing more consilience on ~H than on H, thus partially undermining the force of the cumulative effect of their joint testimonies (see L. McGrew 2017, pp. 564-65). This explains why evidence against collusion or copying is probabilistically helpful.

As I have shown (L. McGrew 2017, p. 564), the total force of a cumulative case by the *l* measure, represented as

can be correctly decomposed into the product of the individual confirmations provided by the items and the correction factor ratio, like this:

(See also Myrvold 2003: 412–413.) Given this decomposition, it follows immediately that, if the individual confirmations the items offer to H are held constant (or even if their product is held constant), an increase in the final ratio (the correction factor) will increase the force of the case for H as measured by *l*. Moreover, if the prior probability of H is held constant, this result also holds for other measures meeting the condition given above. It also follows immediately from this decomposition that, if the correction factor is greater than 1, then, given the *ceteris paribus* conditions, the case from the conjunction of the E_i is stronger than the case from their independent confirmations alone. This point is directly related to a result that I have proven in L. McGrew 2016a--namely, that if the product of the individual Bayes factors for items of evidence, the prior probability of the hypothesis, and the consilience ratio for the evidence and H are all held constant, then a decrease in the coherence of the evidence on ~H (the denominator of the RC ratio) is helpful to the confirmation of H.

With these probabilistic tools in hand, it will be possible to show why some testimonies or historical reports fit together in a way that is especially confirmatory. In some cases I will argue that the undesigned coincidence increases independence given ~H; in others I will also argue that it increases dependence given H. Both of these are helpful to the RC ratio for H, and it should always be assumed, when I speak comparatively, that I am holding constant the individual confirmations of the items of evidence and the prior probability of the hypothesis.

3. Differences between the approach in this paper and Bayesian coherentism

The project known as Bayesian coherentism commonly involves the attempt to see if a satisfactory measure of the coherence of the contents of testimony can be developed and if an

increase in such a measure among items of testimony is truth-conducive (e.g., Bovens and Hartmann 2003, pp. 11-16). There are a variety of proposed measures of coherence (Olsson 2013), and the idea is usually to find out whether an increase in the coherence of the contents (by some measure), with some independently motivated *ceteris paribus* conditions, increases the probability of the conjunction of all of the contents of the testimonies. Sometimes this project is simplified by having all of the witnesses testify to precisely the same thing (e.g., Olsson 2005, pp. 24-25). Both the reliability of the witnesses and the independence of the witnesses are related only to the contents of the witnesses' own testimonies rather than to the H, if those differ (Bovens and Hartmann 2003, pp. 16-17). As mentioned above, the results of this project so far have been largely negative.

My project here deliberately differs in several respects. At no point in the examples is the H under consideration the conjunction of *all* of the contents of *all* witness testimonies. (For some problems with H as the conjunction of all testimony contents, see L. McGrew 2016b, pp. 329-31, 348-9, and Shogenji 2013, pp. 2526-7.) On the other hand, except where the H is a "source hypothesis" (in section 6.4), the H is always a *portion* of the contents of at least *one* of the reports, and sometimes of both.

These aspects of the modeling mean that the details that are not included in H can float free and participate in dependence and independence considerations as discussed in the examples. It also means that the stipulation (which I always make) that the witnesses or sources have at least some credibility for what they testify makes at least one witness testimony relevant directly to the H (except where the H is a reliability hypothesis), without the need to break out the contents separately. When one source indirectly corroborates the claims of another, the way in which it does so is explained in the exposition.

Collusion and copying and their absence are defined here probabilistically relative to ~H or, in some cases, to the negation of a subhypothesis under H. I do not define the relevant independence between reports as independence given *both* the truth and the falsehood of the contents. That move would rule out useful evidential relations for confirmation, since in several cases the reports are positively relevant to each other given the truth of (at least part of) the contents of one report. As Michael Huemer has noted (2007, p. 341), our concern in ruling out collusion is to avoid probabilistic dependence given the *negation* of the contents. We fear that colluding witnesses would have jointly attested to an event even if it never took place. The problem with copied reports is that the copied version would have agreed with the initial version if the event did not happen.

Unlike other authors in the Bayesian coherence literature, I do not *stipulate* full conditional independence given the contents of testimony, the salient H, or ~H. Rather, I discuss ways in which the varied contents of the reports make it plausible, or at least more plausible than it would be otherwise, that the reports are relevantly independent. Dependence and independence are thus allowed to arise from the nature of the reports rather than being stipulated in advance, making the scenarios more realistic.

Shogenji (2013, p. 2526) stresses that it is "wrong-headed" to try to argue that the coherence of the contents of reports is *per se* truth-conducive. He also stresses problems with treating the hypothesis of interest as the conjunction of all the contents of testimony. I heartily agree with both points and have argued the same elsewhere (L. McGrew 2016b). My project differs somewhat from Shogenji's as well. For one thing, I am attempting to illustrate, in a positive vein, how coherence of the reports *for* an hypothesis of interest can be defined and might work in practice. Shogenji also stipulates in his analysis (p. 2529) that the reports are

independent given their own contents (and their negation), whereas (as explained) I do not stipulate even that much independence. I occasionally stipulate that evidence is not negatively relevant given H, and I argue for other screening conditions based on concrete details.

None of this means that the contents of reports are irrelevant to confirmation. My argument alludes crucially to contents of reports. But the relevance of the contents is far more complex and interesting than would be indicated by trying to measure the coherence of their contents and seeing whether an increase in the coherence of contents is *per se* truth-conducive. The argument here shows in practice how a different approach to issues of overlap and variation in testimony contents can work better than the more common Bayesian coherence approach when we are modeling realistic cumulative cases.

4. Types of undesigned coincidences

The 18th-century Anglican clergyman William Paley coined the term "undesigned coincidence" in his *Horae Paulinae* (Paley 1850, pp. 1-8) in the context of biblical studies. Paley's object was to argue both for the genuineness of the letters traditionally attributed to the Apostle Paul and for the accuracy of the book of Acts. In the 19th century the Anglican minister John James Blunt took up the term and wrote an entire book applying similar considerations to other parts of the Bible (Blunt 1851).

Although the term arose in the context of biblical interpretation and discussions of biblical reliability, the concept of an undesigned coincidence involves general considerations that are relevant in many contexts where one is attempting to discover the truth of some historical matter or the reliability of a witness or document.

In this paper I consider several different kinds of such coincidences, while making no claim to have given a mutually exclusive or jointly exhaustive taxonomy.[2] The categories themselves will help to clarify the meaning of the phrase.

1a) Two different witnesses or sources attest to the same core content while diverging (in a non-contradictory manner) in the specific details they mention.

1b) Beyond this simple variation, the force of coincidences in this category is enhanced further when some differing details fit together in such a way as to confirm each other.

2) One witness or source explicitly attests to an event, while another witness or source mentions some circumstance or detail that confirms the first witness's account without explicitly attesting to the event.

3) Two or more accounts tell about an event or series of events that are not at first obviously the same event but that, upon examination, are best explained as the same.

I will discuss examples of each category and give a probabilistic analysis of their value in section 6.

There are two types of hypotheses that these coincidences support--hypotheses about events in the world and hypotheses about the sources that (supposedly) tell about events. Hypotheses about events are just what they sound like--claims that something actually took place. Hypotheses about the sources are claims such as that a source had access to the facts, that the source is generally reliable, that the source is what it purports to be (e.g., a genuine letter or an account by an eyewitness), and so forth. As the analysis of examples will argue, the types of

accounts discussed can have special force for confirming hypotheses, either about events or about sources, for reasons that are explicable in terms of the probabilistic tools discussed in section 2.

5. Collusion, copying, and empirical considerations

Judgements of report independence are always based upon empirical information and an understanding of how that information affects probabilities. The use of undesigned coincidences to support the relevant independence of reports is no different in this respect from any other non-deductive inference. It is always logically possible that reports are actually copied from one another or in some other way causally synchronized but produced deceptively in order to give the appearance of independence, just as it is possible that two witnesses who were thought to have been isolated have found a way to get together and work out their joint story.

I will show how probabilistic tools can be applied in an interesting way to a kind of evidence that intuitively seems to increase the confirmational force of testimonial reports. The tools yield greater understanding of the probabilistic reasons behind the intuitions and can prevent mistakes by practitioners (see note 8). This project is similar to (and in part an instance of) explaining in probabilistic terms why varied evidence is more valuable to confirmation than unvaried evidence--the subject of a large philosophical literature (e.g., Carnap 1945, p. 93, Thagard 1978: 77–78, 81, Horwich 1982, pp. 119–122, Sober 1989, Howson & Urbach 1993, pp. 113–114, Myrvold 1996, 2003, Fitelson 2001a, pp. S130ff).

It is certainly possible to define variety in a way that permits “more varied” evidence to be less valuable to confirmation than “less varied” evidence, and it is also possible to argue over the merits of specific putative undesigned coincidences. As pointed out in the previous section, it

is a *theorem* that, if the prior probability of H and the product of the Bayes factors of individual items of evidence are held constant, an increase in the relative consilience measure increases the confirmation that the set of evidence offers to H. It is also (*ipso facto*) a theorem that, if those other factors are held equal and the consilience ratio on H is held equal, a decrease in consilience given ~H is helpful to confirmation. I have argued elsewhere (L. McGrew 2016a) that this latter is the correct analysis of cases where it seems intuitively that varied evidence is valuable to confirmation. Undesigned coincidences, as I will argue below, involve either evidence permitting decreased consilience on ~H or increased consilience on H, or both. If one wishes to reject the value of either varied evidence or undesigned coincidences, one would need to argue that there are relevant places where those informal terms should be applied but where the theorems do not apply and that this renders the informally understood properties valueless in those cases for confirmation while still meriting the terms.[3] I am starting with cases where the coincidence seems intuitively valuable and showing that this value can be parsed out probabilistically.

Since a bare intuition about empirical cases is somewhat unsatisfying, I note that there is inductive evidence from those who work with witness testimony that too much similarity is a hallmark of collusion.

> [W]here witnesses deny mutual discussion, the cross-examiner should compare the testimonies of each of them for any significant similarity of content and language which points to collusion and if these emerge they will be fertile openings for challenge. ("Challenging the Credibility of Evidence" 2010)

The variation in undesigned coincidence expresses the other side of the same coin. If, on the basis of experience, we consider great similarity in detail among witness testimonies to look like collusion or copying, then variation of details would normally be *helpful* to confirmation by way of allaying those concerns.

Modern cold-case detective J. Warner Wallace makes similar statements concerning variation and truthful witness testimony, carrying this further to the concept of varied details that also dovetail.

> Often, questions an eyewitness raises at the time of the crime are left unanswered until we locate an additional witness years later. This is a common characteristic of true, reliable eyewitness accounts....It's my job to assemble the complete *picture* of what happened at the scene. No single witness is likely to have seen every detail, so I must piece together the accounts, allowing the observations of one eyewitness to *fill in the gaps* that may exist in the observations of another eyewitness....True, reliable eyewitness accounts are never completely parallel and identical. Instead, they are different pieces of the same puzzle, unintentionally supporting and complementing each other to provide all the details related to what really happened. (Wallace 2013, p. 187, emphasis in original)

The empirical evidence indicates that collusion, copying, or falsification, even by people who are not especially stupid or careless, is by no means *expected* to produce minutely differing, casually stated, and/or interlocking details. On the contrary, such causal dependency is empirically disconfirmed by these coincidences, which is another way to say that they *appear* undesigned.[4]

The empirical point is not merely a black-box induction. It is easy to think of reasons why truthful witnesses would mention varied details. Varied interests might cause them to notice and remember different things. Different physical placement and sensory sensitivities would lead to variation in testimony. Moreover, if what they say is true, their details could be expected to fit together for causal reasons and because reality is internally consistent. E.g. Truthful testimony to a car accident and truthful testimony to glass in the road fit together because an accident might well cause broken windows.

On the other side, it is not hard to see why collusion often brings about suspicious similarity and why it is not expected to bring about these kinds of subtle variety. It is not that a deceptive witness is heedless, *per se*, but rather that the attention and energy of fraudulent colluders has to be spent in other ways, especially in making sure that their testimonies do not contradict each other or the source(s) they are using. Not having the convenience of relying on truth to make their testimonies fit together, they must make up a consistent, false tale and rehearse it carefully so as to avoid scrutiny due to inconsistency.

Moreover, a witness who is falsifying on the basis of some other evidence or collusion cannot be at all sure that those who hear the story will even notice or give credit for subtly varied details. Many such small marks of verisimilitude are, in practice, easily overlooked, especially if they are mentioned casually or if they require combining several points in order to see them. If the details dovetail in some way, but each witness notes only one part of the puzzle and draws no attention to the connection, this careful correlation may well be of no more value for purposes of fraud than merely making the stories agree. If the connection would be noted only if the audience had some rather distant additional information, it could be missed. It is *possible* that fakers will go to the trouble to use such extra artistry, but it is an expenditure of energy that they

may understandably not consider worth the time. It could even backfire if one witness did not get his varied story quite right and created a contradiction instead.

In a given case one could have specific evidence that would counter an undesigned appearance. If one had experience with a witness and knew him to be not only a liar but a liar who took pleasure in putting subtle, artistic touches into his lies for his own amusement, this would rightly influence one's probabilities. In general, however, such liars are apparently not very common.

While an undesigned coincidence may need to be part of a cumulative case for the truth of an account or (still more) the reliability of a source, the value of such coincidences for enhancing confirmation makes them worth understanding better.

6. Probabilistic analysis of undesigned coincidence examples

6.1 Category 1

Suppose that two witnesses, one a bank teller and one a customer, both state that they witnessed a bank robbery. Their testimonies overlap in certain core respects. Both say that the perpetrator was a white male in his twenties, about six feet tall, wearing a black leather jacket. But Witness 1 gives details that Witness 2 does not mention--that the perpetrator was wearing glasses and a red shirt. Witness 2 also has details that Witness 1 does not mention--that the robber was carrying a backpack and limped slightly on the right foot.

None of the varying details contradict one another, though it would not necessarily significantly undermine the force of the testimony even if there were some contradictions (see section 7). They are simply different.

This is the simplest type of coincidence I will discuss. The force of the conjunction of testimonies for the hypothesis that the robbery took place approximately as related by both witnesses (the overlapping content) comes from their individual force. Let us suppose that there is no particular reason to disbelieve the witnesses and that we reasonably take each of them to have some individual credibility for the facts he relates and hence for R, the hypothesis that the bank really was robbed by a person meeting the approximate description given. This can be represented by a top-heavy individual Bayes factor,

$$\frac{P(E_n|R)}{P(E_n|\sim R)}.$$

If the testimonies are relevantly independent, these two factors can be simply multiplied to give us the force of their conjunction in favor of R.

It is at this point that the value of the *differences* in the testimonies becomes relevant. If all goes well--that is, if the testimonies can be regarded as independent given ~R and not negatively dependent given R--we can multiply them together. This wouldn't mean that the joint Bayes factor would be *stronger* than the product of their individual factors, but it would be *at least* as strong as that product. And, as noted above, such a case of multiplied, independent testimonies can mount up fairly quickly. But can we treat them as relevantly independent, or do we need to penalize their conjunction in some way because of the worry that the witnesses might be causally dependent on each other in some invidious way?

The variation in the testimonies' details makes their conjunction stronger as evidence for R than the conjunction would be if these varied details were missing, because the variation can allow us to multiply them without penalty or with a significantly lessened penalty. (The prior probability of collusion is obviously relevant here.) Consider the claim that R is false and that

collusion is taking place between the witnesses. (I will call this a "subhypothesis" of ~R. A subhypothesis entails the main hypothesis but is not entailed by it.) Suspicious similarity between the testimonies would require us to treat the testimonies as significantly dependent on the assumption of ~R, so that

would be significantly greater than 1.

But if the bank teller and the customer were colluding to produce a false tale of a bank heist that never occurred, on ordinary evidence they would not be likely to engage in the ruse of deliberately making their stories agree in some respects while varying in minor, non-contradictory details. Even if they were intelligent colluders, their effort could well be directed towards making sure that their stories are sufficiently detailed to seem credible and that they do not disagree (as discussed above). Arguably, the variation of details is good evidence against collusion. We can think of this formally by saying that the subhypothesis of ~R that hypothesizes collusion is more probable than the further subhypothesis that they are not only colluders but exceptionally artistic colluders. (The examples in Appendix 1 spell this out in greater detail with illustrative numbers.) Since the concern about collusion given ~R (if collusion had appreciable probability) would require treating the testimonies as at least somewhat dependent given ~R, the realization that the conjunction of *these specific testimonies*, given their details, could be explained only by a *less probable* subhypothesis of ~R allows us to be far less probabilistically concerned about dependence given ~R.

The varying details between the account of the teller and the customer alleviate concerns about collusion so as to help us to multiply the force of their testimonies without undue worry that this overestimates the force of the case.

This first example explains the meaning of "undesigned" in the phrase "undesigned coincidence." Roughly, a coincidence between or among accounts is *designed* just in case, if one knew the causal process that is the true explanation of the similarity between or interlocking of the contents, one would be rationally required by that knowledge to treat the accounts as dependent given ~H. If the overlap between testimonies is correctly explained by design, such as collusion, one account's copying the other, or both accounts' copying a common source containing the overlapping information, then the cumulative case is greatly weakened. In an extreme case, the force would be no stronger than that of one testimony alone. If a coincidence is undesigned, this is not the case. When the coincidence appears undesigned (e.g., it appears that both reporters were simply trying to tell the truth and just happened to produce the overlapping content, without using a common source), one does not have to treat them as significantly probabilistically dependent given ~H, unless there is some good separate reason to believe that they are causally dependent in a way that undermines their relevant probabilistic independence.

Now let us complicate matters a little more to understand a more interesting kind of coincidence. Suppose that two putative witnesses both say, with some variation of details, that a generally described event took place--e.g. a picnic. Without including all details of both sources in our definite description of the event, suppose that we have a general idea of "the picnic," designating it by (say) some of the people involved, the approximate location and date, and some topics discussed. P is the hypothesis that the picnic happened.

One witness (but not the other) mentions that the tablecloth at the picnic was orange. The second witness alone mentions that one of the people attending had difficulty getting his car started that day. Thus far, this scenario is the same as the first in this category. Each source testifies to P, and non-contradictory details vary between them. But let's add a feature--namely, the fitting-together of some of the varying details. Suppose that Source 1 says that an English professor with a specialty in James Joyce was at the picnic. Source 2 does not mention the presence of a Joyce scholar but says, "Someone made a joke about some guy named Finnegan," implying that he didn't follow the joke. On background information, we know that James Joyce wrote *Finnegans Wake*. Suppose that Source 1 does not mention *Finnegans Wake* or a joke concerning Finnegan.

The probabilistic analysis here begins like the analysis of category 1. Each source (we stipulate) has some degree of credibility for what he says, and hence his report has a somewhat top-heavy individual Bayes factor in favor of P. Their purely varying details provide some evidence of relevant independence between them (that is, independence given ~P).

The indirect fitting-together of the details adds another helpful element. While providing support for the relevant independence, the details that fit together also create fruitful *dependence* given P. Source 1, remember, attests that the picnic happened along with his unique details. Suppose that the picnic did happen, and add his report, containing the claim that there was a Joyce scholar present. In other words, consider the probabilistic situation given the conjunction $(P \ \& \ E_1)$. Given that P (asserted by Source 1) is true, there is *some* reason to think that Source 1 is truthfully reporting about that day, including the extra detail about the Joyce scholar's presence. There is then some reason to expect to find other information that confirms this detail. E_2 (the report of the joke about Finnegan) is some additional (indirect) confirmation of the

presence of a Joyce scholar. The connection between the mention of the Joyce scholar in E_1 and the detail in E_2 about the joke makes the evidence somewhat *positively* dependent conditional on P.

As pointed out in the examples in Appendix 1, this positive relevance given P works only if there is a relevant kind of independence within P. The support from E1 "flows" to E2 "by way of" the subhypothesis that there was an actual Joyce scholar present at the picnic. If the witness merely conjectured that there was a Joyce scholar present because he heard a vague reference to a joke about Finnegan, then there might have been no Joyce scholar present at all, but he would still have testified that there was. The failure of this screening *within* P would cause the dependence given P to fail. Here it is particularly helpful that the detail is confirmed indirectly rather than directly--again, a matter of variation. It seems highly unlikely that the first witness would jump to the conclusion that a Joyce scholar was present and recount it as established fact given merely the information that someone made a joke about someone named Finnegan. Hence we can argue for the relevant screening condition from the details of the testimony.

What about the consilience of these details on ~P? If both sources said outright that there was a Joyce scholar at the picnic, or if both sources mentioned both his presence and a joke he made about *Finnegans Wake*, then these details would just be part of the overlapping contents of the reports and would be consistent with the falsehood of P and copying or collusion. If one source mentioned *both* the presence of the Joyce scholar and the joke, and the other source mentioned only a portion of the information that did not require any special explanation--the presence of the Joyce scholar--then the similarity might be merely a result of incomplete copying. As it is, the way Source 2 tells the story implies that he simply didn't understand the allusion to Finnegan, and Source 1 might not (for all we can tell) have heard the joke at all, all of

which is quite realistic. While it is *possible* that Source 2 knew about Source 1 and made the comment about the joke in an attempt to pick up on the mention of Joyce, or *vice versa* (that Source 1 knew about the mention of Finnegan in Source 2 and invented a Joyce scholar to explain it) that's highly improbable. Here the coincidence is particularly obscure, depending for its force upon known and remembered information about the writings of Joyce, and it might be easily overlooked by the audience, since neither source draws any attention to the connection.

Based on this argument, we can conclude that the correction factor for dependence, favors P over ~P, and it does so for two reasons--the varied details producing independence in the denominator and the mutually confirming details producing positive dependence in the numerator. What this means is that in this example the case for P given E_1 and E_2 is especially helpful in two ways: It is stronger than it would have been if there were not varied details (because this variation can permit it to be treated as at-least-independent), and it is stronger than it would be if the reports were treated as only independent. Given the extra consilience given P, the case is better-than-independent. The correction factor is not merely neutral (permitting us to multiply the individual factors in favor of P) but top-heavy. (See the appendix for some illustrative numbers.)

6.2 Category 2

Suppose, in this scenario, that Witness 1 tells a story of a flat tire in a small town and a kindly stranger who helped with the flat. Assume that Witness 1 has some individual credibility (an individually helpful Bayes factor) for what he relates.

Witness 1 names the town where the flat tire supposedly took place as Grantville. He says that the stranger who helped him looked to be in his sixties, wore glasses, and seemed by his manner to be well-educated. The stranger, according to Witness 1, said that his first name was Marcus.

Later, while browsing the Internet, you happen upon a story from a Grantville news outlet about a recent "thank you" ceremony for a well-loved, retiring high school English teacher named Marcus van Dyke. Van Dyke wears glasses and, says the news story, is sixty-seven years old.

Let us make the salient hypothesis something general rather than including every detail of the testimony. Let the hypothesis be T, the proposition that Witness 1 had a flat tire and that he was helped by a kindly stranger. In that case, the news account about the retirement of Marcus van Dyke has (plausibly) no confirmational value all by itself (or none worth mentioning) for T. The mere existence of a well-loved high school teacher named Marcus van Dyke in Grantville would give us no significant reason *by itself* to think that some particular person had a flat tire and received help from someone like him. Yet the news story seems to have value for confirming the witness's story of the flat tire. How can this appearance be explained?

The probabilistic value of the news story lies in the dependence between it and the witness's testimony, given T, and in their plausible independence given ~T. Suppose that T is true: The witness did have a flat tire and was helped by a kindly stranger. By itself, since it is so non-specific, T still gives us no particular reason to expect to find the news story about Marcus

van Dyke. But if T is true and we also have E_1--the witness's report with its further details--then the witness is (*ex hypothesi*) telling the truth as far as the broad outlines of his story. As in the second scenario in category 1, the conjunction (T & E_1) produces a case for what one might call Witness 1's "local reliability"--that is, his truth-telling in the context of this story. In that case, there is some reason to believe that the details of his story are true as well--that the town where his flat tire took place was Grantville, that the kindly stranger who helped him had the slightly unusual name of Marcus, wore glasses, seemed well-educated, and appeared to be in his sixties. If the witness is telling the truth about the story as a whole, there is no particular reason to believe that he has made up the details out of whole cloth, even though it is possible that he is misremembering something. Therefore, given the conjunction of T and E_1, there is some reason to expect *some* E_2 to emerge that would confirm the truth of a detail included in E_1, and the news report about the retirement party is just such an E_2--that is, evidence for the existence of a person such as E_1 describes. Hence, $P(E_2|E_1 \& T) > P(E_2|T)$. In other words, E_1 and E_2 are positively relevant to each other given T. This means that the consilience ratio

is top-heavy.

Next we have to consider whether the consilience ratio of the two reports *modulo* ~T is top-heavy or not, and whether it is as top-heavy as that for T. Whether or not this story is dated after the supposed flat tire incident is certainly relevant to whether or not Witness 1 *could* have invented his kindly stranger based upon the news item. Obviously, if the news story did not even

exist at the time that Witness 1 told his story, that particular source *could not* have given rise to the "Marcus" in the flat tire account.

But even if the news story predated the flat tire story and thus in theory could have been a source for a false tale by Witness 1, there is no good reason to believe that this is the case. For one thing, Witness 1 could not have counted on the fact that someone would happen to run across another account confirming the existence of Marcus van Dyke answering to his description, and he made no such connection. He did not even give his helper a last name. The casual statement of the details in Witness 1's story and their happening to be confirmed by the news story gives an appearance of *undesignedness* to the coincidence between the two accounts in this scenario.

Suppose that T is false: W1 has *not* recently had a flat tire and help from a kindly stranger. How likely is it, in that case, that the details of the account (such as the existence of a man answering to the description of the Good Samaritan) will be corroborated? If W1 made up the details entirely, it would be sheer coincidence that some of them would have separate, confirming evidence. If, on the other hand, he gave those details because he himself had reason to believe that such a person existed and was trying to give verisimilitude to his falsehood, then this would render the items dependent given ~T. But in that case it is rather surprising that he left it to chance as to whether anyone would give him credit for his artistic genius. If the conditional probability of the relevant sort of dependence on ~T (e.g., T is false but Witness 1 put in true details because he knew about them otherwise) is very, very low, it does little to make the items of evidence dependent given ~T as a whole.

If the items of evidence are independent or virtually independent conditional on ~T, then the consilience ratio

is neutral or virtually neutral.

This, of course, is where empirical considerations come into play. Suppose that you have other reason to think that W1 is an artistic liar who doesn't mind too much if his artistry goes unpraised. Another potentially relevant subhypothesis of ~T is that W1 has appropriated the story of a person who really was helped out along the road by Marcus van Dyke. This would explain the appearance of an undesigned coincidence by the fact that all of this really happened, only not to W1. If the witness has been previously caught doing such a thing, that information would raise the consilience ratio given ~T. However, if W1 has (as stipulated) some individual credibility for his story, presumably this is not something we have good reason to believe. The usefulness of the formalism, considered qualitatively, is that it requires us to think about subhypotheses of ~T and the evidence concerning their probability. This value of the formal analysis does not require the use of specific numbers.

If the items of evidence are significantly more dependent conditional on T than on ~T, then the correction factor

is favorable to T, and the case for T from the conjunction of E_1 and E_2 is thus stronger than the case from E_1 alone, even though E_2 does not have any force for T all by itself.

6.3 Category 3

Suppose that a history book with some positive reliability declares that Abraham Lincoln gave a speech in September, 1859, in Milwaukee, Wisconsin, at the Wisconsin State Fair. An historian investigating Lincoln's life finds, to his frustration, that he is unable to get direct confirmation in other historical sources for this alleged speech. According to the book that describes the speech, Lincoln traveled afterward directly to Ohio.

Seeking more information about this period of Lincoln's life, the historian comes upon a copy of a letter, ostensibly by Lincoln and appearing to bear his signature, handwriting, and style. The letter, unfortunately, has had its date torn off. In it, Lincoln (apparently) writes to a friend and tells the friend that he enjoyed his visit with him "this past fall" in Milwaukee, where (the letter indicates) they discussed slavery in the territories and the question of Kansas's statehood. The letter-writer also implies that he headed to Ohio immediately after that visit. The letter-writer does not mention that he gave a speech.

Since no date survives on the letter, and since the history book does not mention any visit to the friend named in the letter, it is not entirely obvious that the letter and the history book are describing the *same* visit by Lincoln to Wisconsin. But that certainly seems like a reasonable conclusion. In both cases, the alleged visit occurred in the fall. In both cases, Lincoln supposedly went to Ohio immediately after being in Wisconsin. And, on background evidence, the events in Kansas alluded to in the letter are the sort of things that might well have been on Lincoln's mind in 1859.

The historian's interest lies most of all in the question of whether Lincoln gave the speech. Let S be the proposition that Lincoln did give a speech at the state fair in Milwaukee in the fall of 1859. Since both the history book and the letter relate the further detail that Lincoln

went on to Ohio after being in Wisconsin, there is dependence of the kind I have already discussed in the earlier categories. That is, given that S is true, the history book is to be trusted that far, and hence the further details given, such as Lincoln's further travels, may end up being independently confirmed elsewhere. The additional detail about Ohio (not built into S) *is* apparently independently confirmed in the letter. So the letter and the history book report seem to be somewhat positively relevant to each other given S.

As in the picnic case, the positive dependence given S requires that the evidence of the history book and the letter should be screened off from one another by a subhypothesis of S--namely, that not only is S true but that the further detail about Lincoln's travel to Ohio is also true. If the history book writer got the detail about Ohio (innocently) from his own access to the letter, then he might attest to this detail even if it were not true--e.g., even if Lincoln wrote mistakenly or confusingly about going immediately to Ohio. This sort of dependence given the negation of the relevant subhypothesis of S would destroy the positive dependence between the two items of evidence given S. I'll address below the issue of whether or not the author of the history book was dependent on the letter.

If S is false altogether, then, unless the history book is based upon the letter, it is quite surprising that the letter appears accidentally to confirm a detail connected with a report of a speech that never happened.

Another interesting point that leaps to one's notice in this case is the multiple confirmation of the fall visit to Wisconsin. The letter does not directly confirm the speech, nor does it contain the precise year, but it confirms that Lincoln visited Milwaukee in an autumn *around* that time and hence confirms to some extent Lincoln's visiting that city in that particular fall. Thus it confirms S, since Lincoln could not have given a speech in Milwaukee in

September, 1859, if he never went to that city in the autumn during that period of his life at all. Part of the content of S is thus somewhat confirmed by the letter. Given both the letter and the history report, together with our other knowledge of the topics Lincoln was interested in at the time in question, it is difficult to doubt that the visit took place and that the letter and history book are referring to the same visit.

By this analysis, and if the history book did not get the Ohio detail from the letter,

is greater than 1, and both

$$\frac{P(E_1|S)}{P(E_1|\sim S)}$$ and

$$\frac{P(E_2|S)}{P(E_2|\sim S)}$$

are greater than 1, though the confirmation of S by the letter is only partial (confirming a fall visit). All of these are good things, *ceteris paribus*, for the confirmation of S.

What about the consilience of the evidence given ~S? Is

top-heavy? And what about the possibility of dependence between the history book and the letter? This is where, as in the other examples, the differences rather than the similarities of the reports are important. It is possible that the history was based upon the letter, at least in part. Let us assume that the historian cannot show in any *decisive* way that the author of the history did not have access to the letter. History writers quite legitimately use letters as sources. If the writer of the history book merely got the agreeing and mutually confirming details from the letter, it is possible that the story of the speech is false and that the letter does not support it in an independent manner. If that were the case, then the reports would be causally dependent given ~S, and if we have reason to think that is the case, they should not be treated as probabilistically independent given ~S. Moreover, if the Ohio detail came from the letter in the first place, then that detail does not create the relevant kind of dependence given S.

But if that were the way in which the history came to be written, it is surprising that the only overlapping facts in the book are the general time of year, the city Lincoln visited, and the following trip to Ohio. The history writer, I will stipulate, does not say that Lincoln visited the friend named in the letter, nor does he mention a visit to *anyone* with whom Lincoln discussed the topic of Kansas's statehood. The history writer does not put any allusion to that topic into the paraphrase of the alleged speech, only an indirect allusion to slavery, and no reference to the expansion of slavery into the territories. In other words, if the history writer based his account on the letter and otherwise (say) made up his account of Lincoln's speech, he seems to have been surprisingly coy in his use of the information in the letter.

If the history book writer innocently based the Ohio detail on the letter (while not making up the speech), he seems to have been surprisingly uninterested in Lincoln's other doings in the city. The hypothesis of innocent use of the Ohio detail from the letter would be further

disconfirmed if there were other ways in which the history book account either failed to refer to or went beyond the information in the letter, thus making it unlikely that the history book author was trying to report accurately while basing the Ohio detail on the letter without independent access to that fact.

This, again, is where the *apparent* undesignedness of the coincidences comes into play. While one cannot be *certain* that the history writer didn't use the letter to include the similar information, and while much will depend upon prior information about that issue, the differences between the two accounts and the casualness of their intersection give us important evidence that supports treating them as more independent on the assumption of ~S than on the assumption of S. If this is empirically defensible, the correction factor comparing the two consiliences favors S. In other words,

is somewhat top-heavy because its numerator is top-heavy.

At the same time, the individual Bayes factors by which the two reports confirm the portion of S concerning Lincoln's presence in Milwaukee in the fall of 1859 can be multiplied without worry that this significantly *overestimates* the strength of the case for that part of the content.

6.4 Hypotheses about the reliability of the sources

In the Lincoln case, the realization that two different sources, with an appearance of independence between them, both appear to allude to the same trip by Lincoln should give the historian greater confidence in the general veracity of the history book and greater confidence that the letter is by Lincoln. An investigator may be interested in such broad questions about sources, whether written or spoken. Was the writer of this memoir knowledgeable and truthful? Did this account come from eyewitnesses? Is this witness of the bank robbery truthful and a good observer?

Rather than going through each of the above types of undesigned coincidences again, I will sketch a form of argument from undesigned coincidences for the reliability of *two or more* putative sources. "Reliability" need not be defined here with any great precision, though it is most useful to take it to describe a lower bound on how often a source or witness gets something right rather than a point value. Reliability involves such notions as that a reporter is attempting to be literally truthful and has access to the relevant facts--circumstances that would normally be expected to yield a reasonably high degree of accuracy.[5]

While one can focus on the reliability of just one of the sources, the easiest way to think of this argument is to consider the conjunction of reliability claims for both (or all, if more than two) of the sources that participate in the undesigned coincidence.[6] In other words, the hypothesis about sources would be something like,

> Both of the people who claim to have witnessed a bank robbery are generally truthful and reliable.

When this type of conjunctive hypothesis is in question, it is noteworthy that the reports have no *individual* confirmatory force for the hypothesis. While we may already give some credence to what Source 1 says, the mere fact that Source 1 says that Lincoln went to a particular city and gave a certain speech does nothing *all by itself* to confirm further the hypothesis that *both* Source 1 and Source 2 are generally truthful and reliable. If, then, the undesigned coincidence confirms such an hypothesis about the multiple sources, it must do so entirely by way of the consilience ratios for the conjunction of the reports.

This is similar to the case of gravitational lensing discussed by Timothy McGrew (2003, pp. 563-64). As McGrew explains, scientists discovered in 1979 that "two quasar images only five arcseconds apart...[had] identical spectral characteristics." The explanatory hypothesis to which this drew attention was lensing:

> By far the most attractive hypothesis proposed was that the phenomenon consisted of a double image produced when radiation streaming from a single quasar was bent by the gravitational field of some massive object located between us and the quasar—a gravitational lens.

Neither of the two quasar images *by itself* confirmed the hypothesis of lensing. From the fact that *one* image had a particular set of characteristics, there was no reason to think that there was a massive object that was bending radiation to produce two images of the same object only a few arcseconds apart. It was, rather, the fact that both images had the same characteristics that confirmed this hypothesis, by way of consilience. If L (the lensing hypothesis) were true, then the spectral characteristics of one image would be expected to be found again in another image:

The two image reports were dependent given L. If L were false, the two images would be independent, and it would be an astonishing coincidence to find that they had identical spectral characteristics. Hence, the ratio of the consilience ratios (which I have called a correction factor) was top-heavy in favor of L, which produced the *entire case* for L from the similarities between the images.

The case is similar for a conjunctive reliability hypothesis for two or more sources. While no individual report confirms the reliability of the sources, the agreement on some point(s) between or among reports does confirm, and therefore so does their conjunction.

Let H be a conjunctive reliability hypothesis for two or more sources that participate in an undesigned coincidence. Consider the consilience of the reports if H is true. It is characteristic of an undesigned coincidence (represented in all the types I have discussed) that either the central assertions of the reports or one or more of their incidental details are confirmed, directly or indirectly, by all the sources that participate in the coincidence. This does not contradict the point I have made throughout that the sources may not all individually confirm the hypothesis about real-world events, as defined in each category. E.g., in category 2, only one source individually confirms *the flat tire story itself*. But there is always some sort of overlap or dovetailing between the report contents at *some level*, just as in the gravitational lensing case both reports showed the existence of a quasar having a particular set of spectral characteristics.

If an H that asserts the reliability of both sources is *true*, then, if we also have assertions or factual implications in one source (that is, we have one report in hand), we have some reason to think that another reliable source discussing the same or closely related topics may confirm some of them in some fashion. In other words, given a reliability hypothesis that mentions both sources, and given one report in all its detail, the other report is more to be expected than given

the reliability hypothesis and neither report. This does require (as discussed in the previous section) that the negation of the relevant subhypothesis of reliability--namely, the subhypothesis about the statements in the first source--screens off the testimony of the first source from the second. In other words, we need to be able to argue that the first source wasn't (even if he is reliable) simply saying what he did because he knew about what the other source said. Suppose that the conjunctive reliability hypothesis is false. This means that one or both of the sources is not reliable, either because that person or source is not trying to tell the truth, or not trying very hard, or makes mistakes, or for some other reason. In the case of gravitational lensing, if lensing were false and the two images were not both coming from the same quasar, there was no reason to expect them to have the same spectral characteristics. The specifics of the images would be independent, which made the conjunction of the evidence in which they were identical enormously improbable. Similarly, if it is not the case that both of the sources discussed in a coincidence are reliable, and if the sources can be regarded as probabilistically independent, it is highly improbable that they would appear to confirm one another, especially on small, indirect, or difficult details.[7]

As when the hypothesis concerns events, so here: The question of independence brings pressingly into focus the possibility of copying or other causal dependence between or among the reports. If one reporter is *not* generally reliable himself but in this particular instance has copied from a more reliable reporter, then the reliability hypothesis concerning both sources might be false, while we would still find the apparent coincidence between some contents. Even if both reporters are reliable, if one has innocently copied from another, this destroys the relevant dependence given reliability--it provides no additional evidence for that reliability. If we knew for certain that the second source simply copied his report from the first and had no other access

to the events, we would (rightly) not consider the coincidences between them to provide any confirmation at all to their reliability. Since quasars are not sentient beings and cannot copy, this type of issue does not arise in the lensing case, provided that scientists can take as given that the different quasar images did really have those spectral characteristics (i.e., that those describing the images were reporting truly). Something similar is true if both sources got their information from a common source rather than having independent access to the facts.

This, again, is where variation, casualness, and indirectness constitute non-deductive but important evidence against problematic kinds of causal dependence. In the first category 1 example, details are found in each report that are not found in the other. In the second category 1 example, one source explicitly states only that a Joyce scholar was present at the picnic, and the other says only that someone made a joke about "some guy named Finnegan." These fit together only if one possesses relevant background knowledge, which is not given in either source. In the category 2 example, the two reports are not even directly about the same incident (the alleged flat tire); one report, from a newspaper that the alleged witness might not even have read and that his intended audience might never encounter, confirms the existence of a person who fits well with the description given by the alleged witness to the flat tire incident. Moreover, the alleged witness did not give the Good Samaritan's full name. In the category 3 example, one source reports a speech by Lincoln but does not mention a visit to a friend in Milwaukee while the other reports his visit to a friend but does not mention the speech. These divergences give a strong, though fallible, appearance of relevant independence to the accounts.

The apparent undesignedness of the coincidences operates, in arguments about the events, to support treating the reports as probabilistically independent given the negation of the hypothesis of interest. It operates in the same way when the hypothesis of interest concerns the

reliability of the sources. If the sources are not reliable, and if they do not appear to be copied from one another, a result of collusion, or copied from a common source containing the overlapping information, then their contents seem (as far as we can tell) to have their features independently, and their agreeing where they do is highly improbable. If the sources can be treated as independent given ~H (that is, given that they are not both reliable), the consilience ratio for ~H,

is neutral. But as already argued, if they are both reliable, and if one did not innocently copy from the other, the consilience ratio for H,

is top-heavy. Hence, as in the case of gravitational lensing, the overall correction factor,

favors H. The extent to which it favors H will depend both upon the degree of dependence given H and upon the improbability of the independent conjunction given ~H (see note 7). The "correction factor" is the whole of the case from the conjunction of the reports.

7. Conclusion

For each category of undesigned coincidence, I have argued that the coincidence conduces to the confirmation of a salient hypothesis, whether this is an hypothesis about an event or about the reliability of the sources. A successful undesigned coincidence makes evidence *coherent for* an hypothesis, as "coherence for" was defined in section 1, producing a cumulative case for H from the conjunction of the evidence. This sometimes arises from the combination of individually confirmatory items, where the undesigned coincidence makes it legitimate to multiply these individual confirmations without overestimating (category 1a). In more complex cases, the special force of the undesigned coincidence arises from the greater consilience of the items given H than given ~H and sometimes also from individual confirmation from both items.

A fascinating area for further research concerns outright contradiction or apparent contradiction between sources. On the one hand, apparent contradictions provide good arguments for independence. If two witnesses are carefully colluding, it is likely that they will try *not* to contradict each other. If two accounts are based on a common source and have no other access to the events, the derivative reporters will (probably) try not to contradict outright their only source. Hence, contradiction or apparent contradiction usually provides some reason to reject problematic types of causal dependence.

On the other hand, there is a reason why apparent contradictions between accounts are considered epistemically important. We usually assume that, if two witnesses both truthfully attest to the same general event, their details will be accurate as well and hence will not contradict one another. My argument has repeatedly made use of this assumption when I argue for dependence of reports in their details given H. Formally, there is a question here that requires further reflection: Since there is some positive dependence given H between reports that confirm the same ancillary details, does it follow that apparent outright contradiction on matters of detail produces *negative* dependence given H?

Empirically, calibration will be useful in evaluating apparent contradictions. How often do we find that an event is independently shown to have happened, that witnesses accurately recounted the core of the event, but that witnesses made mistakes on other details? Can items attested to be divided into different empirical reference classes--types of facts that witnesses who are trying to be truthful are more likely to get right and those that they are more likely to make mistakes about? On the basis of such empirical information, the extent to which minor contradictions undermine the force of a case from multiple witnessing may be mitigated, while the real or apparent contradictions will continue to support the relevant independence of the witnesses. Since none of the examples discussed here involve contradictory details, the discussion in this paper is unaffected by such results, but they are relevant to further consideration of complex historical texts.

A valuable role of the formal epistemologist is to delineate the nature of an apparently confirmatory argument and to explain the probabilistic rationale for its *prima facie* force. This also allows us to explain clearly why some cases seem stronger or weaker than others and to focus on the probabilistically relevant points when evaluating an argument.[8] My goal has been to

show how undesigned coincidences can create coherence for an hypothesis. In these cases, undesigned coincidences point both to the truth of what the sources attest and to their reliability.

Works Cited

Blunt, John James (1851). *Undesigned Coincidences in the Writings Both of the Old and New Testament*. New York: Robert Carter and Brothers.

Bovens, Luc & Hartmann, Stephan (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.

Carnap, Rudolph (1945). “On Inductive Logic,” *Philosophy of Science* 12:72-97.

Claveau, François and Grenier, Olivier (2017). "The Variety-of-Evidence Thesis: A Bayesian Analysis of Its Surprising Failures," *Synthese* (on-line) DOI 10.1007/s11229-017-1607-5.

“Challenging the Credibility of Evidence in Criminal Trials” (2010), Criminal Law and Justice Weekly, viewed 8/24/17 <https://www.criminallawandjustice.co.uk/features/Challenging-Credibility-Evidence-Criminal-Trials-0>.

Earman, John 2000. *Hume’s Abject Failure: The Argument Against Miracles.* New York: Oxford University Press.

Fitelson, Branden (1998) “The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity,” *Philosophy of Science* 66, Supplement, “Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association,” Part I, pp. S362-S378.

_____ (2001a). “A Bayesian Account of Independent Evidence With Applications.” *Philosophy of Science*, *68*(3), Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association, S123–S140.

_____ (2001b). “Studies in Bayesian Confirmation Theory,” PhD thesis, University of Wisconsin, Madison, WI.

Huemer, Michael (2007). “Weak Bayesian Coherentism.” *Synthese* 157:337-346.

Howson, Colin and Urbach, Peter (1993). *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court.

Holder, Rodney (1998). “Hume On Miracles: Bayesian Interpretation, Multiple Testimony, and the Existence of God,” *British Journal for the Philosophy of Science* 49:49-65.

Licona, Michael (2016). "Licona Responds to Ehrman on New Testament Reliability," < https://thebestschools.org/special/ehrman-licona-dialogue-reliability-new-testament/licona-detailed-response/> (accessed 1/23/18).

McGrew, Lydia (2016a). "Evidential Diversity and the Negation of H: A Probabilistic Account of the Value of Varied Evidence," *Ergo* 3:10.

_____(2016b). "Bayes Factors All the Way: Toward a New View of Coherence and Truth," *Theoria* 82:329-350.

_____(2017). "Accounting for Dependence: Relative Consilience as a Correction Factor in Cumulative Case Arguments," *Australasian Journal of Philosophy* 95:3:560-572.

McGrew, Lydia and McGrew, Timothy (2009). "The Argument from Miracles: A Cumulative Case for the Resurrection of Jesus of Nazareth." In W.L. Craig and J.P. Moreland, eds., *The Blackwell Companion to Natural Theology*. Oxford: Wiley-Blackwell, 593-662.

McGrew, Timothy (2003). "Confirmation, Heuristics, and Explanatory Reasoning," *British Journal for the Philosophy of Science* 54:553-67.

Myrvold, Wayne (1996). "Bayesianism and Diverse Evidence," *Philosophy of Science* 63:661-5.

_____ (2003). "A Bayesian Account of the Virtue of Unification," *Philosophy of Science* 70:399-423.

Olsson, Erik (2005). *Against Coherence*. Clarendon Press, Oxford.

_____. (2013) “Coherentist Theories of Epistemic Justification.” *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), Viewed 8/24/17, <http://plato.stanford.edu/archives/spr2013/entries/justep-coherence/>.

Olsson, Erik and Schubert, Stefan (2007). “Reliability Conducive Measures of Confirmation.” *Synthese* 157:297-308.

Paley, William (1850). *Horae Paulinae, or The Truth of the Scripture History of St. Paul Evinced by a Comparison of the Epistles Which Bear His Name with the Acts of the Apostles and With One Another*. Edited by T.R. Birks. London: The Religious Tract Society.

_____. (1859) *A View of the Evidences of Christianity: In Three Parts.* Edited by Richard Whately. 1859 edition. Murfreesboro, TN: Dehoff Publications, 1952.

Schubert, Stefan (2012) “Is Coherence Conducive to Reliability?” *Synthese* 187:607-21.

Shogenji, Tomoji (2013). “Coherence of the Contents and the Transmission of Probabilistic Support.” *Synthese*, 190:2525-2545.

Sober, Elliott (1989). “Independent Evidence About a Common Cause.” Philosophy of Science 56:275-87.

Thagard, Paul (1978). “The Best Explanation: Criteria for Theory Choice.” *The Journal of Philosophy* 75:76-92.

Wallace, J. Warner (2013). *Cold-Case Christianity: A Homicide Detective Investigates the Claims of the Gospels*. Colorado Springs, CO: David C. Cook.

Wheeler, Gregory (2009). "Focused Correlation and Confirmation," *British Journal for the Philosophy of Science* 60:79-100.

_____ (2012). "Explaining the Limits of Olsson's Impossibility Result," *The Southern Journal of Philosophy* 50:136-150.

Wheeler, Gregory and Schlosshauer, Maximilian (2011). "Focused Correlation, Confirmation, and the Jigsaw Puzzle of Variable Evidence," *Philosophy of Science* 78:376-92.

Appendix 1--Illustrative worked examples

Example 1: Entirely overlapping testimony

Here we have two testimonies E_1' and E_2' that have entirely overlapping content.

Suppose that

and that P(H)=.3 and P(~H) = .7.

Suppose that the probability of collusion (of all kinds) given ~H equals .1. For the illustration, let collusion (of various kinds) be the only dependence concern given ~H. E_1' and E_2' are independent otherwise given ~H. (Other types of hypotheses that would raise dependence problems could also be modeled using the tools in this paper.)

Subdivide collusion into three possible options:

Full copying (FC): All content will be identical. $P(E_2'|E_1'\&\sim H \ \& \ FC) = 1$.

Possible copying and possible variation (PC): Witnesses collude concerning overlapping content and may or may not vary (or even contradict each other accidentally) on other content.

Crafty collusion (Cr): Witnesses agree to overlap on some portions, agree not to overlap on some portions, do not mutually determine what precise details they will give when they don't overlap, but do make sure that these are not contradictory.

Suppose that

P(FC|~H)=.05

P(PC|~H)=.046

P(Cr|~H)=.004

(These sum to .1.)

$P(E_1'\&E_2'|\sim H\&FC) = (.003)$ (This follows from previous stipulations.)

$P(E_1'\&E_2'|Cr) = 0$ (This follows from the fact that the testimonies do not vary.)

Suppose further that

$P(E_1' \& E_2'|\sim H\&PC) = .0003$

The probability of the conjunction ($E_1'\&E_2'$) given ~H and no collusion equals the product of the individual probabilities: .000009, since there are no other dependence considerations.

Then, by the Theorem on Total Probability,

$P(E_1'\&E_2'|\sim H) = .0001719$

The consilience ratio for ($E_1' \& E_2'$) and ~H, then, is

Suppose that E_1' and E_2' are not negatively relevant given H. I will therefore treat them as independent given H for purposes of this calculation, since I am not claiming positive relevance given H as in Example 3. Then, by the odds form of Bayes' Theorem and the formula for including a dependence correction factor,

which yields a posterior probability of H--that is, $P(H|E_1' \& E_2')$--of approximately .6917. For comparison, if the evidence were treated as fully conditionally independent, the posterior probability of H would be approximately .977.

The dependence concern has significantly decreased the force of the two items of evidence.

Example 2: Partially overlapping testimony with simple varied details

Here there is overlapping content and partial variation of details. The details do not contradict each other but do not form any further coincidence. This is like example 1a in the paper. Call these E_1 and E_2.

Keep the individual Bayes Factors the same as in Example 1. Keep the prior probabilities of H and ~H the same as in Example 1.

Keep the probability given ~H of the collusion hypothesis overall and of the particular types of collusion given ~H the same as in Example 1.

Then,

$P(E_1\&E_2)|\sim H\&FC)= 0$, because the testimonies vary.

Suppose that

$P(E_1\&E_2|\sim H \& PC) = .00015$

$P(E_1\&E_2|\sim H\&Cr) = .0015$

As before, the independent probability of $(E_1\&E_2)$ given ~H and no collusion of any kind is .000009.

Then, by the Theorem on Total Probability,

$P(E_1\&E_2|\sim H) = .000021$

The consilience ratio for $(E_1 \& E_2)$ and ~H, then, is

Then, under the same assumptions as in Example 1 concerning relevance given H,

which yields a posterior probability of H--that is, $P(H|E_1 \& E_2)$--of approximately .948. The variation of the details has resulted in a posterior probability of the hypothesis much closer to the posterior using full independence.

Example 3: Variation with dovetailing

Keep the individual Bayes Factors the same as in example 1. Keep the prior probabilities of H and ~H the same as in Example 1.

The testimonies in this example overlap in part and vary in part, as in Example 2, and they also contain details that not only vary but also form a dovetailing coincidence, as in the case of the picnic given in the paper. Call the testimonies here T_1 and T_2.

To understand the positive relevance given H, consider a subhypothesis of H that says that not only did the event named in H occur but also that the additional detail given in one of the witness testimonies (such as the presence of a Joyce scholar at the event) is true. Call this subhypothesis H'. Suppose that

$P(H'|H) = .05$

Suppose that, under the assumption of H, T_1 confirms this subhypothesis and thereby confirms T_2. The confirmation from T_1 to T_2 under the assumption of H can be regarded as collapsing the effect of two steps under the assumption of H. Given H, we envisage learning that the first witness attests to H. This increases this witness's own credibility by (say) a factor of 2 (over the

previous force of 10/1), since he has attested to H, which is given as true. This witness attests to a further detail, which he now confirms fairly substantially. Given this testimony, within this hypothetical distribution P* *modulo* H, P*(H'|H) = .513.

Suppose, given empirical considerations and the nature of the differences in testimony, T_1 and T_2 are independent given (H&H') and (H&~H'). Hence, given this screening, support for T_2 can flow "through" H' from T_1 to T_2, given H.

Suppose that $P(T_2|H')=.3$ (Hence, given the screening described, $P^*(T_2|H') = .3$.)

Suppose that H' is the only relevant dependence hypothesis given H. T_1 and T_2 are otherwise independent given H.

Thus, by the Theorem on Total Probability,

$P(T_2|H \ \& \ T_1) = .16851$

By the general conjunction rule, the consilience ratio for $(T_1 \& T_2)$ and H is

The evidence in this case should certainly be no *more* dependent given ~H than in Example 2 (and plausibly less dependent given ~H). If we use the consilience ratio for ~H calculated in Example 2, then, by the odds form of Bayes' Theorem and applying all correction factors,

which yields a posterior probability of H--that is, $P(H|T_1 \ \& \ T_2)$--of approximately .9904. Arguably, consilience on ~H should be even lower in this case; one could argue that collusion should need to be "ultra-crafty" to bring about the special dovetailing of details, making dependence on ~H negligible.

If all dependence on ~H is eliminated while the positive dependence given H remains, the posterior probability of H is approximately .996. Both .9904 and .996 are even better than the posterior probability we would get by treating the testimonies as independent--approximately .977.

[1] The *s* measure--*P(H|E) - P(H|~E)*--rather surprisingly does not meet this condition. *S* allows for cases where $P(H|E_1) > P(H|E_2)$ but the confirmation of H by E_2 according to *s* is greater than or equal to the confirmation by E_1 (Fitelson 2001b, p. 19).

[2] Some other categories: a) The different statements or different portions of a long statement or document by the *same* witness/author contain casual, incidental details that fit together while appearing undesigned. b) There seems to be some fact "standing behind" and confirmed by different statements, though none of the statements explicitly affirms that fact.

[3] The analysis of alleged failures of the variety-of-evidence thesis by François Claveau and Olivier Grenier (2017) falls into this category. Claveau and Grenier do not define increased variety of evidence in the same way that I do (as increased independence given ~H), nor do they always hold constant the confirmation of H by individual reports between the cases being compared.

[4] One of the above quotations makes reference to similarity and variation both of language and of content. Which of these is more valuable overall is not a matter that admits of *a priori* determination. For one thing, it would be quite difficult to hold all else equal while comparing a case where there is similarity of semantic content in testimony but differences of specific terminology to a case where there is a clear difference of content. It could easily happen that the *individual* confirmations offered by the testimonies to the hypothesis of interest will not be the same between these. Variation of language will tell against certain subhypotheses of ~H while variations of content will disconfirm others, so the value of the variation could depend on the prior probability of these subhypotheses of ~H. Testimonial cases with differing content rather than solely varied language participate more often in undesigned coincidences and are easier to recognize. Most or all of the categories (depending on how one interprets 1a) *require* varied content. Variation of content might turn out to be valuable more often than variation of language, and I will be discussing variation of content throughout this paper. On the other hand, varying content permits the possibility of at least apparent contradiction, which is a complex issue (see section 7). Nothing in my probabilistic analysis of intuitively valuable undesigned coincidences requires that I decide whether differences of content or language are better or worse for confirmation *across the board*.

[5] Schubert and Olsson (2007) have shown that coherence of the contents of reports is "reliability conducive" in a restricted set of cases. In their proof, the fact that two sources report *exactly* the same content tends to raise the probability that each of them is *perfectly* reliable. I do not require that the witnesses report exactly the same content; as emphasized throughout, the variants in their content help to support their independence. Nor am I concerned only with perfect reliability.

[6] My argument here avoids Stefan Schubert's result (2012) that coherence is not *generally* conducive to reliability. In the scenarios I am discussing it is possible for all of the contents to which the sources testify to be true, whereas Schubert's proof involves a scenario in which the set of all testimonial contents is inconsistent. Also, Schubert's proof concerns the reliability of each source taken singly rather than an hypothesis that all sources are reliable (2012, p. 612).

[7] To support a reliability hypothesis to any significant extent, the details in question should be somewhat improbable, not part of general knowledge. If both reporters would certainly have known that the city of Milwaukee exists, even if they were ignorant about Lincoln, it does not create any significant undesigned coincidence if both sources merely agree on the *existence* of Milwaukee. Both sources might be unreliable concerning the events they report while both being right on that point. When the coincidence concerns something difficult to get right, such as that Lincoln traveled to Ohio right after being in Milwaukee, and when the sources are relevantly independent, the probability of the conjunction of testimonies is quite low given ~H.

[8] That this analysis can help practitioners is shown by the fact that scholars in the humanities are sometimes confused in ways that my analysis would prevent. For example, New Testament scholar Michael Licona (2016) speculates that the differences between the infancy accounts about Jesus in Matthew and Luke may be the result of imaginative embellishment by the authors, going beyond a more minimal set of statements held to be factual in the

Christian community at the time the accounts were written. Yet Licona attempts to use the differences between the accounts to argue for multiple, independent attestation to the overlap. The probabilistic analysis in this paper, aside from specific numbers, shows why this is a misstep. On the suggested hypothesis, the differences do not attest to independence given the *negation* of the overlap nor to causally separate access to the truth of the overlap. The shared contents might well be *false* while both authors got them from commonly held tradition, which they elaborated. Independence of fictional imagination is the wrong kind of independence to assist confirmation.