Accounting for Dependence:

Relative Consilience as a Correction Factor in Cumulative Case Arguments

Lydia McGrew

Abstract: I propose a measure of dependence that relates a set of items of evidence to an hypothesis *H* and to its negation. I dub this measure relative consilience (RC) and propose a method for using it as a correction factor for dependence among items of evidence. Using RC, I examine collusion and testimonial independence, the value of diverse evidence, and the strengthening of otherwise weak or non-existent cases. RC provides a valuable tool for formal epistemologists interested in analyzing cumulative case arguments.

Keywords: Bayesianism; coherence; dependence; independence; testimony

1. Introduction

Suppose that two witnesses testify to the same content—say, that they saw Alex commit a murder or that the weapon known to be the murder weapon belongs to Alex. *Prima facie* it seems that we can build a cumulative case against Alex from their testimony, that this case is stronger than if there were only one witness, and that this case would be even stronger if we had even more witnesses agreeing with these two.

Immediately, however, the question of independence arises. If, for example, the witnesses are close friends and have had a chance to talk with each other before giving their testimony, or if one of them has heard the other's testimony, the cumulative case for Alex's guilt immediately

seems weaker. One might have influenced the other to say that it was Alex that he saw or that the weapon is Alex's even if it was not. People are suggestible, after all. Or, if they have talked together in private, perhaps they have colluded for some unknown reason to pin the blame on Alex, even if he is innocent.

Such questions of independence have been especially important in the literature surrounding Bayesian modelling of the coherence of the contents of witness testimony. Generally an independence condition is stipulated so as to give the most plausible model of the intuitive sense that coherence of the contents of witness testimony is beneficial to the hypothesis supported by the items of evidence taken separately. Since it seems intuitively that it is *independent* testimony that is epistemically powerful when its contents are the same or similar, Bayesian coherence theorists try to capture the necessary independence by a condition placed upon the cases under consideration. One suggested independence condition has been independence of the items of testimonial evidence given both H (the hypothesis to which the individual items are positively relevant) and $\sim H$ (see Olsson [2005: 27–28, 45–56]). Another suggested independence stipulation is independence of the testimonies given $\sim H$ [Lewis 1946: 344, 349], without stipulating that they are independent given H. A third suggested stipulation is independence of the testimonies given the contents of the testimonies and the negations of the contents, where the conjunction of the contents might or might not be identical to H [Shogenji 2013: 2529].

One problem, however, is that most real-life cases are messy, and it can be difficult to know if these stipulated conditions hold. When the content of two witness testimonies is identical down to its details, the suspicion immediately arises that there has been some sort of collusion that should lessen their cumulative force. It is often difficult to be sure that we have

prevented witnesses from talking to each other ahead of time or from hearing each other's testimony. Moreover, as Tomoji Shogenji has pointed out [2013: 2529], agreement between witnesses about some fact other than *H* that tends to support *H*—say, that the weapon belongs to Alex—can be explained by scenarios in which they are testifying truly but *H* is still false. Hence, if *H* is the hypothesis that Alex is the murderer and two witnesses both say that the weapon belongs to him, perhaps they are speaking truly, Alex is innocent, but someone else used his weapon to commit the murder. Therefore, even if we know that they have not directly influenced each other, there is reasonable concern that the items of evidence are not relevantly independent and hence that there is a danger of overestimating their force for *H* (see Lydia McGrew [forthcoming]). I will discuss this scenario further below.

On the other hand, there are scenarios where there is literally no case at all for *H* from the individual items of evidence but where their cumulative force strongly supports *H*. Michael Huemer [2007] has argued for this reason that it is important *not* to stipulate that the items of evidence are independent given *H*, and Timothy McGrew [2003] has given a concrete example in the case of gravitational lensing, which I will discuss below. So sometimes dependence among items of evidence tends to favour *H* rather than lessening the impact of the case for *H* (see also Wheeler [2012: 146–7]).

I wish to suggest a measure of the dependence of the items of evidence under consideration, which I call relative consilience. In the case of testimonial evidence, this measure will concern the report propositions—the propositions that someone has testified thus—rather than the *contents* of the testimonies. This measure, I will argue, can be used as a correction factor for a cumulative case after initially calculating the force of the case under the assumption of independence among items of evidence. If the calculation under independence tends to

overestimate the strength of the cumulative case for H, the relative consilience ratio will correct that overestimate, and the same *mutatis mutandis* if the calculation under independence underestimates the strength of the cumulative case for H. The consideration of this dependence factor clarifies a variety of issues, including the value of diverse evidence.[1]

2. Relative consilience

I dub the measure I wish to discuss a measure of relative consilience, or RC. RC compares the extent to which the items of evidence are relevant to each other on the assumption of H, and the direction of that relevance, to the extent and direction of their mutual relevance on the assumption of $\sim H$. RC therefore relates H and $\sim H$ to a fixed body of evidence $\boldsymbol{E}$. For a set of items of evidence $\{E_1 \ldots E_n\}$, $\boldsymbol{E}$ is the conjunction $(E_1 \,\&\, \ldots \,\&\, E_n)$.

The numerator of RC is a ratio that shows to what extent the items of evidence in $\boldsymbol{E}$ are positively relevant to each other or negatively relevant to each other, or whether they are independent of each other, conditional on an hypothesis H.[2] Following the terminology of Timothy McGrew (2003), who calls this 'theoretical consilience', I will refer to this ratio as C, for consilience. For some set of items of evidence $\{E_1 \ldots E_n\}$ and an hypothesis H, the C ratio is

$$P(\boldsymbol{E}|H)/P(E_1|H) \times \ldots \times P(E_n|H).$$

[1] A version of this method was used in Timothy McGrew and Lydia McGrew [2009: 633–34].
[2] See McGrew [2003: 562], Myrvold [1996: 662], Schupbach [2005: 597], and Shogenji [2013: 2532–2533] for discussions of the ratio of theoretical consilience under a variety of names. Myrvold [2003: 412–13] briefly mentions the log of the measure I call RC.

The numerator of the C ratio is the probability of the conjunction of all the items of evidence, conditional on *H*. The denominator is the product of the probabilities of the individual items of evidence, conditional on *H*. If the items of evidence are all independent of each other when we assume *H* to be true, the numerator and the denominator of C will be equal—the conjunction of the items of evidence will be neither more nor less probable, given *H*, than the product of the individual conditional probabilities of the items of evidence, given *H*. In that case, C will be equal to 1. However, if the items of evidence are negatively relevant to each other when we assume *H*, the C ratio will be less than 1. If they are positively relevant to each other when we assume *H*, the C ratio will be greater than 1.

C can be used to relate ***E*** to any hypothesis H_i, and we could use C to compare hypotheses that do not form a probabilistic partition. But for my purposes, I am interested in C as applied to *H* and to its negation, *~H,* and in the relation between these consiliences. The denominator of my proposed measure RC is therefore this same C ratio, *mutatis mutandis*, for ***E*** and *~H*. Therefore, for *H* and a body of evidence ***E***, relative consilience (RC) is defined thus:

$$RC_H(\boldsymbol{E}) =_{df} \quad \frac{P(\boldsymbol{E}|H)/P(E_1|H) \times \ldots \times P(E_n|H)}{P(\boldsymbol{E}|\sim H)/P(E_1|\sim H) \times \ldots \times P(E_n|\sim H).}$$

RC is a ratio of ratios—the ratio of the consilience of *H*, *vis a vis* the evidence, to the consilience of *~H, vis a vis* the same evidence. I call it relative consilience because it is comparative in this way.[3]

As already mentioned, if the items of evidence are independent of each other given each of *H* and *~H,* both the numerator and the denominator of RC are equal to 1. In that case, of

[3] In a series of papers, Gregory Wheeler [2009, 2012] et. al. [Wheeler and Scheines 2010, 2013, Wheeler and Schlosshauer 2011] have promoted a measure that Wheeler calls focused correlation, or FOR. FOR cannot be used as a correction factor when the case under independence is modeled using Bayes factors.

course, RC itself will be equal to 1. There are other ways for RC to be equal to 1, but it is a sufficient condition if its numerator and denominator are both equal to 1.[4,5] If the items of evidence are more positively dependent conditional on *H* than on ~*H*, RC will be greater than 1; it will be less than 1 if they are more positively dependent conditional on ~*H* than on *H*.

3. Relative consilience as a correction factor for dependence

Suppose that each of two items of evidence E_1 and E_2 supports *H*. For purposes of this paper I will use *l*, the likelihood ratio (the Bayes factor),

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

as a measure of confirmation provided by *E* to *H*. I am aware of the controversy surrounding the best measure of confirmation; however, the *l* measure provides a straightforward way to combine the force of independent items of evidence. As I will show, the *l* measure of confirmation can be used with individual items of evidence, a conjunction of evidence, and RC to correct for dependence factors in judging the strength of a cumulative case. Hence, it is the most convenient measure for my purposes here.

Suppose that $P(E_1|H) = .1$ and $P(E_1|\sim H) = .01$ and that the same is true of E_2. Then the *l* measure for each of E_1 and E_2 is 10/1 in favour of *H*. Each item confirms *H* to this extent. If E_1

[4] If, for example, the items of evidence are somewhat positively relevant to each other given *H* and also given ~*H*, and if this degree of conditional positive relevance (as measured by C) is *equal* for *H* and ~*H*, then RC will also be equal to 1.

[5] In the notation used in this paper *H* and ~*H* designate propositions which form a probabilistic partition; probability for purposes of this paper is defined over a propositional language, though it would be possible to translate my argument into the terminology of binary variables. In this paper, *H* does not represent a variable; hence, the phrase '*H* screens off the items of evidence from each other' as I use it *does not* entail that ~*H* also screens.

and E_2 are independent in the relevant sense, the force of the cumulative case they provide for H can be measured quite easily as follows:

$$\frac{P(E_1 \,\&\, E_2|H)}{P(E_1 \,\&\, E_2|{\sim}H)} = \frac{P(E_1|H)}{P(E_1|{\sim}H)} \times \frac{P(E_2|H)}{P(E_2|{\sim}H)} = \frac{.01}{.0001} = \frac{100}{1}$$

This helps to show why a cumulative case made up of independent items of evidence, each of which favours the same hypothesis, can become so surprisingly strong (see Earman [2000: 55ff]).

The independence condition is the sticking point. For this straightforward multiplication of Bayes factors to be a completely unproblematic way to measure the strength of the cumulative case provided by E_1 and E_2, it must be the case that these items are independent of each other such that

$$P(E_1|H \,\&\, E_2) = P(E_1|H)$$

and

$$P(E_1|{\sim}H \,\&\, E_2) = P(E_1|{\sim}H).$$

This relation is known as screening off. If these conditions hold, both H and ~H screen off the items of evidence from each other. This is the paradigmatic situation in which we can be quite sure that multiplying the individual Bayes factors is an accurate measure of the force of the

cumulative case. In terms of the RC ratio, this condition requires that both the numerator and the denominator of RC be 1, making RC as a whole equal to 1.

I propose that, in order to take dependence among items of evidence into account, one can proceed heuristically as follows: Calculate or estimate the force of the cumulative case of the evidence *as if* the items of evidence were screened off from each other given H and given $\sim H$. Then, multiply the ratio thus calculated or estimated by a calculation or estimate of RC to correct for dependence among the items of evidence.

That RC is an accurate correction factor for dependence in a cumulative case, if the force of the case is measured using Bayes factors both for the individual items and for the conjunction of the items, is formally easy to see. As Shogenji [2013: 2532–3] has pointed out, $P(\mathbf{E}|H)$ is equal to the product of the probabilities of the individual items given H times what Shogenji calls the conditional coherence (that is, the C ratio) for the conjunction given H. This follows trivially from the fact that $P(\mathbf{E}|H)$ is the numerator of the C ratio and that the product of the individual conditional probabilities is the denominator. So

$$P(\mathbf{E}|H) = [P(E_1|H) \times \ldots \times P(E_n|H)] \times [P(\mathbf{E}|H)/P(E_1|H) \times \ldots \times P(E_n|H)].$$

And the same is true, *mutatis mutandis*, for $P(\mathbf{E}|\sim H)$. Therefore, the Bayes factor for the cumulative case, that is,

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

is equal to

$$\frac{P(E_1|H)}{P(E_1|\sim H)} \times \ldots \times \frac{P(E_n|H)}{P(E_n|\sim H)} \times \frac{P(\mathbf{E}|H)/P(E_1|H) \times \ldots \times P(E_n|H)}{P(\mathbf{E}|\sim H)/P(E_1|\sim H) \times \ldots \times P(E_n|\sim H).}$$

This representation of the confirmation provided by **E** is a model of the strategy I have suggested—namely, calculating the case as if the items of evidence are independent given both *H* and *~H* and multiplying by RC to correct for dependence.

A toy case of dependence illustrates the way that RC works when there is complete collusion between witnesses. (I discuss partial dependence below.) Imagine that E_1 and E_2 are items of testimonial evidence and that we become aware that they are not independent. In fact, complete collusion is taking place. One witness will copy exactly what the other says concerning Alex's guilt *regardless of whether Alex actually committed the murder or not*. Let us say that we do not know whether either of them was in more of a position than the other to know about Alex's guilt and that the individual conditional probabilities and Bayes factors are as stated above (an individual item of testimony is ten times more likely given *H* than given *~H*). We do know that, whatever one witness says, the other will say.

In that case, $P(\mathbf{E}|H) = P(E_1|H) = .1$, because $P(E_2|H \,\&\, E_1) = 1$. And $P(\mathbf{E}|\sim H) = P(E_1|\sim H) = .01$, because $P(E_2|\sim H \,\&\, E_1) = 1$. Therefore,

$$RC_H(\mathbf{E}) = \frac{.1/.01}{.01/.0001} = \frac{1}{10}$$

If we calculate the force of the cumulative case under independence and use RC as a correction factor, we get precisely the intuitively correct result:

$$\frac{P(E_1 \,\&\, E_2|H)}{P(E_1 \,\&\, E_2|{\sim}H)} = \frac{P(E_1|H)}{P(E_1|{\sim}H)} \times \frac{P(E_2|H)}{P(E_2|{\sim}H)} \times \frac{.1}{1} = \frac{.01}{.0001} \times \frac{.1}{1} = \frac{10}{1}$$

The 'cumulative case' is not a cumulative case anymore where there is known complete collusion as described above. The force of the case is reduced to the force of just one of the items of evidence.

4. More complex cases: Partial Dependence Favouring ~H

The toy case in the previous section illustrates the way that RC works as a correction factor. In any real situation of the kind described, we would simply discount the second item of evidence, given the knowledge of complete dependence, and reduce the force of the duplicate witness testimony to the force of a single item without bothering to go through a calculation.

RC can also be used when there is only partial dependence given *H*, given ~*H*, or given each of these. In examples of such cases of partial dependence, RC reveals a surprising result which helps to explain why the possibility of dependence is so often regarded as 'worse for' the *H* supported by the individual testimonies.

In the simple case above, the probability of complete dependence between the two testimonies was treated as 1 given either *H* or ~*H*—dependence was certain and total. However, dependence yields a correction factor even when dependence is uncertain and only partial. What is surprising is that it is possible to construct a case in which partial dependence is in a fairly intuitive sense equal for *H* and ~*H* but in which it turns out that this equal dependence is helpful to ~*H*.

For purposes of illustration, suppose that the individual likelihoods are as given in the case of complete dependence in section 3. That is, $P(E_1|H) = .1$ (and the same for E_2) and $P(E_1|\sim H) = .01$ (and the same for E_2). If these two items of evidence were screened off by H and by $\sim H$, the conjunction of the evidence would provide a cumulative Bayes factor in favour of H of 100/1. Now suppose that neither H nor $\sim H$ screens off E_1 and E_2 from each other. Instead, given each of H and $\sim H$, E_1 confirms E_2 by 10/1 odds. Just in this sense, we can say that equal dependence has been stipulated given H and given $\sim H$. Call the conjunction of the evidence $\boldsymbol{E'}$. Does this dependence yield an RC ratio which 'eats up' part of the confirmation of the independent cumulative case?

The calculation for answering this question is made somewhat simpler by the fact that, when there are only two items of evidence, the C ratio for H is equal to

$$P(E_2|H \,\&\, E_1)/P(E_2|H),$$

and the same *mutatis mutandis* for $\sim H$.[6]

If $P(E_2|H)$ is .1, and if, given H, E_1 confirms E_2 by a factor of 10/1, then $P(E_2|H \,\&E_1) \approx .5263$. If $P(E_2|\sim H)$ is .01, and if, given $\sim H$, E_1 confirms E_2 by a factor of 10/1, then $P(E_2|\sim H \,\& E_1) \approx .0917$.[7] Plugging these numbers into the C ratios for H and $\sim H$ yields,

$$RC_H(\mathbf{E'}) = \frac{.5263/.1}{.0917/.01} = \frac{5.263}{9.17} \approx \frac{1}{1.74}.$$

[6] $P(E_1\&E_2|H)/[P(E_1|H) \times P(E_2|H)] = [P(E_1|H) \times P(E_2|H \,\&E_1)]/[P(E_1|H) \times P(E_2|H)] = P(E_2|H \,\&\, E_1)/P(E_2|H)$.

[7] $P(E_2|H)/P(\sim E_2|H) \times P(E_1|H \,\&\, E_2)/P(E_1|H \,\&\sim E_2) = P(E_2|H \,\&E_1)$. Plugging in what has been stipulated, this comes to .1/.9 × 10/1 = 1/.9 which yields $P(E_2|H \,\&\, E_1) \approx .5263$. The similar calculation for $\sim H$ is .01/.99 × 10/1 = .1/.99 which yields $P(E_2|\sim H \,\&\, E_1) \approx .0917$.

What does this mean in concrete terms? For illustration, suppose that the prior probability of *H* is low—a mere .01. Then the case from E_1 and E_2 under independence counterbalances the low prior probability of *H* so that its posterior probability is a more respectable .5.[8] But given the amount of dependence in the case just discussed, the case is weakened to such an extent that the posterior probability of *H* (having a prior probability of .01) is only approximately .37.[9,10]

Why is it that an equal (in one sense) amount of dependence conditional on *H* and on ~*H* seems to favour ~*H* in this way? The answer lies in the contrast between the C ratios which form the numerator and denominator of RC. We have defined ~*H* as the hypothesis that gives lower individual probability to the items of evidence. Under an assumption of screening off, the numerator of the C ratio for ~*H*, like the denominator, is the product of the individual probabilities of the items of evidence given ~*H,* and the same *mutatis mutandis* for *H*. Therefore, under the assumption of screening off, the numerator of the C ratio for ~*H* is by definition smaller than the numerator of the C ratio for *H*. When the screening-off assumption is abandoned, conditional positive dependence of the items of evidence 'boosts' the numerator of each of the C ratios. The 10/1 Bayes factor by which E_1 confirms E_2, conditional on either hypothesis, does more work to change the probability of E_2 conditional on ~*H*, and hence does more to change the numerator of the C ratio for ~*H*, since the probability of each item of evidence is lower on ~*H* than on *H*. (See footnote 7.)[11] This means that the C ratio for ~*H* is

[8] By the odds form of Bayes's Theorem, .01/.99 × 100/1 = 100/99 which yields about .5 posterior probability.
[9] By the odds form of Bayes's Theorem, .01/.99 × 100/1 × 1/1.74 = 1/1.72 which yields about .37 posterior probability. Since a proposition's probability is always between 0 and 1, the difference between the posterior probabilities of *H* under independence and dependence in this case will appear more dramatic when the prior probability of *H* is lower.
[10] This point concerning prior and posterior probability remains true if one models the force of the evidence in terms of the log of the Bayes factor(s) rather than the simple ratio(s) and the RC measure as a log instead of a simple ratio. My thanks to a reviewer for raising the question of mathematical scale. My thanks to Timothy McGrew for helping to clarify my thinking concerning questions of scale.
[11] My thanks to Timothy McGrew for helping me to catch an error I was previously making in my analysis at this point.

helped more by this degree of positive dependence among the items of evidence than the C ratio for *H*. This is why the correction for independence, represented by RC, 'eats up' part of the cumulative force of the evidence for *H* in such a case.

This illustration helps to explain why there tends to be more concern in a testimonial case about the possibility that an independence assumption causes an *overestimate* of the cumulative case in favour of *H*. While it is not true *in general* that independence of evidence favours *H* and that dependence is unfavourable to *H* [Wheeler 2012: 146ff, T. McGrew 2003, Huemer 2007], there are plausible circumstances in which a degree of dependence of the evidence that is in *some* understandable sense equal between *H* and ~*H* does reduce the force of a cumulative case for *H*. This may also explain why C. I. Lewis [1946: 344, 349], in discussing the value of what he calls congruence in evidence, stipulates that he is asserting the power of congruent evidence only on the assumption that the items of evidence are screened off from each other by ~*H*.

5. Diverse evidence and relative consilience

Branden Fitelson [2001: S132–4] has argued, *contra* Howson and Urbach [1993: 113–114], that the value of diverse evidence does not appear to lie in the fact that diverse evidence is unconditionally independent. Fitelson suggests [2001: S130ff] that one aspect of the value of diverse evidence is explained by what he calls confirmational independence. As Fitelson defines the phrase [2001: S125], two items of evidence E_1 and E_2 are confirmationally independent regarding *H* iff, for some measure of confirmation *c*, $c(H, E_1|E_2) = c(H, E_1)$ and $c(H, E_2| E_1) = c(H, E_2)$ and the same, *mutatis mutandis*, for ~*H*. That is, *H* is confirmed by E_1 to just the same

extent regardless of whether or not E_2 has already been taken into account, and *vice versa*. Fitelson further points out that a sufficient (though not a necessary) condition for confirmational independence when one is using the likelihood measure of confirmation is that both *H* and ~*H* screen off the items of evidence from each other.[12]

I suggest, using RC, that there is an interesting class of cases in which the value of diverse evidence lies in the ability of diverse evidence to render the items of evidence less dependent given ~*H*—that is, less positively relevant to each other, entirely independent of each other (where they would otherwise be positively relevant), or even negatively relevant to each other, conditional on ~*H*. This use of diverse evidence is important in view of the point of the previous section that what might seem like a small amount of positive dependence conditional on ~*H* reduces the force of a cumulative case for *H*.

Here I want to introduce the concept of a subhypothesis and distinguish two ways in which a subhypothesis might affect the C ratio and hence to the RC measure for some hypothesis. A subhypothesis of *H* is, informally speaking, a way in which *H* might be the case. For my purposes here I will always be using the term for what one might call a 'proper' subhypothesis. That is to say, a subhypothesis of *H* has a probability between 0 and 1, it entails *H*, but it is not entailed by *H*. For example, a subhypothesis of the proposition that Jones shot Smith would be, 'Jones shot Smith, and many people saw the event take place'. A subhypothesis of the proposition that Jones did not shoot Smith would be, 'Jones did not shoot Smith, but several people are colluding to frame Jones'.

[12] We can define 'confirmationally independent or better' for E_1 and E_2 and some *H* as $c(H,E_1|E_2) \geq c(H,E_1)$ and $c(H,E_2|E_1) \geq c(H,E_2)$. When there are only two items of evidence in **E**, it follows trivially from the definition of RC and the definition of the *l* measure that E_1 and E_2 are 'confirmationally independent or better' for *H* by the *l* measure of confirmation just in case $RC_H(\mathbf{E}) \geq 1$. Unfortunately, it is impossible to generalize this result for E_n where $n > 2$. By the definition of RC, the confirmation from **E** *as a whole* is as strong as or stronger than the case under independence just in case $RC_H(\mathbf{E}) \geq 1$. However, if the set of Es is larger than 2, it does not follow that the confirmational force of *each individual item* is as good as or better than its individual force, conditional on all of the others, when $RC_H(\mathbf{E}) \geq 1$. Proof omitted.

In cases where we are concerned that witnesses are colluding, the subhypothesis that creates dependence among items of evidence claims *that* the items of evidence are dependent. The collusion subhypothesis is that the items of evidence are causally connected—the witnesses are influencing one another. It is also possible for a subhypothesis to contribute positively to the C ratio of the hypothesis it entails (which I will sometimes refer to as the 'umbrella hypothesis') even if the subhypothesis does not state that the items are dependent and even if it screens off the items of evidence from each other. This occurs when the subhypothesis gives higher probability to each of the individual items of evidence than the average probability given to each of them by the umbrella hypothesis.[13] There are therefore two different kinds of subhypotheses that help to increase the C ratio for an umbrella hypothesis, and they work in different ways.

The example of Alex's weapon, discussed at the outset, can be seen as the second type of subhypothesis. Even if we have somehow completely ruled out any direct influence between the two testimonies that the weapon belongs to Alex, the two testimonies that the weapon belongs to Alex are still plausibly somewhat dependent given ~*H*—that Alex is innocent. The reason is that, if Alex is innocent but the weapon does belong to him, the subhypothesis of ~*H* that Alex is innocent and that someone else used his weapon to commit the crime provides more reason to expect each of the individual testimonies than ~*H* does on average. Hence, the subhypothesis, 'Alex is innocent, but the murder weapon belongs to him' provides higher probability to the conjunction of the testimonies than ~*H* as a whole, even if direct influence between the witness testimonies has been ruled out.

[13] A subhypothesis that screens and gives higher-than-average probability to some two or more items of the evidence under consideration but not to all of them *may* contribute to the consilience of the umbrella hypothesis, depending on the probability it gives to the other items of evidence.

I will use the notation $(\sim H)_i$ for a subhypothesis *of* $\sim H$. I place the subscript outside of the parentheses to make it clear that this notation does not refer to the negation of H_i but rather to a way in which $\sim H$, specifically, might be true.

Suppose that one were given $\sim H$ with certainty and received E_1—the first testimony that the weapon belongs to Alex. Under the assumption of $\sim H$, the first item of testimony would confirm $(\sim H)_1$—that Alex is innocent but that the weapon does belong to him. That would in turn raise the probability of receiving E_2—a second testimony that the weapon belongs to Alex.[14]

Suppose that we are justifiedly confident that some items of evidence E_1 and E_2 are not *negatively* dependent conditional on H—that Alex is the murderer. That is, the numerator of RC in this case is equal to or greater than 1, though we have not given much thought to whether the evidence really is positively dependent or independent given H. Suppose, moreover, that E_1 and E_2 taken as independent items of evidence for H would constitute a decent cumulative case for H. This will be all the more true in situations when the number of items of evidence that have individual positive relevance to H is greater than two. Our chief concern, then, is whether or not the calculation under independence is overestimating the strength of the case for H.

The value of diverse evidence arises from the fact that more diverse items E_1 and E_2 make it more difficult to find a subhypothesis $(\sim H)_1$ that is helpful to $\sim H$ in either of the senses discussed above. Suppose that, instead of two witnesses who either both testify that they saw Alex commit the murder or both testify that the weapon belongs to Alex, we have one witness who testifies that the weapon belongs to Alex and also a surveillance camera video showing

[14] An interesting and complex corollary to this point is that, if the conditional probability of $(\sim H)_1$ given $\sim H$ rises (holding constant the probability of each of the items of evidence conditional on $(\sim H)_1$), a point is eventually reached at which $(\sim H)_1$ contributes relatively little to the consilience of $\sim H$. At that point $(\sim H)_1$ contributes instead to the conditional probability of the conjunction of the testimonies given $\sim H$ by the direct route of making the *individual* items, and hence their conjunction, more probable given $\sim H$, weakening the individual Bayes factors for the items in favour of H.

Alex at the scene of the crime on the relevant night. Here the issue of collusion between witnesses does not even arise. More to the point, it is difficult to come up with any *single* hypothesis that increases the consilience of these items given ~*H* in either of the senses I have discussed. It is possible that Alex is being framed somehow, and the prior probability of Alex's being framed as a subhypothesis of Alex's innocence may vary widely with the empirical specifics. But in the face of the diverse evidence, the 'somehow' would need to be at least somewhat conjunctive in order to be helpful to ~*H*. For example, the framing hypothesis would have to be something to the effect that someone was able to frame Alex by arranging both that his weapon was used in the crime and that he was present on the night of the crime (or by taking advantage of his just happening to be there that night). But this is really two hypotheses bundled into one, and evidence for one part of this conjunction gives, on many ordinary types of background evidence, relatively little confirmation to the other part of the conjunction. A framer who is able to arrange for the use of Alex's weapon is not *ipso facto* a framer who can arrange for Alex's presence to be caught on a surveillance camera on the night of the murder. If the different parts of a conjunctive framing subhypothesis are *completely* independent of one another, then one piece of evidence is not relevant to the other piece of evidence by way of supporting the framing hypothesis, and the framing subhypothesis is not helpful to the C ratio for ~*H* and hence does not create an RC correction factor that undermines the force of the evidence for *H*. The point will apply *a fortiori* if other items of diverse evidence come in—Alex's fingerprints inside the house where the murder took place, a witness who says that Alex uttered threats against the deceased, and so forth.

It is sometimes possible to invent a framing subhypothesis of ~*H* that gives higher probability (than the average probability on ~*H*) to all the items of evidence, but the cost of

doing so is that such a subhypothesis will have *extremely* low probability conditional on ~*H*. For example, one could hypothesize a vast, powerful, government conspiracy against Alex. This would give higher probability than the average on ~*H* to a certain amount of varied evidence, because one could guess that such conspirators could find a way to collect and plant Alex's fingerprints, steal and use his weapon in the crime, and perhaps even induce him to visit the relevant location on the night of the murder or at least predict when he would be there and arrange to commit the murder on that night. But the cost of inventing such a subhypothesis is that it will have low conditional probability given ~*H*, which will make it difficult for it to assist the overall consilience of ~*H*. The extent to which a subhypothesis $(\sim H)_1$ is helpful to the C ratio for ~*H* is negligible if $P((\sim H)_1|\sim H) \approx 0$.

If there is no remotely plausible available subhypothesis that helps ~*H* to explain the evidence, and if we have no reason to believe that the evidence is negatively dependent given *H*, the strength of the cumulative case can be estimated under the assumption of independence without *overestimating* its force in favour of *H*. When the case for *H* under independence is strong, this may be the most important point for purposes of deciding what is rational to believe.[15]

6. RC and dependence conditional on *H*

The previous sections have focused on applications to cases where the epistemic value of RC lies in drawing attention to or ruling out dependence conditional on ~*H*. Dependence considerations, modelled by RC, are also epistemically useful in drawing attention to positive dependence conditional on *H.* This point has interesting applications not only in the philosophy of science but

[15] For more on dependence conditional on ~*H* and diverse evidence, see Lydia McGrew [2016].

also in the perennial problem of the external world. Sometimes individual items of evidence have little or even no force for either *H* or ~*H* taken by themselves but do have force, or significantly greater force, for *H* when considered jointly. In that case, RC accounts for the force of the cumulative case.

The case of gravitational lensing discussed by Timothy McGrew[2003: 563] is of this kind:

> In 1979 two quasar images only five arcseconds apart, QSO 0957 + 561, were found to have identical spectral characteristics. Data on the spectra of known quasars indicated that there was only a remote probability of such a coincidence on chance; an explanation seemed called for. 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. . . .Why is the lensing explanation (L) so attractive? In large measure, it is because it eliminates a brute coincidence by providing a context in which the spectrum of the first quasar (S_i) is relevant to the spectrum of the second (S_2). By itself, L does nothing to raise the likelihood of either spectrum: $P(S_1/L) = P(S_1)$ and *mutatis mutandis* for S_2. But the coordination effected by L makes the dependence between the two almost complete, for if the two images have been formed by lensing from a single quasar then the spectrum of the one image is virtually guaranteed to match the spectrum of the other, i.e.,

$$P(S_1 \& S_2/L) \approx P(S_I/L) >> P(S_I/L)P(S_2/L)$$

In other words, the C ratio for the hypothesis L is high. McGrew implies, though he does not say it in so many words, that scientists had no reason to consider the two quasar image observations to be dependent given ~L—in other words, the items of evidence were randomly correlated given ~L. These points taken together made the likelihood ratio for their conjunction strongly favourable to L.

> L did not rise to the level of an established fact until subsequent observations from Mauna Kea and Palomar confirmed the presence of sufficient mass in the right place. But the loveliness of the explanation almost certainly motivated the search for that mass. From a heuristic perspective this is perfectly rational: with the high likelihood in place thanks to the consilience of L over randomness with respect to S_1 & S_2, all that remains for an inference is to establish by independent means the plausibility of L itself (McGrew 2003: 564).

The moral that McGrew draws is that an awareness of something like his consilience measure (which I am calling the C measure) played a tacit role in the evaluation of L. I would add that the implied comparison to randomness given ~L means that the complete picture includes the entire RC measure—the ratio of consilience on L to the absence of consilience on ~L. In fact, since the individual items of evidence were irrelevant to L taken separately, the *entire* force of the case from S_1 and S_2 lies in the RC measure.

A consideration of RC also plays an important role when individual items provide fairly weak evidence for *H* while their conjunction provides much stronger evidence than can be

accounted for by the product of the individual Bayes factors. Suppose that I am sometimes subject to visual hallucinations and sometimes subject to tactile hallucinations but have never had a hallucination that realistically correlates visual and tactile sensations. Now suppose that I have an experience in which I seem both to see an apple and to be able to touch and handle the apple normally. Let *H* be the hypothesis, 'I am not now hallucinating'. Arguably, the RC measure for the evidence of my visual and tactile sensations favours *H* over ~*H* in this scenario. If I am not hallucinating, my visual sensation confirms a subhypothesis H_1—that I am not hallucinating and that there really is an apple in front of me—which in turn gives me more reason to expect the tactile evidence than I would have given *H* alone. In other words, the two items of evidence are dependent given *H*. On the other hand, if I am hallucinating, then, given my hallucination history, the visual evidence gives me little or no reason to expect the tactile evidence. Since I have not previously had vivid visual-tactile hallucinations in which the sensations are realistically correlated, it is reasonable to think of the two sensory experiences as independent given ~*H*. If I am hallucinating, what appears to be an apple is quite likely to be impossible to touch, and the tactile sensation of holding an apple is quite likely to be correlated with some quite different visual experience or no visual experience at all.

Given the history of hallucinations, the product of the force of the *individual* items of evidence in this scenario will provide relatively weak support for *H*—the hypothesis that I am not hallucinating. The top-heavy RC factor will therefore be particularly epistemically helpful in concluding that I am not hallucinating, and all the more so if other sensory evidence, such as smell and taste sensation, is added to the case.

This example hints at the way that RC could play a role in the defence of realism against Cartesian Deceiver scenarios, though in that epistemic debate there would be no convenient

frequency data concerning different types of hallucinations. Nonetheless, it is plausible that the hypothesis that an external world exists and that I have fairly reliable memorial and sensory access to it brings together two apparent reports of, say, the existence of Paris better than does its negation, for reasons that are roughly similar to those in the hallucination case. Given realism, the first apparent report supports the subhypothesis that the external world exists *and also contains a real city of Paris*, which gives me more reason to expect another such report experience than one would have given realism alone. If realism is false, I have little or no additional reason to expect a second apparent report of the existence of Paris given one such apparent report.

These cases also illustrate the fact that the RC ratio may be accessible in a somewhat more gestalt fashion than by way of separately estimating or calculating its component parts—for example, the probabilities of the individual items given H and $\sim H$ or even the exact C ratios for H and $\sim H$. After critiquing an analysis of the gravitational lensing case given by Wesley Salmon, Timothy McGrew makes a similar point:

> The moral is that the phenomenology should not be wholly despised. Attention to our pre-theoretical notions of loveliness may at times be a surer guide to a theory's probabilistic merits and the structure of our reasoning than purely algebraic manipulations even in the hands of an acknowledged master (McGrew 2003: 564–5).

In arguing for a contested point, of course it should be possible to break down and explain the probabilistic virtues of the conclusion that one favours But it may be possible to do

this, as I have done above for the hallucination example, by discussing relative consilience without giving a separate account of every part of the measure.

7. Conclusion

Although it is quite simple to show that RC is an accurate correction factor for dependence when the case under independence is measured using Bayes factors, this technically simple point has wide-ranging and important epistemic implications. A consideration of RC illuminates epistemic concerns about collusion and testimonial dependence, including the common assumption that dependence among items of testimony tends to weaken the case for *H*. A consideration of RC and various ways in which subhypotheses of ~*H* might be helpful to the consilience of ~*H* helps to explain precisely how diverse evidence can be helpful to confirmation. Finally, RC sheds light on cases where dependence strengthens or even creates a case for *H*. Relative consilience should therefore become a standard part of the toolkit of formal epistemology in analyzing and evaluating cumulative case arguments.

References

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

Fitelson, Branden. 2001. A Bayesian Account of Independent Evidence With Applications, *Philosophy of Science* 68/3: S123–S140.

Howson, Colin and Peter Urbach 1993. *Scientific Reasoning: The Bayesian Approach*, LaSalle, IL: Open Court Publishing Co.

Huemer, Michael 2007. Weak Bayesian Coherentism, *Synthese* 157/3: 337–346.

Lewis, C. I. 1946. *An Analysis of Knowledge and Valuation,* LaSalle, IL: Open Court Publishing Co.

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

McGrew, Lydia forthcoming. Bayes Factors All the Way: Toward a New View of Coherence and Truth, *Theoria*.

McGrew, Timothy 2003. Confirmation, Heuristics, and Explanatory Reasoning, *British Journal for the Philosophy of Science* 54/4: 553–67.

McGrew, Timothy and Lydia McGrew 2009. The Argument from Miracles: A Cumulative Case for the Resurrection of Jesus of Nazareth, in *The Blackwell Companion to Natural Theology*, ed. William Lane Craig and J. P. Moreland, Oxford: Blackwell Publishing Ltd.: 593–662.

Myrvold, Wayne C. 1996. Bayesianism and Diverse Evidence: A Reply to Andrew Wayne, *Philosophy of Science* 63/4: 661–65.

Myrvold, Wayne C. 2003. A Bayesian Account of the Virtue of Unification, *Philosophy of Science* 70/2: 399–423.

Olsson, Erick J. 2005. *Against Coherence: Truth, Probability, and Justification,* Oxford: Clarendon Press.

Schupbach, Jonah N. 2005. On a Bayesian Analysis of the Virtue of Unification, *Philosophy of Science* 72/4: 594–607.

Shogenji, Tomoji 2013. Coherence of the Contents and the Transmission of Probabilistic Support, *Synthese* 190/13: 2525–2545.

Wheeler, Gregory 2009. Focused Correlation and Confirmation, *British Journal for the Philosophy of Science* 60/1: 79–100.

Wheeler, Gregory 2012. Explaining the Limits of Olsson's Impossibility Result, *The Southern Journal of Philosophy* 50/1: 136–150.

Wheeler, Gregory and Richard Scheines 2010. Causation, Association and Confirmation, Carnegie Mellon University Department of Philosophy. Paper 418. http://repository.cmu.edu/philosophy/418.

Wheeler, Gregory and Richard Scheines 2013. Coherence and Confirmation Through Causation, *Mind* 122/485: 135–170.

Wheeler, Gregory and Maximilian Schlosshauer 2011. Focused Correlation, Confirmation, and the Jigsaw Puzzle of Variable Evidence, *Philosophy of Science* 78/3: 376–92.