When Are Two Witnesses Better Than One?

David Atkinson and Jeanne Peijnenburg
University of Groningen
The Netherlands

Abstract

Even if two testimonies in a criminal trial are independent, they are not necessarily more trustworthy than one. But if they are independent in the sense that they are screened off from one another by the crime, then two testimonies raise the probability of guilt above the level that one testimony alone could achieve. In fact this screening-off condition can be weakened without changing the conclusion. It is however only a sufficient, not a necessary condition for concluding that two witnesses are better than one. We will discuss two different conditions, each of them necessary as well as sufficient, and we conclude that one of them is slightly better than the other.

One witness shall not rise up against a man for any iniquity, or for any sin, in any sin that he sinneth: at the mouth of two witnesses . . . shall the matter be established. *Deuteronomy 19:15.*

1 Introduction

Are two testimonies always better than one? If two witnesses that are generally known to be reliable give incriminating evidence, does this make the probability that the defendant is guilty greater than it would be with only one witness? Clearly the answer is in the negative. For one thing, the two testimonies may contradict one another.

Is it enough that the testimonies are not contradictory? Again the answer is no. Two independent testimonies, each of which by itself would increase the probability of guilt, could together actually conspire to reduce this probability. The next section will contain an example of such a surprising case.

Perhaps screening off, as a special case of the Markov condition, will fit the bill. Suppose that the testimonies are independent of one another, not unconditionally, but conditional on the defendant's being guilty, and also conditional on her being innocent. Is it now the case that the probability of guilt is greater with two incriminating testimonies rather than just one? L. Jonathan Cohen has shown that indeed it is.[1] In fact, Cohen demonstrated that even a weakened version of screening off is sufficient to obtain the result. Conditional on guilt, the testimonies may be either independent or negatively correlated, and conditional on innocence they may be either independent or positively correlated. Assuming that each of the two testimonies is indeed incriminating (i.e. each raises the probability that the defendant is guilty), weakened screening off guarantees that two testimonies are better than one.

Weakened screening off is however by no means necessary to draw this conclusion. This was first shown by L.J. O'Neill, who formulated a condition that is both necessary and sufficient.[2] The same condition was later given by George Schlesinger, who proves it in a different manner.[3] Both proofs in a sense rely on Cohen's demonstration. We will establish a slightly different necessary and sufficient condition, one that does not depend on Cohen's argument and which has some practical advantage over the O'Neill-Schlesinger condition.

We shall proceed as follows. In Section 2 we consider a murder trial and show that in general two independent testimonies may reduce the probability below what it was in the presence of only one testimony. In Section 3 we demonstrate that this is no longer the case if a Markov condition is in place: under that condition, two testimonies make it more probable that the crime has been committed. Section 4 is devoted to a relaxation of the Markov condition; and we explain how Cohen's argument can be considerably simplified. In Section 5 we outline the necessary and sufficient condition as formulated by O'Neill and Schlesinger, followed by an alternative condition that is likewise necessary and sufficient. In Section 6 we explain why the alternative is to be preferred in some cases.

2 Murder Most Foul

Alice and Bob live in a large house and have been married for many years. But one day Bob is found dead in the couple's bedroom. He had been shot

[1] Cohen 1976, Cohen 1977, 104-107.

[2] O'Neill 1982.

[3] Schlesinger 1991, 155-157. Schlesinger fails to mention O'Neill, although he must have been familiar with his result, having seen Cohen's reply to O'Neill (Cohen 1982).

in the head. Alice is a prime suspect as the perpetrator of the crime; she is arraigned and the prosecutor claims that

C : Alice killed her husband.

There are two witnesses for the prosecution, Clara and Deanna. Clara testifies

T_1 : Alice had a gun in her purse on the night of the crime.

Deanna claims that

T_2 : Alice had an argument with Bob on the night of the crime.

Clearly, each testimony increases the probability that Alice committed the crime:

$$P(C|T_1) > P(C) \quad \text{and} \quad P(C|T_2) > P(C)\,. \tag{1}$$

It might appear that both testimonies together should make it more probable that Alice killed her husband than would just one of the testimonies by itself. That is, if (1), it seems as though the following inequalities should hold:

$$P(C|T_1 \wedge T_2) > P(C|T_1) \quad \text{and} \quad P(C|T_1 \wedge T_2) > P(C|T_2)\,. \tag{2}$$

But of course this conclusion is not necessarily true. For suppose that Clara had asked Deanna to trump up some claim if she, Clara, should be asked to produce her testimony. Then it would seem clear that adding Deanna's testimony will not increase the probability of Alice's guilt above what it would have been if only Clara had testified.

One might guess that, if there is on the contrary no collusion or other relevant contact between Clara and Deanna, so that testimonies T_1 and T_2 are independent of one another, then the inequalities (2) should hold. However, independence of the testimonies, that is

$$P(T_1|T_2) = P(T_1)\,, \tag{3}$$

is in fact *not* a sufficient condition for the validity of (2). Here is a counterexample. Look at the probability distribution in the Venn diagram of Figure 1. The triple probabilities can be read off from the diagram; for example $P(T_1 \wedge T_2 \wedge C) = \frac{1}{64}$ and $P(\neg T_1 \wedge \neg T_2 \wedge \neg C) = \frac{33}{64}$.

With this probability distribution we check that (1) and (3) are satisfied, but

$$P(C|T_1 \wedge T_2) = \tfrac{1}{4} \;\; < \;\; \tfrac{7}{16} = P(C|T_1) = P(C|T_2)\,.$$

So in this example the two testimonies together would actually decrease the probability of Alice's guilt from the value that one testimony alone would produce, and this despite the fact that the testimonies are independent.

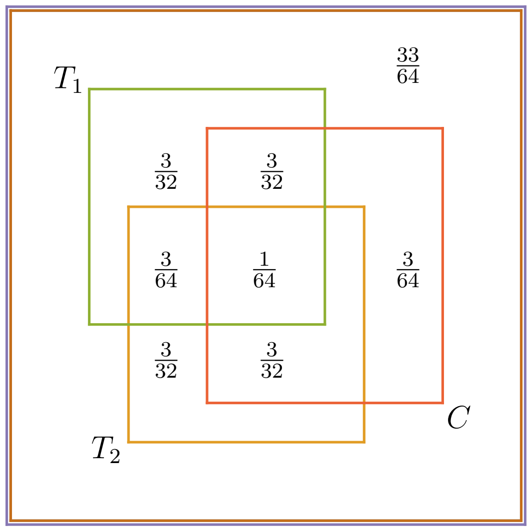

Figure 1: Triple probabilities for independent testimonies

In Section 6 we return to Alice and the hapless Bob; but first we shall look at some further examples in order to explain the condition of screening off.

3 The Markov Condition

The Markov condition has proven to be of use in various contexts: Hans Reichenbach introduces it as the screening-off requirement in his discussion of the common cause; and nowadays it is much applied in DAGs and in algorithms for search engines.[4] In the present discussion, the condition requires that the two testimonies by Clara and Deanna are screened off from one another in the sense that they are independent conditional on the defendant's being guilty, and also independent conditional on her being innocent. The difference between this condition and that of the previous section is that, in-

[4]Reichenbach 1956, 159; Pearl 2000.

stead of one restriction of *un*conditional independence, there are rather two conditional independence conditions.

Under screening off it does follow that two testimonies make the conclusion more probable than just one. Thus Erik Olsson writes:

> [I]n the context of Conditional Independence, Weak Foundationalism does imply Coherence Justification. Indeed, the combined testimonies will, in this case, confer more support upon the conclusion than the testimonies did individually.[5]

'Conditional Independence' for Olsson is precisely the Markov condition; and 'Weak Foundationalism' implies in our example that C is made more probable by one testimony than it was in the absence of a testimony. 'Coherence Justification' means that the combined testimonies also render C more probable than it was in the absence of a testimony. The second sentence of Olsson is the stronger claim that the Markov condition is sufficient to guarantee that the combined testimonies make it more probable that the crime was commited than does just one testimony.

Here is an example of a situation in which screening off holds sway; it applies to the reviewing of a scientific paper rather than a courtroom situation. Consider the following propositions:

C : The paper is of high quality.

T_1 : The expert reviewer #1 recommended publication.

T_2 : The expert reviewer #2 recommended publication.

If we do not know whether any of these statements are true, we will conclude that the paper is more likely of high quality if reviewer 1, or reviewer 2, recommended publication, then if no reviews were made, i.e. $P(C|T_1) > P(C)$ and $P(C|T_2) > P(C)$. Moreover, the probability that reviewer 1 recommended publication, given that the paper is of high quality, is the same as the probability that reviewer 1 recommended publication, given that the paper is of high quality *and* that reviewer 2 recommended publication, for the two reviewers are independent. That is, once we know that the paper is of high quality, it is very probable that reviewer 1 had recommended publication, and the recommendation of reviewer 2 does not affect this probability, so $P(T_1|C \wedge T_2) = P(T_1|C)$.

[5] Olsson 2017, 20.

This is not the full description of screening off, however. To achieve that we must add the corresponding equation in which C is replaced by its negation: $P(T_1|\neg C \wedge T_2) = P(T_1|\neg C)$. This condition means that the (low) probability that reviewer 1 recommended publication, given that the paper is *not* of high quality, is the same as the probability that reviewer 1 recommended publication, given that the paper is not of high quality, *and* that reviewer 2 recommended publication. In this case also, the recommendation of reviewer 2 does not affect the probability that reviewer 1 recommended publication.

L. Jonathan Cohen appears to be the first to have seen that the Markov condition is sufficient to guarantee that the combined testimonies make the defendant's guilt more probable than one would have done. In other words, he saw that two witnesses are better than one, i.e.

$$P(C|T_1 \wedge T_2) \quad > \quad P(C|T_i) \qquad i = 1, 2\,. \tag{4}$$

if the following three conditions are satisfied

$$\begin{aligned} P(C|T_i) &> P(C) \qquad\qquad i = 1, 2\,, \\ P(T_1|C \wedge T_2) &= P(T_1|C) \\ P(T_1|\neg C \wedge T_2) &= P(T_1|\neg C). \end{aligned}$$

Here the second and the third condition together constitute the Markov constraint. The first condition expresses the fact that the testimonies are indeed incriminating. In his book Cohen has expressed reservations about this condition, since it allegedly implies that some positive prior probability $P(C)$ is assignable, which appears to be at odds with the practice in many legal systems.[6] Although Cohen is of course right in stressing the tension between actual judicial custom and formal probability theory, his complaint about the rôle of $P(C)$ is perhaps overemphasized. Firstly, there is no need to attach a precise value to $P(C)$; it is enough to assume that it is not zero, since we are interested in an increase of probability, and only if $P(C)$ is positive can this take place. The numerical size of the increment is here of no concern: it is simply a premise that $P(C|T_i)$ is greater than $P(C)$. Secondly, assuming that $P(C)$ is not zero is in fact quite harmless. It reflects the logical possibility that a person is guilty, or in other words, that her being innocent is not a tautology. It seems to us that such an assumption has to be made, on

[6] Cohen 1977, 65, 107-108.

pain of falling prey to some sort of tunnel vision. Thirdly, $P(C|T_i) > P(C)$ is true if and only $P(T_i|C) > P(T_i|\neg C)$ is true[7], so we may always use the latter rather than the former if we want to avoid the offending $P(C)$.[8]

4 Relaxation of the Markov Condition

In the previous section we have seen that Cohen formally derived (4) from three conditions, but in fact he did more: he proved that (4) follows from conditions that are considerably weaker that these three, as we will now explain.

Suppose again that each of the two testimonies increases the probability that the defendant is guilty: $P(C|T_i) > P(C)$ with $i = 1, 2$. Now weaken the Markov condition to:

$$\begin{aligned} P(T_1|C \wedge T_2) &\geq P(T_1|C) \\ P(T_1|\neg C \wedge T_2) &\leq P(T_1|\neg C)\,. \end{aligned} \tag{5}$$

Note that the two inequalities go in opposite directions, and also that the original Markov condition corresponds to the special case of (5) in which the inequality signs are replaced by equalities.

[7]On condition of course that all these conditional probabilities are defined, i.e. $0 < P(C) < 1$ and $P(T_i) > 0$.

[8]The question 'When are two witnesses better than one?' is sometimes raised in a different sense than the one that we are considering and that Cohen had in mind. For example, Elliott Sober writes (Sober 2008, 42 — we have adjusted the symbols to agree with our notation):

> When are two witnesses better than one? If the witnesses agree that C is true, and the two witnesses go about their business independently, the two pieces of testimony discriminate more powerfully between C and $\neg C$ than either does by itself, in the sense that
>
> $$\frac{P(T_1 \wedge T_2|C)}{P(T_1 \wedge T_2|\neg C)} > \frac{P(T_i|C)}{P(T_i|\neg C)} > 1 \qquad\qquad i = 1, 2$$

Sober is thus interested in another inequality than Cohen's, namely the inequality in which C and $\neg C$ are to the right, rather than to the left of the bar in $P(\,|\,)$. This inequality is easier to prove, since one does not have to go through Cohen's manipulations to move C to the left of the bar (see the next section and our appendix for the details of Cohen's insight). Sober assumes the Markov condition, but does not notice that his inequality is also valid when this condition is relaxed.

According to Cohen, from (5) it follows that two witnesses are better than one, so we do not need the original strict Markov condition to derive (4). He gives the following example to bolster this intuition. Interpret testimonies T_1 and T_2 as follows:

T_1 : The defendant had a motive for killing the victim.

T_2 : The defendant lacked grief at the victim's death.

and let C be

C : The defendant killed the victim.

The probability that the defendant had a motive, given that he killed the victim and also felt no grief, is greater than the probability that he had a motive, given only that he killed the victim. This is precisely what is implied by the first line of (5). As Cohen phrases it:

> . . . if he was the killer, his lack of grief would confirm the strength of his motive and his motive would confirm that his apparent lack of grief was not due to concealment of his feelings.[9]

The interpretation is also consistent with the second line of (5). The probability that the man had a motive, given that he was innocent and lacked grief is clearly equal to the probability that he had a motive, given that he was innocent. After all, if he is innocent, his lack of grief seems to be irrelevant. Maybe the man is in general not very capable of feeling empathy; or maybe he is empathetic, but does not experience grief because he is overwhelmed and absorbed by other feelings, such as fear for the police and anxiety about the Kafkaesque situation that he finds himself in.

The interpretation thus satisfies (5), and it therefore guarantees (4): given that the man has a motive and lacks grief, the probability that he is guilty of the killing is greater than if he only had a motive or only lacks grief, $P(C|T_1 \wedge T_2) > P(C|T_i)$ where $i = 1, 2$. In the words of Cohen:

> . . . a man's having a motive for killing the victim and his lack of grief at the victim's death could converge to raise the probability of his being the killer, even though either fact would increase the probability of the other.[10]

[9] Cohen 1976, 74.
[10] Cohen 1976, 74.

In the above example of Cohen, (5) is satisfied because the first line has a 'greater than' ($>$) whereas the second line uses an equality sign ($=$). Here is an example in which it is the other way around.[11]

C : Kevin is employed by a criminal syndicate as a lover boy.

T_1 : Kevin showers Mary with loving attention and buys her presents, many of them obtainable only on the black market.

T_2 : Kevin often leaves Mary alone with unsavoury men, who take sexual advantage of her.

The probability that Kevin is employed as a lover boy is increased by either of his actions T_1 or T_2: i.e. $P(C|T_1) > P(C)$ and $P(C|T_2) > P(C)$; but in the absence of the information about Kevin's employment, it seems clear that T_1 and T_2 are in tension with each other. They are not mutually incompatible; but there is certainly a negative correlation between them, which is expressed by $P(T_1|T_2) < P(T_1)$. However, $P(T_1|C \wedge T_2) = P(T_1|C)$, since adding the information about the unsavoury men to the knowledge that Kevin is a lover boy does not alter the probability that T_1 is true. On the other hand, conditioning on information that Kevin is *not* employed as a lover boy would not remove the tension between T_1 and T_2, so we would expect $P(T_1|\neg C \wedge T_2) < P(T_1|\neg C)$. So this example satisfies (5) because it corresponds to the situation in which the sign in the first line is $=$ and in the second one is $<$.

Both examples above imply a violation of the original Markov constraint.[12] They help us in developing the intuition that (4) follows from (5), but of course they do not constitute a proof. In his paper Cohen has successfully demonstrated that the above weakening of the Markov condition suffices to ensure that two witnesses are better than one. That is, he proved that (4) can be derived from (5), together with the assumption that the testimonies are incriminating, $P(C|T_i) > P(C)$ with $i = 1, 2$. Although very ingenious, Cohen's proof is long, taking up more than three pages, and also a bit intricate. In our appendix we offer a shortened and more transparent proof.[13]

[11] This example is not one of Cohen's.

[12] The second example illustrates that especially the second leg of the Markov condition — independence conditional on the negation of C — is questionable, and may be downright inappropriate (Author's publication 2013).

[13] After having completed this paper, we discovered the paper of Carl G. Wagner, in which the author likewise complains that Cohen's proof "runs to three pages" (Wagner 2013, 1460). Wagner gives a shorter proof that in part is similar to ours.

5 Necessary and Sufficient Conditions

As was first noted by O'Neill, Cohen's condition (5) is only a sufficient condition, not a necessary one, that two testimonies are better than one.[14] He points out that

$$P(T_1|C \wedge T_2) > P(T_1|\neg C \wedge T_2) \tag{6}$$

is a *necessary and sufficient* condition that

$$P(C|T_1 \wedge T_2) > P(C|T_2)\,. \tag{7}$$

In order to obtain also $P(C|T_1 \wedge T_2) > P(C|T_1)$, we would need $P(T_2|C \wedge T_1) > P(T_2|\neg C \wedge T_1)$ too. If we spell this out in full we find that

$$\begin{aligned} P(T_1|C \wedge T_2) &> P(T_1|\neg C \wedge T_2) \\ P(T_2|C \wedge T_1) &> P(T_2|\neg C \wedge T_1) \end{aligned} \tag{8}$$

is necessary and sufficient for inequality (4), the statement that two witnesses are better than one:

$$P(C|T_1 \wedge T_2) \quad > \quad P(C|T_i) \qquad\quad i = 1, 2\,.$$

In reply to O'Neill, Cohen grants that formally (8) is necessary and sufficient for (4), but he stresses that in forensic contexts (5) will still be indispensable.[15] Carl Wagner however gives an argument based on second-order probabilities to the effect that Cohen's (5) should not be regarded as a necessary condition.[16] Moreover, George N. Schlesinger has presented examples in which one or other of Cohen's inequalities (5) would be violated, but in which (8) is respected. One such example (in our notation) is the following.[17]

Suppose that C is the contention that a patient is strongly addicted to alcohol, let T_2 be the claim that he has gulped down a jigger of denatured alcohol, and T_1 that he has immediately afterwards gulped down another jigger containing a similar substance. Given the aversion that the denaturing of alcohol engenders, we would expect $P(T_1|C \wedge T_2) < P(T_1|C)$, that is, the probability that the alcoholic patient drinks another jigger (T_1), given that he has already drunk one (T_2), and therefore has a feeling of nausea, is less

[14] O'Neill 1982.
[15] Cohen 1982.
[16] Wagner 2013.
[17] Schlesinger 1991, 151.

than the probability that he drinks the second jigger (T_1), given only that he is strongly addicted to alcohol (C). This inequality contradicts Cohen's condition (5) in that it is inconsistent with its first line. Nevertheless it is still true that it is more likely that the patient is strongly addicted to alcohol if he drinks both jiggers than if he only drinks one, i.e. $P(C|T_1 \wedge T_2) > P(C|T_2)$.[18]

We do not want to go into all the details of Cohen's proof. For our purposes it is enough to say that he goes from (5) to (4) via a formula that is equivalent to our (8).[19] Cohen first proves that (5) is sufficient for (8) and then that the latter suffices for (4). In schema:

$$(5) \longrightarrow (8) \longrightarrow (4)\,.$$

O'Neill and Schlesinger however note that (8) and (4) are equivalent:

$$(5) \longrightarrow (8) \longleftrightarrow (4)\,.$$

Since (5) can be false while (8) is true (an example is given by the above story about the denatured alcohol), one might start the derivation with (8) rather than (5), and this is precisely what both O'Neill and Schlesinger do. In this sense their proofs rely on that of Cohen: they pick out, as it were, the (8)-(4) part only of Cohen's (5)-(8)-(4) reasoning.

A new and useful alternative to condition (8) is afforded by the insight that (8) is true if and only if the following inequality holds:

$$P(T_1|C \wedge T_2) > P(T_1|T_2)\,. \qquad (9)$$

Since (9) is also a necessary and sufficient condition for the validity of (4), we have a new schema, which does not use Cohen's (8):

$$(5) \longrightarrow (9) \longleftrightarrow (4)\,.$$

The technical justification of all these claims can be found in the appendix. As we will argue in the next section, in practice our (9) is sometimes preferable to (8).

[18]Ibid. We have replaced Schlesinger's D by C, his E_1 is our T_2, and his E_2 is our T_1; we switched the indices 1 and 2 in order to discuss the example more easily. Three decades after O'Neill and Schlesinger proved that (8) is a necessary and sufficient condition for (4), Michael Huemer derived the same result, apparently unaware of their achievements (Huemer 2007). The fact that he has not seen Schlesinger's attempt to give examples might explain why Huemer is unduly pessimistic about the practical utility of the derivation.

[19]The formula in question is inequality (40) on p. 77 of Cohen 1976, and inequality (39) on p. 106 of Cohen 1977.

The probability distribution depicted in Figure 2 illustrates a situation where Cohen's condition (5) fails, but (8) as well as our (9) applies. The following inequalities hold:

$$\begin{aligned}
P(T_1|C \wedge T_2) = \tfrac{4}{5} &\quad < \quad \tfrac{5}{6} = P(T_1|C) \\
P(T_1|\neg C \wedge T_2) = \tfrac{3}{5} &\quad > \quad \tfrac{1}{2} = P(T_1|\neg C) \\
P(T_1|C \wedge T_2) = \tfrac{4}{5} &\quad > \quad \tfrac{3}{5} = P(T_1|\neg C \wedge T_2) \\
P(T_1|C \wedge T_2) = \tfrac{4}{5} &\quad > \quad \tfrac{7}{10} = P(T_1|T_2) \\
P(C|T_2) = \tfrac{1}{2} &\quad > \quad \tfrac{3}{8} = P(C) \\
P(C|T_1 \wedge T_2) = \tfrac{4}{7} &\quad > \quad \tfrac{1}{2} = P(C|T_2)\,.
\end{aligned}$$

From the first two inequalities we see that both conditions of the weakened Markov condition (5) are violated; while the third and fourth show that the O'Neill-Schlesinger condition (8) and our new condition (9) are satisfied. The fifth inequality shows that the second testimony raises the probability of C above what it originally was; and the sixth indicates that the two testimonies increase the probability even more.

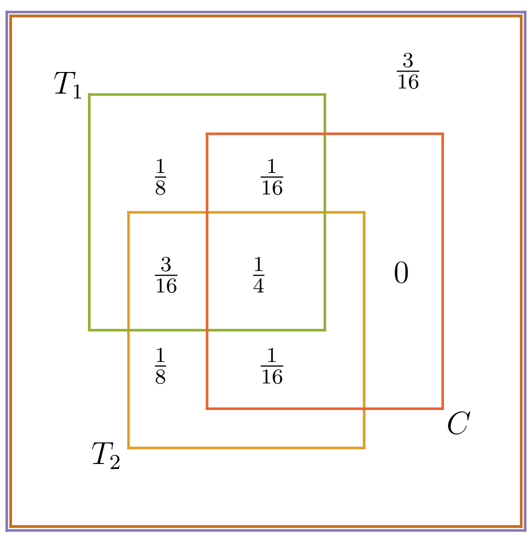

Figure 2: Violation of Cohen's Condition

6 Concluding Remarks

We have seen that under the Markov condition two witnesses are better than one; and we recalled L. Jonathan Cohen's proof that even a weakened version the Markov requirement is enough to draw that conclusion. Cohen's proof is however long and somewhat impenetrable, and we have replaced it by a simpler and more transparant one.

Like the proof of Cohen, ours gives only a sufficient and not a necessary condition for the conclusion that two witnesses are to be preferred. L.J. O'Neill succeeded in finding a condition that is both necessary and sufficient, and Schlesinger illustrated it by giving some real life examples. We showed that their reasoning is actually dependent on Cohen's, in the sense that it uses the second half of Cohen's argument; and we have offered a necessary and sufficient condition which is different from Cohen's reasoning.

Here is how Alice's trial of Section 2 is analyzed when our condition is used. Recall the testimonies by Clara and Deanna respectively:

T_1 : Alice had a gun in her purse on the night of the crime.

T_2 : Alice had an argument with Bob on the night of the crime.

$P(T_1|T_2)$ is the probability that Clara's testimony was trustworthy, on condition that there was indeed a connubial altercation. In the first instance we supposed that the testimonies of Clara and Deanna were independent of one another: $P(T_1|T_2) = P(T_1)$. In many contexts this seems a natural assumption to make. After all, Alice may simply have been exercising her second amendment rights, and then having a gun in her purse had nothing to do with her quarrel with Bob. Independence of the two testimonies is however neither necessary nor sufficient for the conclusion that two testimonies are better than one would have been.

Often $P(T_1|T_2)$ is not equal to $P(T_1)$. Alice may have taken the gun out of its box with the explicit purpose of effecting the quietus of her quarrelsome spouse; and if we add the hypothesis that Alice actually did shoot Bob to the condition that Deanna told the truth, the probability that Clara's testimony was accurate will surely be increased:

$$P(T_1|C \wedge T_2) > P(T_1|T_2)\,.$$

As we have seen, this inequality is a necessary and sufficient condition that the two testimonies are superior to one.

We believe there are scenarios where our necessary and sufficient condition fares better than that of O'Neill or Schlesinger. For example, imagine a special sale of 100 television sets is announced in a downtown store, but due to a miscommunication between the employees, nobody checked whether the customers actually paid for what they took home. Consider the following propositions:

C : 99 of the 100 television sets carried out of this store have been stolen.

T_1 : Jones has not paid for a television set.

T_2 : Jones carried a television set out of this store.

The conditional probability that Jones has not paid anything, given that 99 of the 100 television sets were stolen, and that he was seen leaving the store with a television set, namely $P(T_1|C \wedge T_2)$, is at least 0.99.[20] On the other hand, the conditional probability that Jones has not paid anything, given merely that he was seen leaving the store with a television set, $P(T_1|T_2)$ (without the information that 99 sets have been stolen), is certainly considerably less than 0.99. Thus $P(T_1|C \wedge T_2) > P(T_1|T_2)$, which is our necessary and sufficient condition for $P(C|T_1 \wedge T_2) > P(C|T_2)$ to hold. For the O'Neill-Schlesinger inequality we would need to look at $P(T_1|\neg C \wedge T_2)$, the conditional probability that he has not paid anything, given that it is not the case that 99 of the 100 sets were stolen, and that he was seen leaving the store with a television set. This is clearly an awkward condition: does it allow perhaps that 98 sets were stolen, or maybe even all 100? Our condition, in which we do not need to consider what $\neg C$ implies, is more straightforward in this case. It is of course true that if our inequality is satisfied, so is the one of O'Neill and Schlesinger, and *vice versa*: it is simply a matter of which is the more convenient.

[20] $P(T_1|C \wedge T_2) = \frac{P(T_1 \wedge T_2|C)}{P(T_2|C)} = \frac{0.99}{P(T_2|C)} \geq 0.99.$

Appendix

Theorem

If the following conditions are satisfied

$$\begin{aligned} P(C|T_i) &> P(C) \qquad i=1,2\,, \qquad (10) \\ P(T_1|C\wedge T_2) &\geq P(T_1|C) \qquad (11) \\ P(T_1|\neg C\wedge T_2) &\leq P(T_1|\neg C)\,, \qquad (12) \end{aligned}$$

then

$$P(C|T_1\wedge T_2) > P(C|T_i) \qquad i=1,2\,. \qquad (13)$$

The probabilities are assumed to be *regular*, i.e. the values 0 and 1 are excluded.

Proof

$$\begin{aligned} P(T_1|C)-P(T_1|\neg C) &= \frac{P(T_1\wedge C)-P(T_1)P(C)}{P(C)P(\neg C)} \\ &= \frac{[P(C|T_1)-P(C)]P(T_1)}{P(C)P(\neg C)}\,. \end{aligned}$$

From (10) we see that $P(C|T_1)-P(C)$ is positive, and it therefore follows from the above equation that $P(T_1|C)-P(T_1|\neg C)$ is positive. In other words,

$$P(T_1|C) > P(T_1|\neg C)\,;$$

and from (11) and (12) we conclude that

$$P(T_1|C\wedge T_2) \geq P(T_1|C) > P(T_1|\neg C) \geq P(T_1|\neg C\wedge T_2)\,, \qquad (14)$$

so that

$$\frac{P(T_1\wedge C\wedge T_2)}{P(C\wedge T_2)} > \frac{P(T_1\wedge\neg C\wedge T_2)}{P(\neg C\wedge T_2)}\,. \qquad (15)$$

After cross-multiplication and use of total probability, we find

$$\begin{aligned} &P(T_1\wedge C\wedge T_2)[P(T_2)-P(C\wedge T_2)] \\ &\qquad > P(C\wedge T_2)[P(T_1\wedge T_2)-P(T_1\wedge C\wedge T_2)]\,. \qquad (16) \end{aligned}$$

Multiplying the parentheses out, we see that two terms cancel, leaving

$$P(T_1 \wedge C \wedge T_2)P(T_2) \quad > \quad P(C \wedge T_2)P(T_1 \wedge T_2)\,. \tag{17}$$

In view of the regularity condition, we may divide both sides freely by probabilities, obtaining two useful expressions:

$$\frac{P(T_1 \wedge C \wedge T_2)}{P(C \wedge T_2)} \quad > \quad \frac{P(T_1 \wedge T_2)}{P(T_2)}\,, \tag{18}$$

and also

$$\frac{P(T_1 \wedge C \wedge T_2)}{P(T_1 \wedge T_2)} \quad > \quad \frac{P(C \wedge T_2)}{P(T_2)}\,.$$

The latter inequality implies that $P(C|T_1 \wedge T_2) > P(C|T_2)$, which is one half of what we wanted to demonstrate. We prove $P(C|T_1 \wedge T_2) > P(C|T_1)$ by interchanging T_1 and T_2 in the whole demonstration.

Corollary
A particular case of conditions (11) and (12) is obtained by requiring independence of the two testimonies, conditional on C and on $\neg C$, which is the classic Markov constraint:

$$P(T_1|C \wedge T_2) = P(T_1|C) \qquad \text{and} \qquad P(T_1|\neg C \wedge T_2) = P(T_1|\neg C).$$

The conclusion (13) is the same.

Comment
From (14) we have

$$P(T_1|C \wedge T_2) \; > \; P(T_1|\neg C \wedge T_2)\,, \tag{19}$$

this being the Schlesinger constraint; and the conditions (11) and (12) can be replaced by (19). From (18) we see that

$$P(T_1|C \wedge T_2) \; > \; P(T_1|T_2)\,. \tag{20}$$

This inequality holds if and only if (19) does, and it constitutes a *necessary and sufficient* condition for $P(C|T_1 \wedge T_2) > P(C|T_2)$ to hold. This is clear, since if (20) does *not* hold, i.e. if

$$P(T_1|C \wedge T_2) \; \leq \; P(T_1|T_2)\,,$$

then

$$\frac{P(T_1 \wedge C \wedge T_2)}{P(C \wedge T_2)} \leq \frac{P(T_1 \wedge T_2)}{P(T_2)},$$

so

$$\frac{P(T_1 \wedge C \wedge T_2)}{P(T_1 \wedge T_2)} \leq \frac{P(C \wedge T_2)}{P(T_2)},$$

which means that $P(C|T_1 \wedge T_2) \leq P(C|T_2)$. In other words, if (20) does not hold, then it is not true that $P(C|T_1 \wedge T_2) > P(C|T_2)$.

We conclude that $P(T_1|C \wedge T_2) > P(T_1|T_2)$ is a necessary and sufficient condition for the validity of $P(C|T_1 \wedge T_2) > P(C|T_2)$. By interchanging T_1 and T_2, we see also that $P(T_2|C \wedge T_1) \leq P(T_2|T_1)$ is a necessary and sufficient condition for the validity of $P(C|T_1 \wedge T_2) > P(C|T_1)$.

References

Author's publication. 2013.

Cohen, L. Jonathan. 1976. How Can One Testimony Corroborate Another? in *Essays in Memory of Imre Lakatos*, eds. R.S. Cohen, P.K. Feyerabend, and M.W. Wartofsky, 65-78. Dordrecht/Boston: D. Reidel Publishing Company.

Cohen, L. Jonathan. 1977. *The Probable and the Provable.* Oxford: Oxford University Press.

Cohen, L. Jonathan. 1982. Corroborating Testimonies. *British Journal for the Philosophy of Science* 33: 161-164.

Huemer, Michael. 2007. Weak Bayesian Coherentism. *Synthese* 157: 337-346.

Olsson, Erik. Coherentist Theories of Epistemic Justification. *The Stanford Encyclopedia of Philosophy.* Spring 2017 Edition. Edward N. Zalta (ed.) URL: <https://plato.stanford.edu/archives/spr2017/entries/justep-coherence/>.

O'Neill, L.J. 1982. Corroborating Testimonies. *British Journal for the Philosophy of Science* 33: 60-63.

Pearl, Judea. 2000. *Causality.* Cambridge: Cambridge University Press.

Reichenbach, Hans. 1956. *The Direction of Time*. Berkeley: University of California Press.

Schlesinger, George N. 1991. *The Sweep of Probability.* Notre Dame: University of Notre Dame Press.

Sober, Elliott. 2008. *Evidence and Evolution. The Logic Behind the Science.* Cambridge: Cambridge University Press.

Wagner, Carl G. 2013. The Corroboration Paradox. *Synthese* 190: 1455-1469.