Dwindling Confirmation

William Roche and Tomoji Shogenji*†

We show that as a chain of confirmation becomes longer, confirmation dwindles under screening-off. For example, if E confirms H1, H1 confirms H2, and H1 screens off E from H2, then the degree to which E confirms H2 is less than the degree to which E confirms H1. Although there are many measures of confirmation, our result holds on any measure that satisfies the Weak Law of Likelihood. We apply our result to testimony cases, relate it to the Data-Processing Inequality in information theory, and extend it in two respects so that it covers a broader range of cases.

1. Introduction. Think of a series of propositions E, H1, H2, . . . , H*n* such that E supports H1, H1 supports H2, . . . , and H*n*-1 supports H*n*. If the support relation is logical entailment, then, regardless of the length of the series, E supports H*n* since logical entailment is transitive. We cannot say the same though if the support relation is *confirmation* in the incremental sense (hereafter simply "confirmation"), where, for any two propositions E and H, E confirms H just in case E increases the probability of H (i.e., $\Pr(H|E) > \Pr(H)$). Even if each member (except the last) supports the next member in the series, E might not support H*n*.[1] This is because confirma-

Received April 2013; revised September 2013.

*To contact the authors, please write to: William Roche, Department of Philosophy, Texas Christian University, Fort Worth, TX; e-mail: w.roche@tcu.edu. Tomoji Shogenji, Department of Philosophy, Rhode Island College, Providence, RI; e-mail: tshogenji@ric.edu.

†We wish to thank Elliott Sober and Le Zhang for helpful discussion and three anonymous reviewers for detailed and extremely helpful comments.

1. A few comments are in order here. First, here and throughout the article we suppress reference to background information. We also assume that E and H are contingent and that the probability function Pr is regular so that $\Pr(H|E) = 1$ and $\Pr(E \wedge \neg H) = 0$ if and only if (iff) E logically entails H. Second, confirmation in the incremental sense is equivalent to confirmation in the "relevance" sense: E incrementally confirms H just in case E is positively relevant to H. Put formally, $\Pr(H|E) > \Pr(H)$ just in case $\Pr(H \wedge E) > \Pr(H)\Pr(E)$. Third, and finally, confirmation in the incremental sense is neither sufficient nor necessary for confirmation in the "absolute" sense, where E confirms H iff the proba-

Philosophy of Science, 81 (January 2014) pp. 114–137. 0031-8248/2014/8101-0007$10.00
Copyright 2014 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

tion, unlike logical entailment, is nontransitive. To illustrate the point, suppose a card is randomly drawn from a standard deck of cards, where E = the card drawn is a heart, H1 = the card drawn is a red jack, and H2 = the card drawn is a diamond. Proposition E confirms H1 since $\Pr(H1|E) = 1/13 > 1/26 = \Pr(H1)$, and H1 confirms H2, as $\Pr(H2|H1) = 1/2 > 1/4 = \Pr(H2)$, but E does not confirm H2 because $\Pr(H2|E) = 0 < 1/4 = \Pr(H2)$. It is somewhat troubling that confirmation is nontransitive since we often construct a chain of reasoning even when the support relation is only confirmation. In fact, such reasoning seems unproblematic in many cases. Take a case of testimony: E = Smith said that the card drawn is a red, where H1 = the card drawn is a red and H2 = the card drawn is a heart. Proposition E confirms H1, at least on certain ways of filling in the details; H1 confirms H2 since $\Pr(H2|H1) = 1/2 > 1/4 = \Pr(H2)$; and it seems clear that E confirms H2.

A second issue of interest, regarding a series E, H1, H2, . . . , H*n* where each member (except the last) confirms the next, concerns *degree of confirmation*. Suppose, unlike in the first card case above, E confirms each of the other members in the series. Intuitively, at least in many cases of this sort, confirmation *dwindles* in that the degree to which E confirms H*n* is less than the degree to which it confirms H*n*-1, the degree to which it confirms H*n*-1 is less than the degree to which it confirms H*n*-2, and so on.[2] Here is an example: E = Smith said that miracle M occurred, where H1 = miracle M occurred and H2 = God exists. Proposition E confirms H1, at least on certain ways of filling in the details; H1 in turn confirms H2; and, it seems, E confirms H2 (although the probability of H2 given E need not be high and, in fact, might be quite low). But, it seems, the degree to which E confirms H2 is less than the degree to which E confirms H1. The sense of dwindling confirmation becomes stronger as the chain gets longer, for example, if E = Jones said that Smith said that miracle M occurred, H1 = Smith said that miracle M occurred, H2 = miracle M occurred, and H3 = God exists. It turns out, however, that confirmation does not always dwindle. Returning to the card case, suppose E = the card drawn is a face card or the ace of hearts, H1 = the card drawn is a heart, and H2 = the card drawn is a king, a queen, or the ace of hearts. It is hard to deny that the degree to which E confirms H2 is greater than the degree to which E confirms H1, for E only slightly raises

bility of H given E is sufficiently high. See Carnap (1962, preface) on "concepts of increase in firmness" and "concepts of firmness." See also Roche and Shogenji (2013) for discussion of yet further notions of confirmation.

2. This claim about degree of confirmation is distinct from Plantinga's claim of "dwindling probabilities" (2000, 2006). The latter (put in terms of E, H1, . . . , H*n*) concerns just the probabilities $\Pr(H1|E), \Pr(H2|E), \ldots, \Pr(Hn|E)$ and not degrees of confirmation (i.e., incremental confirmation). See McGrew (2004) and Swinburne (2004) for forceful objections to Plantinga's argument for dwindling probabilities.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

the probability of H1 (from 1/4 to 4/13) but substantially raises the probability of H2 (from 9/52 to 9/13). Another point (not unrelated to the first) in favor of this judgment is that $\Pr(H2) < \Pr(H1)$, while $\Pr(H2|E) > \Pr(H1|E)$; in other words, the probability of H2 is initially lower than that of H1, but it ends up higher given E, so that the increase in probability is greater for H2 than for H1.[3]

It would be welcome, then, if there were conditions sufficient for transitivity in confirmation and conditions sufficient for dwindling confirmation. The former conditions could be used to verify in a given case, where E confirms H1, H1 confirms H2, . . . , and Hn-1 confirms Hn, that E confirms Hn. The latter could be used to verify that, even when E confirms Hn, the degree to which E confirms Hn is less than the degree to which E confirms Hn-1.

It turns out, fortunately, that confirmation is transitive under the following condition:

> **Screening-Off Condition (SOC).** $\Pr(Hk|Hk\text{-}1 \wedge E) = \Pr(Hk|Hk\text{-}1)$ and $\Pr(Hk|\neg Hk\text{-}1 \wedge E) = \Pr(Hk|\neg Hk\text{-}1)$.[4]

SOC says in effect that Hk-1 screens off E from Hk in that given the truth or falsity of Hk-1, E has no impact on the probability of Hk.[5] In other words, E affects the probability of Hk only indirectly through its impact on the probability of Hk-1. SOC is a condition sufficient for transitivity in confirmation in that if $\Pr(Hk\text{-}1|E) > \Pr(Hk\text{-}1)$, $\Pr(Hk|Hk\text{-}1) > \Pr(Hk)$, and SOC holds, then $\Pr(Hk|E) > \Pr(Hk)$.

The main question we aim to answer is whether SOC is also a condition sufficient for dwindling confirmation. We argue in the affirmative. The article proceeds as follows. In section 2, we provide a list of the main confirmation measures in the literature (13 in total) and introduce an adequacy condition on confirmation measures, which is called the Weak Law of Likelihood (WLL) and which is satisfied by each of the measures in the list. Then, in section 3, we show that confirmation dwindles under SOC on any confirmation measure meeting WLL, hence on any adequate confirmation measure. This is the main result of the article.[6] We apply our result to testi-

3. It should be noted too that each of the main confirmation measures discussed in the literature, and listed below in sec. 2, implies that the degree to which E confirms H2 is indeed greater than the degree to which E confirms H1.

4. See Shogenji (2003). See also Sober (2009) for an equivalent result and applications.

5. SOC does not hold in the first card case above since $\Pr(H2|H1 \wedge E) = 0 < 1/2 = \Pr(H2|H1)$ and $\Pr(H2|\neg H1 \wedge E) = 0 < 6/25 = \Pr(H2|\neg H1)$.

6. Douven (2011) presents counterexamples to nondwindling under SOC; i.e., he shows that on three measures—the difference measure, the log-likelihood measure, and the

https://doi.org/10.1086/674346 Published online by Cambridge University Press

mony cases of the sort given above (where the subject uses information obtained from testimony for further inference) and relate our result to the Data-Processing Inequality in information theory. In section 4, we extend our result in two respects. Finally, in section 5, we conclude.

2. Preliminaries. It is a matter of controversy among Bayesian confirmation theorists how exactly to measure confirmation. The main confirmation measures in the literature are as follows:[7]

$$C(H, E) = Pr(H \wedge E) - Pr(H)Pr(E); \tag{1}$$

$$D(H, E) = Pr(H|E) - Pr(H); \tag{2}$$

$$G(H, E) = \frac{Pr(\neg H)}{Pr(\neg H|E)}; \tag{3}$$

$$J(H, E) = \frac{\log[Pr(H|E)] - \log[Pr(H)]}{-\log[Pr(H)]}; \tag{4}$$

$$K(H, E) = \frac{Pr(E|H) - Pr(E|\neg H)}{Pr\,(E|H) + Pr(E|\neg H)}; \tag{5}$$

Kemeny-Oppenheim measure—regardless of the minimum degree required for substantial confirmation, there are probability distributions on which the degree to which E confirms H1 is substantial, the degree to which H1 confirms H2 is substantial, and SOC holds, and yet the degree to which E confirms H2 is not substantial. For related discussion, see Roche (2012a). Our result is different in three main respects. First, our result is not limited to cases where E substantially confirms H1 and H1 substantially confirms H2. Second, our result concerns all confirmation measures meeting WLL and thus all of the main confirmation measures from the literature. Third, our result goes beyond mere counterexamples to nondwindling. We establish the positive universal claim that, when a certain condition concerning logical entailment holds, confirmation invariably dwindles under SOC.

7. For discussion and references regarding measures (1)–(3) and (5)–(13), see Festa (1999), Eells and Fitelson (2002), and Crupi, Tentori, and Gonzalez (2007). Measure J is given in Shogenji (2012). See also Tornebohm (1966).

https://doi.org/10.1086/674346 Published online by Cambridge University Press

$$L(H, E) = \frac{Pr(E|H)}{Pr(E|\neg H)}; \tag{6}$$

$$L^*(H, E) = \log\left[\frac{Pr(E|H)}{Pr(E|\neg H)}\right]; \tag{7}$$

$$M(H, E) = Pr\,(E|H) - Pr(E); \tag{8}$$

$$N(H, E) = Pr\,(E|H) - Pr(E|\neg H); \tag{9}$$

$$R(H, E) = \frac{Pr(H|E)}{Pr(H)}; \tag{10}$$

$$R^*(H, E) = \log\left[\frac{Pr(H|E)}{Pr(H)}\right]; \tag{11}$$

$$S(H, E) = Pr(H|E) - Pr(H|\neg E); \tag{12}$$

$$Z(H, E) = \begin{cases} \dfrac{Pr(H|E) - Pr(H)}{1 - Pr(H)} & \text{if } Pr(H|E) \geq Pr(H) \\ \dfrac{Pr(H|E) - Pr(H)}{Pr(H)} & \text{otherwise} \end{cases}. \tag{13}$$

Some of these measures are ordinally equivalent to each other: they impose the same ordering in degrees of confirmation on any two ordered pairs of propositions $\langle H, E\rangle$ and $\langle H', E'\rangle$.[8] For example, K, L, and L* are ordinally equivalent to each other. The measures taken as a group, however, are motley.

8. Put formally, measures X and X* are ordinally equivalent to each other just in case for any two ordered pairs of propositions $\langle H, E\rangle$ and $\langle H', E'\rangle$, $X(H, E) \leq X(H', E')$ iff $X^*(H, E) \leq X^*(H', E')$. Note that it follows by contraposition that $X(H, E) > X(H', E')$ iff $X^*(H, E) > X^*(H', E')$.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Consider now the following condition on any adequate confirmation measure X:

> **Weak Law of Likelihood (WLL).** If (a) $\Pr(E|H) > \Pr(E|H^*)$ and (b) $\Pr(E|\neg H) < \Pr(E|\neg H^*)$, then $X(H, E) > X(H^*, E)$.

We argue in the next section that, under certain conditions, confirmation dwindles under SOC on all measures meeting WLL. First, though, some comments are in order regarding WLL.

Bayesian confirmation theorists, while disagreeing about how exactly to measure confirmation, all agree (as far as we are aware) that any adequate confirmation measure should satisfy WLL.[9] This is manifest in the fact that each of measures (1)–(13) meets WLL. This point is easiest to see with respect to measures (1)–(3) and (5)–(13). If $\Pr(E|\neg H) < \Pr(E|\neg H^*)$, then

$$G(H, E) = \frac{\Pr(\neg H)}{\Pr(\neg H|E)} = \frac{\Pr(E)}{\Pr(E|\neg H)} > \frac{\Pr(E)}{\Pr(E|\neg H^*)} = \frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} = G(H^*, E).$$

So, G meets WLL. Similarly, if $\Pr(E|H) > \Pr(E|H^*)$, then

$$R(H, E) = \frac{\Pr(H|E)}{\Pr(H)} = \frac{\Pr(E|H)}{\Pr(E)} > \frac{\Pr(E|H^*)}{\Pr(E)} = \frac{\Pr(H^*|E)}{\Pr(H^*)} = R(H^*, E).$$

Thus, R too meets WLL. To see that the same is true of measures (1), (2), (5)–(9), and (11)–(13), consider the following equalities:

$$C(H, E) = \frac{1}{(1/\{1 - [1/G(H, E)]\}) + \{1/[R(H, E) - 1]\}} \Pr(E);$$

$$D(H, E) = \frac{1}{(1/\{1 - [1/G(H, E)]\}) + \{1/[R(H, E) - 1]\}};$$

9. Bayesian confirmation theorists also all agree on a second adequacy condition on confirmation measures: there is a real number t such that (a) $X(H, E) > t$ iff $\Pr(H|E) > \Pr(H)$, (b) $X(H, E) = t$ iff $\Pr(H|E) = \Pr(H)$, and (c) $X(H, E) < t$ iff $\Pr(H|E) < \Pr(H)$. This condition, called Qualitative Discrimination in Crupi, Festa, and Buttasi (2010), says in effect that X discriminates among cases of confirmation, cases of neutrality (where E neither confirms nor disconfirms H), and cases of disconfirmation, by assigning values above t to cases of confirmation, t to cases of neutrality, and values below t to cases of disconfirmation.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

$$K(H, E) = \frac{R(H, E)G(H, E) - 1}{R(H, E)G(H, E) + 1};$$

$$L(H, E) = R(H, E)G(H, E);$$

$$L^*(H, E) = \log[R(H, E)G(H, E)];$$

$$M(H, E) = \Pr(E)[R(H, E) - 1];$$

$$N(H, E) = \left[R(H, E) - \frac{1}{G(H, E)}\right]\Pr(E);$$

$$R^*(H, E) = \log[R(H, E)];$$

$$S(H, E) = \frac{1/[(1/\{1 - [1/G(H, E)]\}) + \{1/[R(H, E) - 1]\}]}{1 - \Pr(E)};$$

$$Z(H, E) = \begin{cases} 1 - \dfrac{1}{G(H, E)} & \text{if } \Pr(H|E) \geq \Pr(H) \\ R(H, E) - 1 & \text{otherwise} \end{cases}.$$

Take the first equality: $C(H, E)$ is fully determined by G, R, and $\Pr(E)$ and is a monotonically increasing function of G and R. So, since $\Pr(E)$ is a constant in the inequalities in WLL, and since $G(H, E)$ and $R(H, E)$ increase, respectively, as $\Pr(E|\neg H)$ decreases and $\Pr(E|H)$ increases, it follows that $C(H, E)$ also increases as $\Pr(E|\neg H)$ decreases and $\Pr(E|H)$ increases. But that is just what satisfaction of WLL demands. So C satisfies WLL. Similar remarks can be made with respect to measures (2), (5)–(9), and (11)–(13). Measure J (as far as we are aware) is not "reducible" to G, R, and $\Pr(E)$. But it too meets WLL (see app. A for proof).

WLL is similar to the "weak likelihood principle," which is endorsed by Joyce (2008) and which is equivalent to

> **WLL*.** If (*a*) $\Pr(E|H) > \Pr(E|H^*)$ and (*b*) $\Pr(E|\neg H) \leq \Pr(E|\neg H^*)$, then $X(H, E) > X(H^*, E)$.[10]

10. Joyce (2008, sec. 3) also gives a second construal of the weak likelihood principle (on which the principle is stronger than on the first construal). It is equivalent to this: if (*a*) $\Pr(E|H) \geq \Pr(E|H^*)$, (*b*) $\Pr(E|\neg H) \leq \Pr(E|\neg H^*)$, and (*c*) one of the two inequalities is strict, then $X(H, E) > X(H^*, E)$ and $X(\neg H, \neg E) > X(\neg H^*, \neg E)$. Each of these likelihood conditions—WLL and WLL* on its two construals—is much weaker than the Law of Likelihood (LL): if $\Pr(E|H) > \Pr(E|H^*)$, then $X(H, E) > X(H^*, E)$. See also Fitelson (2007, 2011), Crupi et al. (2010), Sober (2011), and Brossel (2013).

https://doi.org/10.1086/674346 Published online by Cambridge University Press

In defense of WLL*, Joyce writes:

> (2.1e) [i.e., the weak likelihood principle] captures one core message of Bayes' Theorem for theories of confirmation. Let's say that H is *uniformly better* than H^* as predictor of E's truth-value when (a) H predicts E more strongly than H^* does, and (b) $\sim H$ predicts $\sim E$ more strongly than $\sim H^*$ does. According to the weak likelihood principle, hypotheses that are uniformly better predictors of the data are better supported by the data. For example, the fact that little Johnny is a Christian is better evidence for thinking that his parents are Christian than for thinking that they are Hindu because (a) a far higher proportion of Christian parents than Hindu have Christian children, and (b) a far higher proportion of non-Christian parents than non-Hindu parents have non-Christian children.
>
> In fact, the weak likelihood principle (2.1e) encapsulates a minimal form of Bayesianism to which all parties can agree. . . . Indeed, the weak likelihood principle must be an integral part of any account of evidential relevance that deserves the title "Bayesian." To deny it is to misunderstand the central message of Bayes' Theorem for questions of evidence: namely, that hypotheses are confirmed by the data they predict. (2008, sec. 3)

WLL* is slightly stronger than WLL, as one of the inequalities in its antecedent is weakened slightly from "less than" in *b* of WLL to "less than or equal to" in *b* of WLL*. This fact, it turns out, is rather significant for our purposes. Unlike WLL, WLL* is not met by all of measures (1)–(13).[11] So, since not all Bayesian confirmation theorists will accept WLL*, we want to remain neutral on whether confirmation measures ought to satisfy WLL*. However, what Joyce says on behalf of WLL* carries over to WLL. If $\Pr(E|H) > \Pr(E|H^*)$ and $\Pr(E|\neg H) < \Pr(E|\neg H^*)$, then H is "uniformly better than" H* at predicting E. The key insight behind WLL is that if H is uniformly better than H* at predicting E, then E provides more confirmation to H than to H*.

One final comment on WLL is in order. While WLL is a rather undemanding adequacy condition on confirmation measures, it is not entirely toothless. Consider the measure

$$D'(H, E) = \Pr(H|E)\Pr(H|E) - \Pr(H)\Pr(H). \tag{14}$$

11. Measures G and Z do not meet WLL*. To see this in the case of G, suppose $\Pr(E|H) > \Pr(E|H^*)$ and $\Pr(E|\neg H) = \Pr(E|\neg H^*)$. It should be the case by WLL* that $G(H, E) > G(H^*, E)$, but it is not since $G(H, E) = \Pr(\neg H)/\Pr(\neg H|E) = \Pr(E)/\Pr(E|\neg H) = \Pr(E)/\Pr(E|\neg H^*) = \Pr(\neg H^*)/\Pr(\neg H^*|E) = G(H^*, E)$. The case of Z is similar.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Measure D′ is similar in an obvious respect to D, which is a popular measure. But D′ does not meet WLL. So, it is significant that the main confirmation measures in the literature, motivated by various disparate considerations, all satisfy WLL.

The question of dwindling confirmation under SOC can now be put precisely. Let E, H1, H2, . . . , Hn be a series of propositions. Consider the conditions:

A. H1 is confirmed by E, and Hk is confirmed by Hk-1 for all $2 \leq k \leq n$;
B. E is screened off in the sense of SOC from Hk by Hk-1 for all $2 \leq k \leq n$;[12]
C. E neither entails nor is entailed by H1, and Hk-1 neither entails nor is entailed by Hk for all $2 \leq k \leq n$.[13]

The question of dwindling confirmation under SOC is this: Is it the case that if A, B, and C hold and X meets WLL, then $X(Hk, E) < X(Hk\text{-}1, E)$ for all $2 \leq k \leq n$?

3. Main Result. The task now is to answer the question of dwindling confirmation under SOC. We do this in section 3.1. We then turn to testimony cases in section 3.2 and to the Data-Processing Inequality in section 3.3.

3.1. Dwindling Confirmation under SOC. The answer to the question of dwindling confirmation under SOC is affirmative:

> **Theorem 1.** If A, B, and C hold and X meets WLL, then $X(Hk, E) < X(Hk\text{-}1, E)$ for all $2 \leq k \leq n$.

(See app. B for proof.) The key in the proof of theorem 1 is the relations of likelihood that if A, B, and C hold, then (i) $Pr(E|Hk\text{-}1) > Pr(E|Hk)$ and (ii) $Pr(E|\neg Hk\text{-}1) < Pr(E|\neg Hk)$ for all $2 \leq k \leq n$. Theorem 1 follows immediately from i and ii by WLL.

Each of measures (1)–(13) meets WLL. So, theorem 1 entails that confirmation dwindles under SOC on each of measures (1)–(13). It is well known that certain claims about confirmation hold on some measures but

12. The qualification "in the sense of SOC" is needed because we introduce a second screening-off condition in sec. 4.1.

13. Recall the regularity assumption mentioned in n. 1. It follows from C that $Pr(H1|E) < 1$, $Pr(E|H1) < 1$, $Pr(Hk|Hk\text{-}1) < 1$, and $Pr(Hk\text{-}1|Hk) < 1$.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

not on others (Fitelson 1999; Brossel 2013). There is no such "problem of measure sensitivity" with dwindling confirmation under SOC: confirmation dwindles under SOC on all of the main confirmation measures in the literature and, indeed, on any adequate confirmation measure extant or otherwise.

3.2. Testimony Cases. Theorem 1 has application in many cases involving testimony where E = S said that H1, but the proposition of interest is not H1 itself but H2 that is related to it. In many such cases, E neither entails nor is entailed by H1, H1 neither entails nor is entailed by H2, but E confirms H1, which in turn confirms H2. Moreover, and crucially, SOC holds in many such cases: given the truth of H1 (which is the proposition to which S testified) or given the falsity of H1, that S said that H1 has no impact on the probability of H2 (Shogenji 2003). The second testimony case from section 1, where E = Smith said that miracle M occurred, H1 = miracle M occurred, and H2 = God exists, is plausibly such a case. Proposition E increases the probability of H1 (at least on some ways of filling in the details of the case)—even if just barely. Proposition H1 in turn increases the probability of H2, E (we can suppose) neither entails nor is entailed by H1, and H1 neither entails nor is entailed by H2.[14] Further, it seems, SOC holds: E has an impact on the probability of H2 only indirectly through H1. To put this in terms of rational credence, if one already knows whether miracle M occurred, then learning further that Smith said that miracle M occurred has no impact on one's credence in God's existence. On this construal of the case, theorem 1 applies so that E confirms the existence of God less than it does the occurrence of miracle M.

Things are much the same in cases involving *higher-order* testimonies. For example, the first member of the series is E = S1 said that S2 said that H2, the second member is H1 = S2 said that H2, the third member is H2 itself, and the fourth member is some proposition H3 related to H2. In many such cases SOC still holds at each step, so that theorem 1 applies if other conditions are also met. Since confirmation keeps dwindling as the chain becomes longer, there is less confirmation, other things being equal, in cases involving more intermediary steps.

A caveat is called for here. Theorem 1 by itself has no implications concerning how much dwindling there is in a given case. Even if the degree to which E confirms Hk is less than the degree to which E confirms Hk-1 for all $2 \leq k \leq n$, it might be that the amount of dwindling at each step is slight and that the total amount of dwindling is not substantial. It is worth noting in this regard that measures (1)–(13) differ in terms of how much

14. We are assuming here that a miracle can occur in principle at least, even if God does not exist.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

dwindling they allow in a given case. Consider R and Z, for instance, and a three-member series E, H1, H2 such that A, B, and C hold. There are distributions of probabilities on which each of $\mathrm{R}(\mathrm{H1}, \mathrm{E})$ and $\mathrm{R}(\mathrm{H2}, \mathrm{H1})$ is very high, say, roughly equal to 100,000, but $\mathrm{R}(\mathrm{H2}, \mathrm{E})$ is very low, say, roughly equal to 1.01. Degree of confirmation can drop off precipitously on R in a single step of mediation under SOC. Things are quite different with Z. When SOC holds, $\mathrm{Z}(\mathrm{H2}, \mathrm{E})$ equals the product of $\mathrm{Z}(\mathrm{H1}, \mathrm{E})$ and $\mathrm{Z}(\mathrm{H2}, \mathrm{H1})$.[15] So, if each of $\mathrm{Z}(\mathrm{H1}, \mathrm{E})$ and $\mathrm{Z}(\mathrm{H2}, \mathrm{H1})$ is very high, then it cannot be the case that $\mathrm{Z}(\mathrm{H2}, \mathrm{E})$ is very low.

3.3. Data-Processing Inequality. The Data-Processing Inequality concerns "mutual information." Let X and Y be discrete variables. The mutual information $\mathrm{I}(X, Y)$ between X and Y is the (weighted) average amount of information provided about (a proposition specifying) the state of the one variable by (a proposition specifying) the state of the other variable:

$$\sum_{x \in X} \sum_{y \in Y} \Pr(x \wedge y) \log \left(\frac{\Pr(x \wedge y)}{\Pr(x) \Pr(y)} \right).$$

Here x and y are propositions specifying the states of X and Y, respectively. Mutual information is symmetric in that $\mathrm{I}(X, Y) = \mathrm{I}(Y, X)$. Now let D, P, and E be discrete variables (D for "distal cause," P for "proximate cause," E for "effect") and suppose SOC holds in that each state of P screens off each state of E from each state of D.[16] Then, by the Data-Processing Inequality, the mutual information between E and D is less than the mutual information between E and P: $\mathrm{I}(E, D) < \mathrm{I}(E, P)$.[17] This result can be glossed: mutual information dwindles under SOC.

The Data-Processing Inequality has application to the issue of information loss in various settings. Consider, for example, an evolutionary process in which each member of a population has trait A or trait B (but not both).[18] Suppose the state of the population (the distribution of A and B in the population) at a given time screens off the state of the population at

15. This is clear from (B3) in app. B.

16. Screening-off is symmetric. So, as each state of P screens off each state of E from each state of D, it follows that each state of P screens off each state of D from each state of E.

17. See Cover and Thomas (2006, chap. 2) for explanation of mutual information and the Data-Processing Inequality (theorem 2.8.1). Strictly speaking, in order for $\mathrm{I}(E, D)$ to be less than $\mathrm{I}(E, P)$, it also needs to be the case that $\mathrm{I}(E, P|D)$ is greater than 0. This condition should be understood throughout this subsection.

18. See Sober and Steel (2013) for discussion of the Data-Processing Inequality, along with the Markov Chain Convergence Theorem and the issue of how evolutionary processes affect how much information the present provides about the past.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

any earlier time. In other words, the state of the population at a given time has a direct impact only on the immediately succeeding state, so that if it has an impact on any subsequent states, it does so only indirectly through its impact on the immediately succeeding state. It follows from the Data-Processing Inequality that the current state of the population provides more information on average about the state of the population at more recent times than about the state of the population at less recent times.

This result is interesting and resembles dwindling confirmation under SOC, but there are important differences. First, mutual information (as defined above) is the *expected* amount of information—the amount of information about the state of one variable we should expect to receive upon learning the state of the other variable.[19] Suppose we learn E, where E is a specific proposition about the current state of the population in the evolutionary case above, and our interest shifts from the expected amount of information to the degree of confirmation provided by E. Suppose, more specifically, we want to know whether E confirms H1 better than H2, where H1 is a proposition about the state of the population at an earlier time $t1$ and H2 is a proposition about the state of the population at an even earlier time $t2$. The Data-Processing Inequality is of no help here since, as explained above, the Data-Processing Inequality concerns mutual information, and mutual information is the expected amount of information. Of course, information theory also allows us to determine how much information particular proposition x provides about particular proposition y in the form of $\log(\Pr(x \wedge y)/\Pr(x)\Pr(y))$. So, we can compare the amounts of information E provides about H1 and H2, respectively. But, notice that the measure used here is the log-ratio measure of confirmation R*. So, when we translate the result into the language of confirmation theory, we are forced to use the specific measure R*. Theorem 1 above reveals not only that E provides less information about H2 than about H1 under SOC but also that there is no need to use the specific measure R*. If E confirms H1, H1 confirms H2, SOC holds, E neither entails nor is entailed by H1, and H1 neither entails nor is entailed by H2, then by theorem 1 the degree to which E confirms H2 is less than the degree to which E confirms H1 on any adequate confirmation measure. Theorem 1 thus serves as a welcome supplement to the Data-Processing Inequality.

4. Further Results. We now show that theorem 1 can be extended in two respects. The first respect (discussed in sec. 4.1) concerns a screening-off

19. Terminology varies. Fano (1961), e.g., distinguishes between mutual information and expected mutual information. Mutual information in our terminology is expected mutual information in Fano's terminology. Mutual information for Fano is not an average.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

condition weaker than SOC. The second respect (discussed in sec. 4.2) concerns cases where condition C—the condition that E neither entails nor is entailed by H1, and Hk-1 neither entails nor is entailed by Hk for all $2 \leq k \leq n$—does not hold.

4.1. Dwindling Confirmation under PISOC. SOC is the standard screening-off condition. But there are others. A less restrictive screening-off condition is

> **Positive Impact Screening-Off Condition (PISOC).** $\Pr(H_k|H_{k-1} \wedge E) \leq \Pr(H_k|H_{k-1})$ and $\Pr(H_k|\neg H_{k-1} \wedge E) \leq \Pr(H_k|\neg H_{k-1})$.

PISOC says in effect that given the truth or falsity of Hk-1, E has no positive impact on the probability of Hk—either E has no impact on the probability of Hk or the impact is negative. PISOC is weaker than SOC, in that PISOC holds if SOC holds but not vice versa.

PISOC, unlike SOC, is not a condition sufficient for transitivity in confirmation.[20] But PISOC, like SOC, is a condition sufficient for dwindling confirmation. Consider the condition:

> B*. E is screened off in the sense of PISOC from Hk by Hk-1 for all $2 \leq k \leq n$.

Condition B* is simply B with SOC replaced by PISOC. PISOC is a condition sufficient for dwindling confirmation in that

> **Theorem 2.** If A, B*, and C hold and X meets WLL, then $X(H_k, E) < X(H_{k-1}, E)$ for all $2 \leq k \leq n$.

(See app. C for proof.) The crucial thing here is that A, B*, and C—as with A, B, and C—together imply that (i) $\Pr(E|H_{k-1}) > \Pr(E|H_k)$ and (ii) $\Pr(E|\neg H_{k-1}) < \Pr(E|\neg H_k)$ for all $2 \leq k \leq n$.[21]

20. Recall the first card case given in sec. 1, where a card is randomly drawn from a standard deck of cards, and E = the card drawn is a heart, H1 = the card drawn is a red jack, and H2 = the card drawn is a diamond. We used the case as a counterexample to transitivity, but PISOC holds in it; $\Pr(H2|H1 \wedge E) = 0 < 1/2 = \Pr(H2|H1)$ and $\Pr(H2|\neg H1 \wedge E) = 0 < 6/25 = \Pr(H2|\neg H1)$. So, PISOC is not sufficient for transitivity in confirmation.

21. There is yet a third screening-off condition: **Negative Impact Screening-Off Condition** (NISOC)—$\Pr(H_k|H_{k-1} \wedge E) \geq \Pr(H_k|H_{k-1})$ and $\Pr(H_k|\neg H_{k-1} \wedge E) \geq \Pr(H_k|\neg H_{k-1})$. NISOC is like PISOC in that given the truth or falsity of Hk-1, E may still have some impact

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Theorem 2 marks a further respect in which dwindling confirmation under SOC differs from the Data-Processing Inequality. The latter cannot be extended to the case where PISOC holds but SOC does not.[22]

4.2. Nonincreasing Confirmation under SOC and PISOC. Recall that each of the main confirmation measures in the literature—measures (1)–(13)—meets WLL. The same is true of a slightly different condition:

> **WLL**.** If (*a*) $\Pr(E|H) \geq \Pr(E|H^*)$ and (*b*) $\Pr(E|\neg H) \leq \Pr(E|\neg H^*)$, then $X(H, E) \geq X(H^*, E)$.

First, if $\Pr(E|\neg H) \leq \Pr(E|\neg H^*)$, then

$$G(H, E) = \frac{\Pr(\neg H)}{\Pr(\neg H|E)} = \frac{\Pr(E)}{\Pr(E|\neg H)} \geq \frac{\Pr(E)}{\Pr(E|\neg H^*)} = \frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} = G(H^*, E),$$

and if $\Pr(E|H) \geq \Pr(E|H^*)$, then

$$R(H, E) = \frac{\Pr(H|E)}{\Pr(H)} = \frac{\Pr(E|H)}{\Pr(E)} \geq \frac{\Pr(E|H^*)}{\Pr(E)} = \frac{\Pr(H^*|E)}{\Pr(H^*)} = R(H^*, E).$$

---

on the probability of H*k*. But whereas PISOC precludes the impact from being positive, NISOC precludes it from being negative. NISOC, like SOC, is a condition sufficient for transitivity in confirmation (Roche 2012b). But it is not the case that confirmation dwindles under NISOC on all measures meeting WLL. In fact, for each of measures (1)–(13), there are distributions of probabilities on which E confirms H*k*-1, H*k*-1 confirms H*k*, NISOC holds, E neither entails nor is entailed by H*k*-1, H*k*-1 neither entails nor is entailed by H*k*, and yet the confirmation provided to H*k* by E is greater than or equal to the confirmation provided by E to H*k*-1.

22. There are (at least) two ways to try to extend the Data-Processing Inequality to the case where PISOC holds but SOC does not. Consider a case involving dichotomous variables *X*, *Y*, *Z*, where $x1 = X$ has value 1, $x2 = X$ has value 2, and similarly for $y1$, $y2$, $z1$, $z2$. First, one can require that each member of the partition $\{x1, x2\}$ be screened off in the sense of PISOC by each member of the partition $\{y1, y2\}$ from each member of the partition $\{z1, z2\}$. Second, one can require just that $x1$ be screened off in the sense of PISOC from $z1$ by $y1$. Both options are problematic. The first option fails because there is no case in which each member of $\{x1, x2\}$ is screened off in the sense of PISOC by each member of $\{y1, y2\}$ from each member of $\{z1, z2\}$ but each member of $\{x1, x2\}$ is not screened off in the sense of SOC by each member of $\{y1, y2\}$ from each member of $\{z1, z2\}$. For, if, say, $\Pr(z1|y1 \wedge x1) \leq \Pr(z1|y1)$, then $\Pr(z2|y1 \wedge x1) \geq \Pr(z2|y1)$. So, $\Pr(z1|y1 \wedge x1) \leq \Pr(z1|y1)$ and $\Pr(z2|y1 \wedge x1) \leq \Pr(z2|y1)$ only if $\Pr(z1|y1 \wedge x1) = \Pr(z1|y1)$ and $\Pr(z2|y1 \wedge x1) = \Pr(z2|y1)$, as required by SOC. The second option fails because there are counterexamples; more specifically, there are cases where $x1$ is screened off in the sense of PISOC from $z1$ by $y1$ and yet $I(X, Y)$ is less than $I(X, Z)$.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

So G and R meet WLL**. Next, given that G and R meet WLL**, and given the list of equalities provided in section 2, it follows that C, D, K, L, L*, M, N, R*, S, and Z meet WLL**. These measures are fully determined by G, R, and Pr(E) and are monotonically increasing functions of G, R, or both. As noted in section 2, J is not "reducible" to G, R, and Pr(E), as far as we can tell, but it can be shown (in a way parallel to the proof given in app. A) that J also meets WLL**.

A few remarks are in order concerning WLL** as it relates to WLL and WLL*. First, all inequalities in WLL are strict (">" or "<"), while all inequalities in WLL** are nonstrict ("≥" or "≤"). WLL states in essence that if H is uniformly better than H* at predicting E, then E confirms H more than it does H*. Similarly, WLL** states in essence that if H is uniformly at least as good as H* at predicting E, then E confirms H at least as well as it does H*. WLL** is every bit as intuitive as WLL. Any adequate confirmation measure should satisfy WLL**. Second, whereas all inequalities in WLL** are nonstrict, this is not the case with WLL*—the first and third are strict, while the second is nonstrict. WLL* states in essence that if H is in some way better than H* and in another way as good as H* at predicting E, then E confirms H better than it does H*. Recall that not all of measures (1)–(13) meet WLL*. This is because on some such measures the first condition in the antecedent of WLL* is irrelevant to the truth or falsity of the consequent so that WLL* is tantamount to the dubious claim that if H is at least as good as H* at predicting E, then E confirms H better than it does H*. Things are different with WLL**. The measures in question imply that the first condition in the antecedent of WLL** is irrelevant to the truth or falsity of the consequent, but they also imply that WLL** is tantamount to the plausible claim that if H is at least as good as H* at predicting E, then E confirms H at least as well as it does H*.

Recall condition C that E neither entails nor is entailed by H1, and Hk-1 neither entails nor is entailed by Hk for all $2 \leq k \leq n$. It is essential to theorem 1 that C holds in that if A and B hold but C does not, then, even supposing that X meets WLL, it might be that $\mathrm{X}(\mathrm{H}k, E) \prec \mathrm{X}(\mathrm{H}k\text{-}1, \mathrm{E})$ for some $2 \leq k \leq n$. However, we now have

> **Theorem 3.** If A and B hold and X meets WLL**, then $\mathrm{X}(\mathrm{H}k, \mathrm{E}) \leq \mathrm{X}(\mathrm{H}k\text{-}1, \mathrm{E})$ for all $2 \leq k \leq n$.

(See app. D for proof.) Theorem 3, unlike theorems 1 and 2, is not limited to cases where C holds. As a result, in some instances of theorem 3, $\mathrm{X}(\mathrm{H}k, \mathrm{E}) = \mathrm{X}(\mathrm{H}k\text{-}1, \mathrm{E})$, and thus confirmation does not dwindle. Theorem 3 can be glossed: confirmation in a chain of support never increases under SOC.

It is straightforward to show that, similarly, confirmation never increases under PISOC:

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Theorem 4. If A and B* hold and X meets WLL**, then $X(Hk, E) \leq X(Hk\text{-}1, E)$ for all $2 \leq k \leq n$.

(The proofs for theorems 2 and 3 can be readily adapted to establish theorem 4.)

5. Conclusion. To state the main result of the article in nonformal terms, as the chain of probabilistic support becomes longer, confirmation dwindles provided that the first proposition supports each of the subsequent propositions only indirectly through the chain. This is true on each of the main confirmation measures in the literature and, more generally, on any adequate confirmation measure. This result applies to many testimony cases, and it serves as a welcome supplement to the Data-Processing Inequality. One question for future research is whether expected confirmation dwindles under SOC—not just on the log-ratio measure of confirmation, as the Data-Processing Inequality implies, but also on any adequate confirmation measure. If so, we can make claims about expected confirmation without taking a stand on which of the many possible confirmation measures is to be preferred.

Appendix A

Proof of J's Meeting WLL. Suppose the antecedent of WLL holds. Then, since G and R meet WLL, it follows that $G(H^*, E) < G(H, E)$ and $R(H^*, E) < R(H, E)$. Given that $R(H^*, E) < R(H, E)$, there are four cases to consider:

$$1 < R(H^*, E) < R(H, E); \tag{A1}$$

$$0 < R(H^*, E) \leq 1 < R(H, E); \tag{A2}$$

$$0 < R(H^*, E) < 1 \leq R(H, E); \tag{A3}$$

$$0 < R(H^*, E) < R(H, E) < 1. \tag{A4}$$

We begin with cases (A2) and (A3) and then turn to cases (A1) and (A4).

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Rewrite $J(H, E)$ and $J(H^*, E)$ as follows:

$$J(H, E) = \frac{\log[R(H, E)]}{-\log[Pr(H)]};$$
$$J(H^*, E) = \frac{\log[R(H^*, E)]}{-\log[Pr(H^*)]}.$$

Since $0 < -\log[Pr(H)]$ and $0 < -\log[Pr(H^*)]$, it follows from (A2) and (A3), respectively, that

$$J(H^*, E) \leq 0 < J(H, E); \tag{A2*}$$

$$J(H^*, E) < 0 \leq J(H, E). \tag{A3*}$$

So, $J(H^*, E) < J(H, E)$ in cases (A2) and (A3).

Cases (A1) and (A4)—the two remaining cases—imply, respectively, that

$$0 < \log[R(H^*, E)] < \log[R(H, E)]; \tag{A1*}$$

$$\log[R(H^*, E)] < \log[R(H, E)] < 0. \tag{A4*}$$

In both case (A1) and case (A4), either (*a*) $0 < Pr(H^*) \leq Pr(H) < 1$ or (*b*) $0 < Pr(H) < Pr(H^*) < 1$, from which it follows, respectively, that

*a**) $\log[Pr(H^*)] \leq \log[Pr(H)] < 0$;

*b**) $\log[Pr(H)] < \log[Pr(H^*)] < 0$.

To show that $J(H^*, E) < J(H, E)$ in case (A1), it suffices to show that $J(H^*, E) < J(H, E)$ in the two subcases: (A1)*a* and (A1)*b*. Likewise, to show that $J(H^*, E) < J(H, E)$ in case (A4), it suffices to show that $J(H^*, E) < J(H, E)$ in the two subcases: (A4)*a* and (A4)*b*.

First subcase, (A1)*a*.—It follows immediately from (A1*)*a** that $J(H^*, E) < J(H, E)$. *Second subcase*, (A1)*b*.—Let $Pr(H^*) = Pr(H)^{\alpha}$, where $0 < \alpha < 1$ from *b*. Since $G(H^*, E) < G(H, E)$, $Pr(\neg H^*)/Pr(\neg H^*|E) < Pr(\neg H)/Pr(\neg H|E)$. Given this, and given that

https://doi.org/10.1086/674346 Published online by Cambridge University Press

$$\frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} < \frac{\Pr(\neg H)}{\Pr(\neg H|E)} \text{ iff}$$

$$\frac{1 - \Pr(H|E)}{1 - \Pr(H^*|E)} < \frac{1 - \Pr(H)}{1 - \Pr(H^*)} \text{ iff}$$

$$\frac{1 - \Pr(H)R(H, E)}{1 - \Pr(H^*)R(H^*, E)} < \frac{1 - \Pr(H)}{1 - \Pr(H^*)} \text{ iff}$$

$$\frac{1 - \Pr(H)R(H, E)}{1 - \Pr(H)^{\alpha}R(H^*, E)} < \frac{1 - \Pr(H)}{1 - \Pr(H)^{\alpha}},$$

we have

$$\frac{1 - \Pr(H)R(H, E)}{1 - \Pr(H)^{\alpha}R(H^*, E)} < \frac{1 - \Pr(H)}{1 - \Pr(H)^{\alpha}}. \tag{A5}$$

Meanwhile, it follows from $\alpha < 1$ and $0 < \Pr(H) < 1$ that $(1 - x)/(1 - x^{\alpha})$ is a strictly increasing function of x. But $\Pr(H1) < \Pr(H1)R(H1, E)$ from (A1). So,

$$\frac{1 - \Pr(H)}{1 - \Pr(H)^{\alpha}} < \frac{1 - \Pr(H)R(H, E)}{1 - [\Pr(H)R(H, E)]^{\alpha}}. \tag{A6}$$

It follows from (A5) and (A6) that

$$\frac{1 - \Pr(H)R(H, E)}{1 - \Pr(H)^{\alpha}R(H^*, E)} < \frac{1 - \Pr(H)R(H, E)}{1 - [\Pr(H)R(H, E)]^{\alpha}}. \tag{A7}$$

Hence, $R(H^*, E) < R(H, E)^{\alpha}$. So,

$$\log[R(H^*, E)] < \alpha \log[R(H, E)]. \tag{A8}$$

It follows from (A8) and $0 < -\log[\Pr(H)]$ that

$$\frac{\log[R(H^*, E)]}{-\log[\Pr(H)]} < \frac{\alpha \log[R(H, E)]}{-\log[\Pr(H)]}. \tag{A9}$$

Equation (A9) implies

$$\frac{\log[R(H^*, E)]}{-\alpha \log[\Pr(H)]} < \frac{\log[R(H, E)]}{-\log[\Pr(H)]}. \tag{A10}$$

https://doi.org/10.1086/674346 Published online by Cambridge University Press

It then follows that $J(H^*, E) < J(H, E)$ as follows:

$$\frac{\log[R(H^*, E)]}{-\log[Pr(H^*)]} = \frac{\log[R(H^*, E)]}{-\log[Pr(H)^{\alpha}]} = \frac{\log[R(H^*, E)]}{-\alpha \log[Pr(H)]} < \frac{\log[R(H, E)]}{-\log[Pr(H)]}. \tag{A11}$$

Hence, $J(H^*, E) < J(H, E)$ in case (A1).

Third subcase, (A4)*a*.—Let $Pr(H^*) = Pr(H)^{\alpha}$, where $1 \leq \alpha$ from *a*. Equation (A5) holds by the same reasoning as in the second subcase. Meanwhile, it follows from $1 \leq \alpha$ that $(1 - x)/(1 - x^{\alpha})$ is a strictly decreasing function of x. But $Pr(H1) > Pr(H1)R(H1, E)$ from (A4). So (A6) holds. The argument then continues in parallel to the argument given in the second subcase. *Fourth subcase*, (A4)*b*.—It follows immediately from (A4*)*b** that $J(H^*, E) < J(H, E)$. Thus, $J(H^*, E) < J(H, E)$ in case (A4). So, as $J(H^*, E) < J(H, E)$ in cases (A1)–(A4), it follows that J meets WLL. QED

Appendix B

Proof of Theorem 1: We establish theorem 1 in two steps. First, we establish a lemma:

> **Lemma.** If A, B, and C hold, then (*a*) $Pr(E|Hk\text{-}1) > Pr(E|Hk)$ and (*b*) $Pr(E|\neg Hk\text{-}1) < Pr(E|\neg Hk)$ for all $2 \leq k \leq n$.

Second, we argue from the lemma to theorem 1.

Consider a three-member series E, H1, H2 and suppose the antecedent of the lemma holds. It is known (Shogenji 2003) that if H1 screens off E from H2, as B implies, then

$$Pr(H2|E) - Pr(H2) = \frac{(Pr(H1|E) - Pr(H1))(Pr(H2|H1) - Pr(H2))}{1 - Pr(H1)}. \tag{B1}$$

Equation (B1) implies

$$\frac{Pr(H2|E) - Pr(H2)}{Pr(H2)} = \left(\frac{Pr(H1|E) - Pr(H1)}{Pr(H1)}\right)\left(\frac{Pr(H1)(Pr(H2|H1) - Pr(H2))}{Pr(H2)(1 - Pr(H1))}\right); \tag{B2}$$

https://doi.org/10.1086/674346 Published online by Cambridge University Press

$$\frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} = \frac{(\Pr(H1|E) - \Pr(H1))(\Pr(H2|H1) - \Pr(H2))}{(1 - \Pr(H1))(1 - \Pr(H2))}. \tag{B3}$$

But, given C, we have

$$\begin{aligned} \Pr(H1)(\Pr(H2|H1) - \Pr(H2)) &= \Pr(H1 \wedge H2) - \Pr(H1)\Pr(H2) \\ &< \Pr(H2) - \Pr(H2)(H1) = \Pr(H2)(1 - \Pr(H1)). \end{aligned} \tag{B4}$$

Equation (B2) and inequality (B4) together entail

$$\frac{\Pr(H2|E) - \Pr(H2)}{\Pr(H2)} < \frac{\Pr(H1|E) - \Pr(H1)}{\Pr(H1)}. \tag{B5}$$

Inequality (B5) implies

$$R(H2, E) = \frac{\Pr(H2|E)}{\Pr(H2)} < \frac{\Pr(H1|E)}{\Pr(H1)} = R(H1, E). \tag{B6}$$

Next, from C, we have

$$\Pr(H2|H1) - \Pr(H2) < 1 - \Pr(H2). \tag{B7}$$

Equation (B3) and inequality (B7) together imply

$$\frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} < \frac{\Pr(H1|E) - \Pr(H1)}{1 - \Pr(H1)}. \tag{B8}$$

But, we have

$$\begin{aligned} &\frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} < \frac{\Pr(H1|E) - \Pr(H1)}{1 - \Pr(H1)} \text{ iff} \\ &\left(1 - \frac{\Pr(\neg H2|E)}{\Pr(\neg H2)}\right) < \left(1 - \frac{\Pr(\neg H1|E)}{\Pr(\neg H1)}\right) \text{ iff} \\ &\frac{\Pr(\neg H2|E)}{\Pr(\neg H2)} > \frac{\Pr(\neg H1|E)}{\Pr(\neg H1)} \text{ iff} \\ &\frac{\Pr(\neg H2)}{\Pr(\neg H2|E)} < \frac{\Pr(\neg H1)}{\Pr(\neg H1|E)}. \end{aligned} \tag{B9}$$

https://doi.org/10.1086/674346 Published online by Cambridge University Press

It follows from (B8) and (B9) that

$$G(H2, E) = \frac{Pr(\neg H2)}{Pr(\neg H2|E)} < \frac{Pr(\neg H1)}{Pr(\neg H1|E)} = G(H1, E). \quad (B10)$$

So, if the antecedent of the lemma holds for a three-member series E, H1, H2, then $R(H2, E) < R(H1, E)$ and $G(H2, E) < G(H1, E)$.

Now we generalize this result to the n-member series $E, H1, H2, \ldots, Hn$. Suppose the antecedent of the lemma holds. SOC is a condition sufficient for transitivity in confirmation, so, given A and B, it follows that E confirms Hk for all $1 \leq k \leq n$. Next, given A–C, and given the result from the previous paragraph, it follows that $R(Hk, E) < R(Hk\text{-}1, E)$ and $G(Hk, E) < G(Hk\text{-}1, E)$.

The lemma is but a short step away: $R(H, E) = Pr(H|E)/Pr(H) = Pr(E|H)/Pr(E)$ and $G(H, E) = Pr(\neg H)/Pr(\neg H|E) = Pr(E)/Pr(E|\neg H)$. So, as $R(Hk, E) < R(Hk\text{-}1, E)$ and $G(Hk, E) < G(Hk\text{-}1, E)$, we have

$$\frac{Pr(E|Hk)}{Pr(E)} = \frac{Pr(Hk|E)}{Pr(Hk)} < \frac{Pr(Hk\text{-}1|E)}{Pr(Hk\text{-}1)} = \frac{Pr(E|Hk\text{-}1)}{Pr(E)}; \quad (B11)$$

$$\frac{Pr(E)}{Pr(E|\neg Hk)} = \frac{Pr(\neg Hk)}{Pr(\neg Hk|E)} < \frac{Pr(\neg Hk\text{-}1)}{Pr(\neg Hk\text{-}1|E)} = \frac{Pr(E)}{Pr(E|\neg Hk\text{-}1)}. \quad (B12)$$

It follows immediately from (B11) and (B12), respectively, that $Pr(E|Hk\text{-}1) > Pr(E|Hk)$ and $Pr(E|\neg Hk\text{-}1) < Pr(E|\neg Hk)$. The lemma is thus established.

Suppose now the antecedent of theorem 1 holds. Then, the antecedent of the lemma also holds. So, it follows by the lemma that (*a*) $Pr(E|Hk\text{-}1) > Pr(E|Hk)$ and (*b*) $Pr(E|\neg Hk\text{-}1) < Pr(E|\neg Hk)$. It then follows, as X meets WLL, that $X(Hk, E) < X(Hk\text{-}1, E)$. QED

Appendix C

Proof of Theorem 2: Suppose the antecedent of theorem 2 holds. Then, by reasoning similar to that for (B1), we have

$$Pr(H2|E) - Pr(H2) \leq \frac{(Pr(H1|E) - Pr(H1))(Pr(H2|H1) - Pr(H2))}{1 - Pr(H1)}. \quad (C1)$$

https://doi.org/10.1086/674346 Published online by Cambridge University Press

It follows from (C1) that

$$\frac{\Pr(\mathrm{H2}|\mathrm{E}) - \Pr(\mathrm{H2})}{\Pr(\mathrm{H2})} \leq \left(\frac{\Pr(\mathrm{H1}|\mathrm{E}) - \Pr(\mathrm{H1})}{\Pr(\mathrm{H1})}\right)\left(\frac{\Pr(\mathrm{H1})(\Pr(\mathrm{H2}|\mathrm{H1}) - \Pr(\mathrm{H2}))}{\Pr(\mathrm{H2})(1 - \Pr(\mathrm{H1}))}\right); \tag{C2}$$

$$\frac{\Pr(\mathrm{H2}|\mathrm{E}) - \Pr(\mathrm{H2})}{1 - \Pr(\mathrm{H2})} \leq \frac{(\Pr(\mathrm{H1}|\mathrm{E}) - \Pr(\mathrm{H1}))(\Pr(\mathrm{H2}|\mathrm{H1}) - \Pr(\mathrm{H2}))}{(1 - \Pr(\mathrm{H1}))(1 - \Pr(\mathrm{H2}))}. \tag{C3}$$

With (C2) and (C3), instead of (B2) and (B3), the proof continues in parallel to the proof given above for theorem 1 (details omitted so as to avoid repetition). First, it is shown that if A and B* hold for a three-member series E, H1, H2, then $\mathrm{R}(\mathrm{H2}, \mathrm{E}) < \mathrm{R}(\mathrm{H1}, \mathrm{E})$ and $\mathrm{G}(\mathrm{H2}, \mathrm{E}) < \mathrm{G}(\mathrm{H1}, \mathrm{E})$. Next, it is shown that the point generalizes to the *n*-member series $\mathrm{E}, \mathrm{H1}, \mathrm{H2}, \ldots,$ H*n*, and thus $\Pr(\mathrm{E}|\mathrm{H}k\text{-}1) > \Pr(\mathrm{E}|\mathrm{H}k)$ and $\Pr(\mathrm{E}|\neg\,\mathrm{H}k\text{-}1) < \Pr(\mathrm{E}|\neg\,\mathrm{H}k)$. Finally, since X meets WLL, it follows from the antecedent of theorem 2 that $\mathrm{X}(\mathrm{H}k, \mathrm{E}) < \mathrm{X}(\mathrm{H}k\text{-}1, \mathrm{E})$. QED

Appendix D

Proof of Theorem 3: The proof for theorem 1 can be adapted to establish theorem 3. Suppose the antecedent of theorem 3 holds. It follows from A and B that (B1), (B2), and (B3) hold. Then, instead of (B4), which is a strict inequality, we have

$$\begin{aligned}&\Pr(\mathrm{H1})(\Pr(\mathrm{H2}|\mathrm{H1}) - \Pr(\mathrm{H2})) = \Pr(\mathrm{H1} \wedge \mathrm{H2}) - \Pr(\mathrm{H1})\Pr(\mathrm{H2})\\ &\leq \Pr(\mathrm{H2}) - \Pr(\mathrm{H2})(\mathrm{H1}) = \Pr(\mathrm{H2})(1 - \Pr(\mathrm{H1})).\end{aligned} \tag{D1}$$

Equation (B2) and inequality (D1) imply

$$\frac{\Pr(\mathrm{H2}|\mathrm{E}) - \Pr(\mathrm{H2})}{\Pr(\mathrm{H2})} \leq \frac{\Pr(\mathrm{H1}|\mathrm{E}) - \Pr(\mathrm{H1})}{\Pr(\mathrm{H1})}. \tag{D2}$$

Inequality (D2) entails

$$\mathrm{R}(\mathrm{H2}, \mathrm{E}) = \frac{\Pr(\mathrm{H2}|\mathrm{E})}{\Pr(\mathrm{H2})} \leq \frac{\Pr(\mathrm{H1}|\mathrm{E})}{\Pr(\mathrm{H1})} = \mathrm{R}(\mathrm{H1}, \mathrm{E}). \tag{D3}$$

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Further, instead of (B7), which is a strict inequality, we have

$$\Pr(\mathrm{H2}|\mathrm{H1}) - \Pr(\mathrm{H2}) \leq 1 - \Pr(\mathrm{H2}). \tag{D4}$$

Equation (B3) and inequality (D4) together entail

$$\frac{\Pr(\mathrm{H2}|\mathrm{E}) - \Pr(\mathrm{H2})}{1 - \Pr(\mathrm{H2})} \leq \frac{\Pr(\mathrm{H1}|\mathrm{E}) - \Pr(\mathrm{H1})}{1 - \Pr(\mathrm{H1})}. \tag{D5}$$

But we have

$$\begin{gathered}
\frac{\Pr(\mathrm{H2}|\mathrm{E}) - \Pr(\mathrm{H2})}{1 - \Pr(\mathrm{H2})} \leq \frac{\Pr(\mathrm{H1}|\mathrm{E}) - \Pr(\mathrm{H1})}{1 - \Pr(\mathrm{H1})} \text{ iff} \\
\left(1 - \frac{\Pr(\neg\,\mathrm{H2}|\mathrm{E})}{\Pr(\neg\,\mathrm{H2})}\right) \leq \left(1 - \frac{\Pr(\neg\,\mathrm{H1}|\mathrm{E})}{\Pr(\neg\,\mathrm{H1})}\right) \text{ iff} \\
\frac{\Pr(\neg\,\mathrm{H2}|\mathrm{E})}{\Pr(\neg\,\mathrm{H2})} \geq \frac{\Pr(\neg\,\mathrm{H1}|\mathrm{E})}{\Pr(\neg\,\mathrm{H1})} \text{ iff} \\
\frac{\Pr(\neg\,\mathrm{H2})}{\Pr(\neg\,\mathrm{H2}|\mathrm{E})} \leq \frac{\Pr(\neg\,\mathrm{H1})}{\Pr(\neg\,\mathrm{H1}|\mathrm{E})}.
\end{gathered} \tag{D6}$$

It follows from (D5) and (D6) that

$$\mathrm{G}(\mathrm{H2}, \mathrm{E}) = \frac{\Pr(\neg\,\mathrm{H2})}{\Pr(\neg\,\mathrm{H2}|\mathrm{E})} \leq \frac{\Pr(\neg\,\mathrm{H1})}{\Pr(\neg\,\mathrm{H1}|\mathrm{E})} = \mathrm{G}(\mathrm{H1}, \mathrm{E}). \tag{D7}$$

So if A and B hold for a three-member series E, H1, H2, then $\mathrm{R}(\mathrm{H2}, \mathrm{E}) \leq \mathrm{R}(\mathrm{H1}, \mathrm{E})$ and $\mathrm{G}(\mathrm{H2}, \mathrm{E}) \leq \mathrm{G}(\mathrm{H1}, \mathrm{E})$. The argument now proceeds in parallel to the argument given above for theorem 1 (details omitted so as to avoid repetition). First, it is shown that the point just established about a three-member series can be generalized to the *n*-member series $\mathrm{E}, \mathrm{H1}, \mathrm{H2}, \ldots,$ H*n*, and thus $\Pr(\mathrm{E}|\mathrm{H}k\text{-}1) \geq \Pr(\mathrm{E}|\mathrm{H}k)$ and $\Pr(\mathrm{E}|\neg\,\mathrm{H}k\text{-}1) \leq \Pr(\mathrm{E}|\neg\,\mathrm{H}k)$. Then, from the antecedent of theorem 3, which says in part that X meets WLL**, it follows that $\mathrm{X}(\mathrm{H}k, \mathrm{E}) \leq \mathrm{X}(\mathrm{H}k\text{-}1, \mathrm{E})$. QED

REFERENCES

Brossel, Peter. 2013. "The Problem of Measure Sensitivity Redux." *Philosophy of Science* 80: 378–97.

Carnap, Rudolf. 1962. *Logical Foundations of Probability.* 2nd ed. Chicago: University of Chicago Press.

https://doi.org/10.1086/674346 Published online by Cambridge University Press

Cover, Thomas, and Joy Thomas. 2006. *Elements of Information Theory*. 2nd ed. Hoboken, NJ: Wiley.
Crupi, Vincenzo, Roberto Festa, and Carlo Buttasi. 2010. "Towards a Grammar of Bayesian Confirmation." In *Epistemology and Philosophy of Science*, ed. Mauricio Suarez, Mauro Dorato, and Miklos Redei, 73–93. Dordrecht: Springer.
Crupi, Vincenzo, Katya Tentori, and Michel Gonzalez. 2007. "On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues." *Philosophy of Science* 74:229–52.
Douven, Igor. 2011. "Further Results on the Intransitivity of Evidential Support." *Review of Symbolic Logic* 4:487–97.
Eells, Ellery, and Branden Fitelson. 2002. "Symmetries and Asymmetries in Evidential Support." *Philosophical Studies* 107:129–42.
Fano, Robert. 1961. *Transmission of Information: A Statistical Theory of Communications*. Cambridge, MA: MIT Press.
Festa, Roberto. 1999. "Bayesian Confirmation." In *Experience, Reality, and Scientific Explanation*, ed. Maria Galavotti and Alessandro Pagnini, 55–87. Dordrecht: Kluwer.
Fitelson, Branden. 1999. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (Proceedings): S362–S378.
———. 2007. "Likelihoodism, Bayesianism, and Relational Confirmation." *Synthese* 156:473–89.
———. 2011. "Favoring, Likelihoodism, and Bayesianism." *Philosophy and Phenomenological Research* 83:666–72.
Joyce, James. 2008. "Bayes' Theorem." In *Stanford Encyclopedia of Philosophy*, ed. Edward Zalta. Stanford, CA: Stanford University. http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/.
McGrew, Timothy. 2004. "Has Plantinga Refuted the Historical Argument?" *Philosophia Christi* 6:7–26.
Plantinga, Alvin. 2000. *Warranted Christian Belief*. New York: Oxford University Press.
———. 2006. "Historical Arguments and Dwindling Probabilities." *Philosophia Christi* 8:7–22.
Roche, William. 2012a. "Transitivity and Intransitivity in Evidential Support: Some Further Results." *Review of Symbolic Logic* 5:259–68.
———. 2012b. "A Weaker Condition for Transitivity in Probabilistic Support." *European Journal for Philosophy of Science* 2:111–18.
Roche, William, and Tomoji Shogenji. 2013. "Confirmation, Transitivity, and Moore: The Screening-Off Approach." *Philosophical Studies*. doi:10.1007/s11098-013-0161-3.
Shogenji, Tomoji. 2003. "A Condition for Transitivity in Probabilistic Support." *British Journal for the Philosophy of Science* 54:613–16.
———. 2012. "The Degree of Epistemic Justification and the Conjunction Fallacy." *Synthese* 184:29–48.
Sober, Elliott. 2009. "Absence of Evidence and Evidence of Absence: Evidential Transitivity in Connection with Fossils, Fishing, Fine-Tuning, and Firing Squads." *Philosophical Studies* 143:63–90.
———. 2011. "Responses to Fitelson, Sansom, and Sarkar." *Philosophy and Phenomenological Research* 83:692–704.
Sober, Elliott, and Mike Steel. 2013. "Time and Knowability in Evolutionary Processes." Unpublished manuscript, arXiv.org. http://arxiv-web3.library.cornell.edu/abs/1301.6470.
Swinburne, Richard. 2004. "Natural Theology, Its 'Dwindling Probabilities' and 'Lack of Rapport.'" *Faith and Philosophy* 21:533–46.
Tornebohm, Hakan. 1966. "Two Measures of Evidential Strength." In *Aspects of Inductive Logic*, ed. Jaakko Hintikka and Patrick Suppes, 81–95. Amsterdam: North-Holland.

https://doi.org/10.1086/674346 Published online by Cambridge University Press