

Dwindling Confirmation

William Roche* and Tomoji Shogenji†

* Department of Philosophy, Texas Christian University, Fort Worth, TX, USA, e-mail: w.roche@tcu.edu

† Department of Philosophy, Rhode Island College, Providence, RI, USA, e-mail: tshogenji@ric.edu

ABSTRACT: 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.

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

1. Introduction. Think of a series of propositions E, H1, H2, ..., Hn such that E supports H1, H1 supports H2, ..., and Hn-1 supports Hn. If the support relation is logical entailment, then, regardless of the length of the series, E supports Hn 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 Hn.¹ This is because confirmation, unlike logical entailment,

¹ A few comments are in order here. First, here and throughout the paper 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 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 if and only if

is non-transitive. 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. 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 non-transitive 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. $H1$ = the card drawn is a Red. $H2$ = the card drawn is a Heart. 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 E confirms $H2$.

A second issue of interest, regarding a series $E, H1, H2, \dots, Hn$ 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 Hn is less than the degree to which it confirms $Hn-1$, the degree to which it confirms $Hn-1$ is less than the degree to which it confirms $Hn-2$, and so on.² Here is an example. E = Smith said that miracle M occurred. $H1$ = miracle M occurred. $H2$ = God exists. E confirms $H1$, at least on certain ways of filling in the details, $H1$ in turn confirms $H2$, and, it seems, E confirms $H2$ (though the probability of $H2$ given E need not be high, 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 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 the probability of $H1$

the probability of H given E is sufficiently high. See Carnap (1962, Preface to the Second Edition) 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 $E, H1, \dots, Hn$) concerns just the probabilities $\Pr(H1 | E), \Pr(H2 | E), \dots, \Pr(Hn | E)$ 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.

(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 condition:

Screening-Off Condition (SOC). $\Pr(Hk | Hk-1 \wedge E) = \Pr(Hk | Hk-1)$ and $\Pr(Hk | \neg Hk-1 \wedge E) = \Pr(Hk | \neg Hk-1)$.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 $\Pr(Hk-1 | E) > \Pr(Hk-1)$, $\Pr(Hk | Hk-1) > \Pr(Hk)$, 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 paper proceeds as follows. In Section 2, we provide a list of the main confirmation measures in the literature (thirteen 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 paper.⁶ We apply our result to testimony cases of

³ It should be noted too that each of the main confirmation measures discussed in the literature, and listed below in Section 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 non-dwinding under (SOC), i.e., he shows that on three measures—the difference measure, the log-likelihood measure, and the

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 the following:

- (i) $C(H, E) = \Pr(H \wedge E) - \Pr(H)\Pr(E)$;
- (ii) $D(H, E) = \Pr(H|E) - \Pr(H)$;
- (iii) $G(H, E) = \frac{\Pr(\neg H)}{\Pr(\neg H|E)}$;
- (iv) $J(H, E) = \frac{\text{Log}_2[\Pr(H|E)] - \text{Log}_2[\Pr(H)]}{-\text{Log}_2[\Pr(H)]}$;
- (v) $K(H, E) = \frac{\Pr(E|H) - \Pr(E|\neg H)}{\Pr(E|H) + \Pr(E|\neg H)}$;
- (vi) $L(H, E) = \frac{\Pr(E|H)}{\Pr(E|\neg H)}$;
- (vii) $L^*(H, E) = \text{Log} \left[\frac{\Pr(E|H)}{\Pr(E|\neg H)} \right]$;
- (viii) $M(H, E) = \Pr(E|H) - \Pr(E)$;
- (ix) $N(H, E) = \Pr(E|H) - \Pr(E|\neg H)$;
- (x) $R(H, E) = \frac{\Pr(H|E)}{\Pr(H)}$;
- (xi) $R^*(H, E) = \text{Log} \left[\frac{\Pr(H|E)}{\Pr(H)} \right]$;
- (xii) $S(H, E) = \Pr(H|E) - \Pr(H|\neg E)$;

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 (2012b). 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 non-dwindling. We establish the positive universal claim that, when a certain condition concerning logical entailment holds, confirmation *invariably* dwindles under (SOC).

$$(xiii) \quad Z(H, E) = \begin{cases} \frac{\Pr(H|E) - \Pr(H)}{1 - \Pr(H)} & \text{if } \Pr(H|E) \geq \Pr(H) \\ \frac{\Pr(H|E) - \Pr(H)}{\Pr(H)} & \text{otherwise} \end{cases} \quad ^7$$

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.

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 below 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 (i)-(xiii) meets (WLL). This point is easiest to see with respect to measures (i)-(iii) and (v)-(xiii). If $\Pr(E | \neg H) < \Pr(E | \neg H^*)$, then $G(H, E) = \frac{\Pr(\neg H)}{\Pr(\neg H|E)} = \frac{\Pr(E)}{\Pr(E|\neg H)} > \frac{\Pr(E)}{\Pr(E|\neg H^*)} = \frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} = G(H^*,$

⁷ For discussion and references regarding measures (i)-(iii) and (v)-(xiii), see Crupi et al. (2007), Eells and Fitelson (2002), and Festa (1999). J is given in Shogenji (2012). See also Tornebohm (1966).

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 $X^*(H, E) > X^*(H', E')$.

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 et al. (2010), says in effect that X discriminates between 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.

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 (i), (ii), (v)-(ix), and (xi)-(xiii), consider the following equalities:

$$\begin{aligned}
C(H, E) &= \left[\frac{1}{\frac{1}{1-\frac{1}{G(H,E)}} + \frac{1}{R(H,E)-1}} \right] \Pr(E); \\
D(H, E) &= \frac{1}{\frac{1}{1-\frac{1}{G(H,E)}} + \frac{1}{R(H,E)-1}}; \\
K(H, E) &= \frac{R(H,E)G(H,E)-1}{R(H,E)G(H,E)+1}; \\
L(H, E) &= R(H, E)G(H, E); \\
L^*(H, E) &= \text{Log}[R(H, E)G(H, E)]; \\
M(H, E) &= \Pr(E) [R(H, E)-1]; \\
N(H, E) &= \left[R(H, E) - \frac{1}{G(H,E)} \right] \Pr(E); \\
R^*(H, E) &= \text{Log}[R(H, E)]; \\
S(H, E) &= \frac{\left[\frac{1}{\frac{1}{1-\frac{1}{G(H,E)}} + \frac{1}{R(H,E)-1}} \right]}{1-\Pr(E)}; \\
Z(H, E) &= \begin{cases} 1 - \frac{1}{G(H,E)} & \text{if } \Pr(H|E) \geq \Pr(H) \\ R(H, E) - 1 & \text{otherwise} \end{cases}.
\end{aligned}$$

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 (ii), (v)-(ix), and (xi)-(xiii). J (as far as we are aware) is not “reducible” to G, R, and Pr(E). But it too meets (WLL) (see Appendix A for proof).

(WLL) is similar to “the weak likelihood principle” which is endorsed by Joyce (2008) and which is equivalent to the following:

(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

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. (2008, Sect. 3, emphasis Joyce's) 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, Sect. 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),

¹⁰ Joyce (2008, Sect. 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 the following: 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 Brossel (2013), Crupi et al. (2010), Fitelson (2007, 2011), and Sober (2011).

(WLL*) is not met by all of measures (i)-(xiii).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:

$$(xiv) \quad D'(H, E) = \Pr(H|E) \Pr(H|\neg E) - \Pr(H) \Pr(\neg H).$$

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 $X(Hk, E) < X(Hk-1, E)$ for all $2 \leq k \leq n$?

¹¹ 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 $G(H, E) = \frac{\Pr(\neg H)}{\Pr(\neg H|E)} = \frac{\Pr(E)}{\Pr(E|\neg H)} = \frac{\Pr(E)}{\Pr(E|\neg H^*)} = \frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} = G(H^*, E)$. The case of Z is similar.

¹² The qualification “in the sense of (SOC)” is needed because we introduce a second screening-off condition in Subsection 4.1 below.

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

3. Main Result. The task now is to answer the question of dwindling confirmation under (SOC). We do this in Subsection 3.1. We then turn to testimony cases in Subsection 3.2, and to the Data-Processing Inequality in Subsection 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 $X(\text{Hk}, E) < X(\text{Hk}-1, E)$ for all $2 \leq k \leq n$.

(See Appendix 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) $\Pr(E \mid \text{Hk}-1) > \Pr(E \mid \text{Hk})$ and (ii) $\Pr(E \mid \neg\text{Hk}-1) < \Pr(E \mid \neg\text{Hk})$ for all $2 \leq k \leq n$. Theorem 1 follows immediately from (i) and (ii) by (WLL).

Each of measures (i)-(xiii) meets (WLL). So, Theorem 1 entails that confirmation dwindles under (SOC) on each of measures (i)-(xiii). It is well known that certain claims about confirmation hold on some measures but not on others (Brossel 2013 and Fitelson 1999). 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 $E = \text{Smith said that miracle } M \text{ occurred}$, $H1 = \text{miracle } M \text{ occurred}$, and $H2 = \text{God exists}$, is plausibly such a case. E increases the probability of $H1$ (at least on some ways of filling in the details of the case)—even if just barely. $H1$ in turn increases the probability of $H2$. E (we can suppose) neither entails nor is entailed by $H1$. $H1$ 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

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

one already knows whether miracle M occurred or not, 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 H_k is less than the degree to which E confirms H_{k-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 (i)-(xiii) differ in terms of how much 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 R(H1, E) and R(H2, H1) is very high, say, roughly equal to 100,000, but R(H2, 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, Z(H2, E) equals the product of Z(H1, E) and Z(H2, H1).¹⁵ So, if each of Z(H1, E) and Z(H2, H1) is very high, then it cannot be the case that Z(H2, E) is very low.

3.3. The Data-Processing Inequality. The Data-Processing Inequality concerns "mutual information." Let X and Y be discrete variables. The mutual information $I(X, Y)$ 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:

$$\sum_{x \in X} \sum_{y \in Y} \Pr(x \wedge y) \text{Log}_2 \left(\frac{\Pr(x \wedge y)}{\Pr(x)\Pr(y)} \right).$$

Here "x" and "y" are propositions specifying the states of X and Y respectively. Mutual information is symmetric in that $I(X, Y) = I(Y, X)$. Now let D , P , and E be discrete variables ("D" for "distal cause," "P" for "proximate cause," "E" for "effect") and

¹⁵ This is clear from (14) in Appendix B.

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 : $I(E, D) < I(E, P)$.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 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

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, Ch. 2) for explanation of mutual information and the Data-Processing Inequality (Theorem 2.8.1). Strictly speaking, in order for $I(E, D)$ to be less than $I(E, P)$, it also needs to be the case that $I(E, P | D)$ is greater than 0. This condition should be understood throughout this subsection.

¹⁸ See Sober and Steel (unpublished) 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.

¹⁹ Terminology varies. Fano (1961), for example, 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.

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 $\text{Log}_2 \left(\frac{\text{Pr}(x \wedge y)}{\text{Pr}(x)\text{Pr}(y)} \right)$. 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 Subsection 4.1) concerns a screening-off condition weaker than (SOC). The second respect (discussed in Subsection 4.2) concerns cases where condition (C)—the condition that E neither entails nor is entailed by $H1$, and H_{k-1} neither entails nor is entailed by H_k 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). $\text{Pr}(H_k | H_{k-1} \wedge E) \leq \text{Pr}(H_k | H_{k-1})$ and $\text{Pr}(H_k | \neg H_{k-1} \wedge E) \leq \text{Pr}(H_k | \neg H_{k-1})$.

(PISOC) says in effect that given the truth or falsity of $H1$, E has no *positive* impact on the probability of $H2$ —either E has no impact on the probability of $H2$ 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:

²⁰ Recall the first card case given in Section 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

(B*) E is screened off in the sense of (PISOC) from H_k by H_{k-1} for all $2 \leq k \leq n$.

(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 Appendix 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

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

Jack, and H_2 = the card drawn is a Diamond. We used the case as a counterexample to transitivity, but (PISOC) holds in it; $\Pr(H_2 | H_1 \wedge E) = 0 < 1/2 = \Pr(H_2 | H_1)$ and $\Pr(H_2 | \neg H_1 \wedge E) = 0 < 6/25 = \Pr(H_2 | \neg H_1)$. So, (PISOC) is not sufficient for transitivity in confirmation.

²¹ There is yet a third screening-off condition:

Negative Impact Screening-Off Condition (NISOC). $\Pr(H_2 | H_1 \wedge E) \geq \Pr(H_2 | H_1)$ and $\Pr(H_2 | \neg H_1 \wedge E) \geq \Pr(H_2 | \neg H_1)$.

(NISOC) is like (PISOC) in that given the truth or falsity of H_1 , E may still have some impact on the probability of H_2 . 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 2012a). But it is not the case that confirmation dwindles under (NISOC) on all measures meeting (WLL). In fact, for each of measures (i)-(xiii), there are distributions of probabilities on which E confirms H_1 , H_1 confirms H_2 , (NISOC) holds, E neither entails nor is entailed by H_1 , H_1 neither entails nor is entailed by H_2 , and yet the confirmation provided to H_2 by E is greater than or equal to the confirmation provided by E to H_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 $x_1 = X$ has value 1, $x_2 = X$ has value 2, and similarly for y_1, y_2, z_1, z_2 . First, one can require that each member of the partition $\{x_1, x_2\}$ be screened off in the sense of (PISOC) by each member of the partition $\{y_1, y_2\}$ from each member of

4.2. *Non-Increasing Confirmation under (SOC) and (PISOC)*. Recall that each of the main confirmation measures in the literature—measures (i)-(xiii)—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 $R(H, E) = \frac{\Pr(H|E)}{\Pr(H)} = \frac{\Pr(E|H)}{\Pr(E)} \geq \frac{\Pr(E|H^*)}{\Pr(E)} = \frac{\Pr(H^*|E)}{\Pr(H^*)} = R(H^*, E)$. 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 and/or R. 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 Appendix 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 non-strict (“≥” 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 non-strict, this is not the case with (WLL*)—the first and third are strict while the second is non-strict. (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*

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, $\Pr(z1 | y1 \wedge x1) \leq \Pr(z2 | y1)$, then $\Pr(\neg z2 | y1 \wedge x1) \geq \Pr(\neg z2 | y1)$. So, $\Pr(\neg z2 | y1 \wedge x1) \leq \Pr(\neg z2 | y1)$ only if $\Pr(\neg z2 | y1 \wedge x1) = \Pr(\neg z2 | y1)$, 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 $I(X, Z)$.

it does H^* . Recall that not all of measures (i)-(xiii) 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 H_1 , and H_{k-1} neither entails nor is entailed by H_k 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 $X(H_k, E) \not\leq X(H_{k-1}, E)$ for some $2 \leq k \leq n$. However, we now have:

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

(See Appendix 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, $X(H_k, E) = X(H_{k-1}, E)$ 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):

Theorem 4. If (A) and (B^*) hold and X meets (WLL^{**}) , then $X(H_k, E) \leq X(H_{k-1}, E)$ 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 paper in non-formal terms: as the chain of probabilistic support becomes longer, confirmation dwindles provided 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 meeing (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) $1 < R(H^*, E) < R(H, E)$;
- (2) $0 < R(H^*, E) \leq 1 < R(H, E)$;
- (3) $0 < R(H^*, E) < 1 \leq R(H, E)$;
- (4) $0 < R(H^*, E) < R(H, E) < 1$.

We begin with cases (2) and (3), and then turn to cases (1) and (4).

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

$$J(H, E) = \frac{\text{Log}_2[R(H, E)]}{-\text{Log}_2[\text{Pr}(H)]}$$

$$J(H^*, E) = \frac{\text{Log}_2[R(H^*, E)]}{-\text{Log}_2[\text{Pr}(H^*)]}$$

Since $0 < -\text{Log}_2[\text{Pr}(H)]$ and $0 < -\text{Log}_2[\text{Pr}(H^*)]$, it follows from (2) and (3), respectively, that:

- (2*) $J(H^*, E) \leq 0 < J(H, E)$;
- (3*) $J(H^*, E) < 0 \leq J(H, E)$.

So, $J(H^*, E) < J(H, E)$ in cases (2) and (3).

Cases (1) and (4)—the two remaining cases—imply, respectively, that:

- (1*) $0 < \text{Log}_2[R(H^*, E)] < \text{Log}_2[R(H, E)]$;
- (4*) $\text{Log}_2[R(H^*, E)] < \text{Log}_2[R(H, E)] < 0$.

In both case (1) and case (4), either (a) $0 < \text{Pr}(H^*) \leq \text{Pr}(H) < 1$ or (b) $0 < \text{Pr}(H) < \text{Pr}(H^*) < 1$, from which it follows, respectively, that:

- (a*) $\text{Log}_2[\text{Pr}(H^*)] \leq \text{Log}_2[\text{Pr}(H)] < 0$;
- (b*) $\text{Log}_2[\text{Pr}(H)] < \text{Log}_2[\text{Pr}(H^*)] < 0$.

To show that $J(H^*, E) < J(H, E)$ in case (1), it suffices to show that $J(H^*, E) < J(H, E)$ in the two sub-cases: (1) and (a); (1) and (b). Likewise, to show that $J(H^*, E) < J(H, E)$ in case (4), it suffices to show that $J(H^*, E) < J(H, E)$ in the two sub-cases: (4) and (a); (4) and (b).

First sub-case: (1) and (a). It follows immediately from (1*) and (a*) that $J(H^*, E) < J(H, E)$. Second sub-case: (1) and (b). Let $\Pr(H^*) = \Pr(H)^\alpha$, where $0 < \alpha < 1$ from (b). $G(H^*, E) < G(H, E)$, so $\frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} < \frac{\Pr(\neg H)}{\Pr(\neg H|E)}$. Given this, and given that

$$\begin{aligned} \frac{\Pr(\neg H^*)}{\Pr(\neg H^*|E)} &< \frac{\Pr(\neg H)}{\Pr(\neg H|E)} \text{ iff} \\ \frac{1-\Pr(H|E)}{1-\Pr(H^*|E)} &< \frac{1-\Pr(H)}{1-\Pr(H^*)} \text{ iff} \\ \frac{1-\Pr(H)R(H,E)}{1-\Pr(H^*)R(H^*,E)} &< \frac{1-\Pr(H)}{1-\Pr(H^*)} \text{ iff} \\ \frac{1-\Pr(H)R(H,E)}{1-\Pr(H)^\alpha R(H^*,E)} &< \frac{1-\Pr(H)}{1-\Pr(H)^\alpha}, \end{aligned}$$

we have:

$$(5) \quad \frac{1-\Pr(H)R(H,E)}{1-\Pr(H)^\alpha R(H^*,E)} < \frac{1-\Pr(H)}{1-\Pr(H)^\alpha}.$$

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

$$(6) \quad \frac{1-\Pr(H)}{1-\Pr(H)^\alpha} < \frac{1-\Pr(H)R(H,E)}{1-[\Pr(H)R(H,E)]^\alpha}.$$

It follows from (5) and (6) that

$$(7) \quad \frac{1-\Pr(H)R(H,E)}{1-\Pr(H)^\alpha R(H^*,E)} < \frac{1-\Pr(H)R(H,E)}{1-[\Pr(H)R(H,E)]^\alpha}.$$

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

$$(8) \quad \text{Log}_2[R(H^*, E)] < \alpha \text{Log}_2[R(H, E)].$$

It follows from (8) and $0 < -\text{Log}_2[\Pr(H)]$ that

$$(9) \quad \frac{\text{Log}_2[R(H^*, E)]}{-\text{Log}_2[\Pr(H)]} < \frac{\alpha \text{Log}_2[R(H, E)]}{-\text{Log}_2[\Pr(H)]}.$$

(9) implies:

$$(10) \frac{\text{Log}_2[\text{R}(\text{H}^*, \text{E})]}{-\alpha \text{Log}_2[\text{Pr}(\text{H})]} < \frac{\text{Log}_2[\text{R}(\text{H}, \text{E})]}{-\text{Log}_2[\text{Pr}(\text{H})]}.$$

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

$$(11) \frac{\text{Log}_2[\text{R}(\text{H}^*, \text{E})]}{-\text{Log}_2[\text{Pr}(\text{H}^*)]} = \frac{\text{Log}_2[\text{R}(\text{H}^*, \text{E})]}{-\text{Log}_2[\text{Pr}(\text{H})^\alpha]} = \frac{\text{Log}_2[\text{R}(\text{H}^*, \text{E})]}{-\alpha \text{Log}_2[\text{Pr}(\text{H})]} < \frac{\text{Log}_2[\text{R}(\text{H}, \text{E})]}{-\text{Log}_2[\text{Pr}(\text{H})]}.$$

Hence $J(\text{H}^*, \text{E}) < J(\text{H}, \text{E})$ in case (1).

Third sub-case: (4) and (a). Let $\text{Pr}(\text{H}^*) = \text{Pr}(\text{H})^\alpha$, where $1 \leq \alpha$ from (a). (5) holds by the same reasoning as in the second sub-case. Meanwhile, it follows from $1 \leq \alpha$ that $\frac{1-x}{1-x^\alpha}$ is a strictly decreasing function of x . But $\text{Pr}(\text{H}1) > \text{Pr}(\text{H}1)\text{R}(\text{H}1, \text{E})$ from (4). So (6) holds. The argument then continues in parallel to the argument given in the second sub-case. Fourth sub-case: (4) and (b). It follows immediately from (4*) and (b*) that $J(\text{H}^*, \text{E}) < J(\text{H}, \text{E})$. Thus $J(\text{H}^*, \text{E}) < J(\text{H}, \text{E})$ in case (4).

So, as $J(\text{H}^*, \text{E}) < J(\text{H}, \text{E})$ in cases (1)-(4), it follows that J meets (WLL). ■

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) $\text{Pr}(\text{E} \mid \text{H}_{k-1}) > \text{Pr}(\text{E} \mid \text{H}_k)$ and (b) $\text{Pr}(\text{E} \mid \neg \text{H}_{k-1}) < \text{Pr}(\text{E} \mid \neg \text{H}_k)$ for all $2 \leq k \leq n$.

Second, we argue from Lemma to Theorem 1.

Consider a three-member series $\text{E}, \text{H}_1, \text{H}_2$ and suppose the antecedent of Lemma holds. It is known (Shogenji 2003) that if H_1 screens off E from H_2 , as (B) implies, then:

$$(12) \text{Pr}(\text{H}_2 \mid \text{E}) - \text{Pr}(\text{H}_2) = \frac{(\text{Pr}(\text{H}_1 \mid \text{E}) - \text{Pr}(\text{H}_1))(\text{Pr}(\text{H}_2 \mid \text{H}_1) - \text{Pr}(\text{H}_2))}{1 - \text{Pr}(\text{H}_1)}.$$

(12) implies:

$$(13) \frac{\text{Pr}(\text{H}_2 \mid \text{E}) - \text{Pr}(\text{H}_2)}{\text{Pr}(\text{H}_2)} = \left(\frac{\text{Pr}(\text{H}_1 \mid \text{E}) - \text{Pr}(\text{H}_1)}{\text{Pr}(\text{H}_1)} \right) \left(\frac{\text{Pr}(\text{H}_1)(\text{Pr}(\text{H}_2 \mid \text{H}_1) - \text{Pr}(\text{H}_2))}{\text{Pr}(\text{H}_2)(1 - \text{Pr}(\text{H}_1))} \right);$$

$$(14) \frac{\text{Pr}(\text{H}_2 \mid \text{E}) - \text{Pr}(\text{H}_2)}{1 - \text{Pr}(\text{H}_2)} = \frac{(\text{Pr}(\text{H}_1 \mid \text{E}) - \text{Pr}(\text{H}_1))(\text{Pr}(\text{H}_2 \mid \text{H}_1) - \text{Pr}(\text{H}_2))}{(1 - \text{Pr}(\text{H}_1))(1 - \text{Pr}(\text{H}_2))}.$$

But, given (C), we have:

$$(15) \Pr(H1)(\Pr(H2 | H1) - \Pr(H2)) = \Pr(H1 \wedge H2) - \Pr(H1)\Pr(H2) \\ < \Pr(H2) - \Pr(H2)(H1) = \Pr(H2)(1 - \Pr(H1)).$$

(13) and (15) together entail:

$$(16) \frac{\Pr(H2|E) - \Pr(H2)}{\Pr(H2)} < \frac{\Pr(H1|E) - \Pr(H1)}{\Pr(H1)}.$$

(16) implies:

$$(17) R(H2, E) = \frac{\Pr(H2|E)}{\Pr(H2)} < \frac{\Pr(H1|E)}{\Pr(H1)} = R(H1, E).$$

Next, from (C), we have:

$$(18) \Pr(H2 | H1) - \Pr(H2) < 1 - \Pr(H2).$$

(14) and (18) together imply:

$$(19) \frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} < \frac{\Pr(H1|E) - \Pr(H1)}{1 - \Pr(H1)}.$$

But, we have:

$$(20) \frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} < \frac{\Pr(H1|E) - \Pr(H1)}{1 - \Pr(H1)} \text{ iff} \\ \left(1 - \frac{\Pr(\neg H2|E)}{\Pr(\neg H2)}\right) < \left(1 - \frac{\Pr(\neg H1|E)}{\Pr(\neg H1)}\right) \text{ iff} \\ \frac{\Pr(\neg H2|E)}{\Pr(\neg H2)} > \frac{\Pr(\neg H1|E)}{\Pr(\neg H1)} \text{ iff} \\ \frac{\Pr(\neg H2)}{\Pr(\neg H2|E)} < \frac{\Pr(\neg H1)}{\Pr(\neg H1|E)}.$$

It follows from (19) and (20) that:

$$(21) G(H2, E) = \frac{\Pr(\neg H2)}{\Pr(\neg H2|E)} < \frac{\Pr(\neg H1)}{\Pr(\neg H1|E)} = G(H1, E)$$

So, if the antecedent of 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, H_1, H_2, \dots, H_n . Suppose the antecedent of Lemma holds. (SOC) is a condition sufficient for transitivity in confirmation, so, given (A) and (B), it follows that E confirms H_k for all $1 \leq k \leq n$. Next, given (A)-(C), and given the result from the previous paragraph, it follows that $R(H_k, E) < R(H_{k-1}, E)$ and $G(H_k, E) < G(H_{k-1}, E)$.

Lemma is but a short step away. $R(H, E) = \frac{\Pr(H|E)}{\Pr(H)} = \frac{\Pr(E|H)}{\Pr(E)}$ and $G(H, E) = \frac{\Pr(\neg H)}{\Pr(\neg H|E)} = \frac{\Pr(E)}{\Pr(E|\neg H)}$. So, as $R(H_k, E) < R(H_{k-1}, E)$ and $G(H_k, E) < G(H_{k-1}, E)$, we have:

$$(22) \quad \frac{\Pr(E|H_k)}{\Pr(E)} = \frac{\Pr(H_k|E)}{\Pr(H_k)} < \frac{\Pr(H_{k-1}|E)}{\Pr(H_{k-1})} = \frac{\Pr(E|H_{k-1})}{\Pr(E)};$$

$$(23) \quad \frac{\Pr(E)}{\Pr(E|\neg H_k)} = \frac{\Pr(\neg H_k)}{\Pr(\neg H_k|E)} < \frac{\Pr(\neg H_{k-1})}{\Pr(\neg H_{k-1}|E)} = \frac{\Pr(E)}{\Pr(E|\neg H_{k-1})}.$$

It follows immediately from (22) and (23), respectively, that $\Pr(E | H_{k-1}) > \Pr(E | H_k)$ and $\Pr(E | \neg H_{k-1}) < \Pr(E | \neg H_k)$. Lemma is thus established.

Suppose now the antecedent of Theorem 1 holds. Then, the antecedent of Lemma also holds. So, it follows by Lemma that (a) $\Pr(E | H_{k-1}) > \Pr(E | H_k)$ and (b) $\Pr(E | \neg H_{k-1}) < \Pr(E | \neg H_k)$. It then follows, as X meets (WLL), that $X(H_k, E) < X(H_{k-1}, E)$.

■

Appendix C: Proof of Theorem 2. Suppose the antecedent of Theorem 2 holds. Then, by reasoning similar to that for (12), we have:

$$(12^*) \Pr(H_2|E) - \Pr(H_2) \leq \frac{(\Pr(H_1|E) - \Pr(H_1))(\Pr(H_2|H_1) - \Pr(H_2))}{1 - \Pr(H_1)}.$$

It follows from (12*) that:

$$(13^*) \frac{\Pr(H_2|E) - \Pr(H_2)}{\Pr(H_2)} \leq \left(\frac{\Pr(H_1|E) - \Pr(H_1)}{\Pr(H_1)} \right) \left(\frac{\Pr(H_1)(\Pr(H_2|H_1) - \Pr(H_2))}{\Pr(H_2)(1 - \Pr(H_1))} \right);$$

$$(14^*) \frac{\Pr(H_2|E) - \Pr(H_2)}{1 - \Pr(H_2)} \leq \frac{(\Pr(H_1|E) - \Pr(H_1))(\Pr(H_2|H_1) - \Pr(H_2))}{(1 - \Pr(H_1))(1 - \Pr(H_2))}.$$

With (13*) and (14*), instead of (13) and (14), 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, H_1, H_2 , then $R(H_2, E) < R(H_1, E)$ and $G(H_2, E) < G(H_1, E)$. Next, it is shown that the point generalizes to the n-member series E, H_1, H_2, \dots, H_n and thus $\Pr(E | H_{k-1}) > \Pr(E | H_k)$ and $\Pr(E | \neg H_{k-1}) < \Pr(E | \neg H_k)$.

$\neg H_k$). Finally, since X meets (WLL), it follows from the antecedent of Theorem 2 that $X(H_k, E) < X(H_{k-1}, E)$. ■

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 (12), (13), and (14). Then, instead of (15) which is a strict inequality, we have:

$$(15^*) \Pr(H1)(\Pr(H2 | H1) - \Pr(H2)) = \Pr(H1 \wedge H2) - \Pr(H1)\Pr(H2) \\ \leq \Pr(H2) - \Pr(H2)(H1) = \Pr(H2)(1 - \Pr(H1)).$$

(13) and (15*) imply:

$$(16^*) \frac{\Pr(H2|E) - \Pr(H2)}{\Pr(H2)} \leq \frac{\Pr(H1|E) - \Pr(H1)}{\Pr(H1)}.$$

(16*) entails:

$$(17^*) R(H2, E) = \frac{\Pr(H2|E)}{\Pr(H2)} \leq \frac{\Pr(H1|E)}{\Pr(H1)} = R(H1, E).$$

Further, instead of (18) which is a strict inequality, we have:

$$(18^*) \Pr(H2 | H1) - \Pr(H2) \leq 1 - \Pr(H2).$$

(14) and (18*) together entail:

$$(19^*) \frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} \leq \frac{\Pr(H1|E) - \Pr(H1)}{1 - \Pr(H1)}.$$

But we have:

$$(20^*) \frac{\Pr(H2|E) - \Pr(H2)}{1 - \Pr(H2)} \leq \frac{\Pr(H1|E) - \Pr(H1)}{1 - \Pr(H1)} \text{ iff} \\ \left(1 - \frac{\Pr(\neg H2|E)}{\Pr(\neg H2)}\right) \leq \left(1 - \frac{\Pr(\neg H1|E)}{\Pr(\neg H1)}\right) \text{ iff} \\ \frac{\Pr(\neg H2|E)}{\Pr(\neg H2)} \geq \frac{\Pr(\neg H1|E)}{\Pr(\neg H1)} \text{ iff} \\ \frac{\Pr(\neg H2)}{\Pr(\neg H2|E)} \leq \frac{\Pr(\neg H1)}{\Pr(\neg H1|E)}.$$

It follows from (19*) and (20*) that:

$$(21^*)G(H2, E) = \frac{\Pr(\neg H2)}{\Pr(\neg H2|E)} \leq \frac{\Pr(\neg H1)}{\Pr(\neg H1|E)} = G(H1, E).$$

So if (A) and (B) hold for a three-member series $E, H1, H2$, then $R(H2, E) \leq R(H1, E)$ and $G(H2, E) \leq G(H1, E)$. 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 $E, H1, H2, \dots, Hn$, and thus $\Pr(E | Hk-1) \geq \Pr(E | Hk)$ and $\Pr(E | \neg Hk-1) \leq \Pr(E | \neg Hk)$. Then, from the antecedent of Theorem 3, which says in part that X meets (WLL**), it follows that $X(Hk, E) \leq X(Hk-1, E)$. ■

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.
- Cover, Thomas, and Joy Thomas. 2006. *Elements of Information Theory* (2nd ed.). Hoboken: John Wiley & Sons.
- 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, Mass.: 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 *The Stanford Encyclopedia of Philosophy* (Fall 2008 ed.), ed. Edward Zalta. URL = <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. "A Weaker Condition for Transitivity in Probabilistic Support." *European Journal for Philosophy of Science* 2: 111-18.
- 2012b. "Transitivity and Intransitivity in Evidential Support: Some Further Results." *Review of Symbolic Logic* 5: 259-68.
- 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. unpublished. "Time and Knowability in Evolutionary Processes."
- 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 Publishing.