When Confirmation Transmits

Stephen Yablo[1] Branden Fitelson[2]

[1]Philosophy @ MIT

yablo@mit.edu

[2]Philosophy & Religion @ Northeastern

branden@fitelson.org

- Carnap [1] discusses 2 types of (probabilistic) confirmation.
  - **Firmness**. E confirms$_f$ H iff

$$\Pr(H \mid E) > t, \text{ where } t \geq 1/2.$$

  - **Increase in Firmness**. E confirms$_i$ H iff

$$\Pr(H \mid E) > \Pr(H).$$

- Carnap also proposed (tentatively) a particular way of measuring *the degree to which* E confirms$_i$ H:

$$d(H,E) \overset{\text{def}}{=} \Pr(H \mid E) - \Pr(H).$$

- Confirms$_f$ & confirms$_i$ exhibit many theoretical *divergences* [15, 5, 10]. One of the most important of these divergences involves Hempel's [13] *Special Consequence Condition*.

(SCC) If E confirms H_1 and $H_1 \vDash H_2$, then E confirms H_2.

- Carnap [1, Ch. VI] discusses the fact that confirms$_f$ (generally) satisfies (SCC); but, confirms$_i$ does not.

- Following Dretske [7], we may say that an epistemic operator $\mathcal{O}(H,E)$ is a *penetrating operator* just in case $\mathcal{O}(H,E)$ is always transmitted by deductive entailment.
- Hempel's (SCC) asserts that confirmation $C(H,E)$ is a penetrating operator. Carnap shows that firmness $C_f(H,E)$ is penetrating, while increase in firmness $C_i(H,E)$ is not.
- Dretske thought *knowledge* was not a penetrating operator (*viz.*, that knowledge isn't closed under entailment).
- We will take no stand on knowledge closure here. But, it is worth noting that confirmation [$C_i(H,E)$] is a *propositional* relation, whereas knowledge is a *doxastic* relation (*e.g.*, for one thing, E may not capture the agent's *total* evidence).
- Having said that, our discussion may be of some relevance to these broader epistemic questions, since some (putative) failures of knowledge transmission may involve (*i.e.*, implicitly trade on) failures of $C_i(H,E)$-transmission.

- Dretske [7] discusses an example he thinks shows that knowledge is not a penetrating operator.

  Zebra. You're at the zoo, and in the pen in front of you is a striped horse-like animal (which happens to be a zebra). The sign on the pen says "Zebra." Do you know it's a zebra?

- Dretske says: Well, what about the possibility that it's just a mule painted to look like a zebra? Do you know that the animal is not a cleverly-disguised mule?
- Let $E \overset{\text{def}}{=}$ your perceptual evidence (from observing the animal in the pen), $H_1 \overset{\text{def}}{=}$ the animal before you is a zebra, and $H_2 \overset{\text{def}}{=}$ the animal before you is not a cleverly-disguised mule.
- Dretske seems to be suggesting (among other things) that, while E confirms H_1 and $H_1 \vDash H_2$, E does *not* confirm H_2. At least: E *does not favor* H_1 *over* $\neg H_2$ (and *vice versa*).
- This basic Dretskean intuition leads to a simple sufficient condition for confirmation$_i$-transmission *failure*.

Fact. Suppose E confirms$_i$ H_1 and $H_1 \vDash H_2$. Then, the following is a sufficient condition for the *failure* of confirmation$_i$-transmission (*i.e.*, for E to *not* confirm$_i$ H_2).

Heavyweight. $\Pr(E \mid H_1) = \Pr(E \mid \neg H_2)$.

- **Heavyweight** is a natural way to explicate the claim that evidence E does not favor H_1 over $\neg H_2$ and *vice versa* [2].
- This way of understanding what Dretske means by "$\neg H_2$ is a heavyweight proposition" [6] is somewhat crude.
- For one thing, if E confirms$_i$ H_1, then **Heavyweight** entails that E *dis*confirms$_i$ H_2 — *whether or not* $H_1 \vDash H_2$.
- This makes **Heavyweight** not super interesting (for us). More interesting: conditions which (a) trade on $H_1 \vDash H_2$, and (b) are compatible with E being *irrelevant to* H_2.
- We will examine some more interesting conditions (in these and other senses) shortly. First, we will discuss some other ways in which confirmation$_i$-transmission can fail.

- Here is a counterexample to confirmation$_i$-transmission that has a different structure than Dretske's **Zebra** case.

Ace. You are going to draw a single card at random from a standard deck. Let $E \overset{\text{def}}{=}$ the card is black, $H_1 \overset{\text{def}}{=}$ the card is the ace of spades, and $H_2 \overset{\text{def}}{=}$ the card is an ace.

- In **Ace**, E confirms$_i$ H_1, since $\Pr(H_1 \mid E) = 1/26 > 1/52 = \Pr(H_1)$. Moreover, $H_1 \vDash H_2$. However, E is *irrelevant to* H_2, since $\Pr(H_2 \mid E) = 2/26 = 4/52 = \Pr(H_2)$. [Note: $H_1 \vDash E$ in **Ace**.]
- Much more *extreme* failures of confirmation$_i$ transmission are possible. To wit, there are cases such that (see Extras 14)

(1) E *strongly* confirms$_i$ H_1 [$d(H_1, E) \gg 0$].[1]

(2) $H_1 \vDash H_2$ [more precisely, $\Pr(H_2 \mid H_1) = 1$].

(3) E *strongly dis*confirms$_i$ H_2 [$d(H_2, E) \ll 0$].

[1] There are *limits* on *how badly* (SCC) can fail (in this sense). Specifically, if we understand $x \gg y$ as $x - y \geq t$, then we must have $t < 1/2$ in (1) & (3).

- We will say that a probabilistic condition $\mathcal{X}$ is *sufficient* for confirmation$_i$-transmission, just in case the following holds.

Sufficiency. There are no probability functions $\Pr(\cdot)$ s.t.

(S_1) $\Pr(H_1 \,\&\, \neg H_2) = 0$, and these are *the only zeros* of $\Pr(\cdot)$.

(S_2) $\Pr(H_1 \mid E) > \Pr(H_1)$. [$E$ confirms$_i$ H_1, wrt $\Pr(\cdot)$]

(S_3) $\Pr(\cdot)$ satisfies $\mathcal{X}$.

(S_4) $\Pr(H_2 \mid E) \leq \Pr(H_2)$. [$E$ does not confirm$_i$ H_2, wrt $\Pr(\cdot)$]

- Kotzen [14] has an illuminating discussion of confirmation$_i$ transmission in which he identifies the following sufficient condition for confirmation$_i$ transmission.

Dragging. $\Pr(H_2) < \Pr(H_1 \mid E)$.

- It is easy to see why **Dragging** is sufficient for transmission.

Proof. (S_1) implies $\Pr(H_2 \mid E) \geq \Pr(H_1 \mid E)$. ($S_2$) and **Dragging** then imply $\Pr(H_2 \mid E) > \Pr(H_2)$, which contradicts ($S_4$). □

- We discovered the following sufficient condition, which is independent of Kotzen's **Dragging** condition — even in the presence of (S_1) and (S_2) above (see Extras 15–16).

Non-confirmation of Exhaustive Alternatives (NEA). *E does not confirm$_i$ $H_2 \supset H_1$* [*viz.*, $d(H_2 \supset H_1, E) \leq 0$].

- We call this **Non-confirmation of Exhaustive Alternatives** because it involves the non-confirmation of a claim which asserts that $\neg H_2$ and H_1 are *exhaustive alternatives*.
- For instance, in **Zebra**, $H_2 \supset H_1$ asserts that the animal before you is *either* a cleverly-disguised mule *or* a zebra.
- In **Zebra**, whether E supports the exhaustivity of H_1 and $\neg H_2$ (as alternative hypotheses) seems probative (perhaps this relates to whether $\neg H_2$ is a "relevant alternative"?).
- Anyhow, in **Zebra**, E *may not* confirm$_i$ $H_2 \supset H_1$. And, *if* it doesn't, *then* it turns out that E *must* (also) *confirm*$_i$ H_2.

- A probabilistic condition $\mathcal{X}$ (*e.g.*, ¬**Heavyweight**) is *necessary* for confirmation$_i$ transmission just in case

  Necessity. There are no probability functions $\Pr(\cdot)$ s.t.

  (S_1) $\Pr(H_1 \,\&\, \neg H_2) = 0$, and these are *the only zeros* of $\Pr(\cdot)$.

  (S_2) $\Pr(H_1 \mid E) > \Pr(H_1)$. [$E$ confirms$_i$ H_1, wrt $\Pr(\cdot)$]

  $\neg(S_3)$ $\Pr(\cdot)$ does *not* satisfy $\mathcal{X}$.

  $\neg(S_4)$ $\Pr(H_2 \mid E) > \Pr(H_2)$. [$E$ confirms$_i$ H_2, wrt $\Pr(\cdot)$]

- Kotzen [14, *p*. 70] voices skepticism about the existence of an interesting *necessary and sufficient* condition for confirmation$_i$-transmission. We think we've found one.

  Relative Disconfirmation of Exhaustive Alternatives (RDEA). E confirms$_i$ H_1 *more strongly than* E confirms$_i$ $H_2 \supset H_1$, according to Carnap's d [*i.e.*, $d(H_1, E) > d(H_2 \supset H_1, E)$].

☞ The confirmation E provides for H_1 transmits to H_2 ***iff*** E raises H_1's probability *more* (as measured by d) than it does the claim that H_1 and $\neg H_2$ are exhaustive alternatives.

- The fact that (RDEA) is necessary and sufficient for transmission of confirmation$_i$ is a corollary of the following general, quantitative result (see Extras 12 for a proof of it).

  Theorem. If $\Pr(H_2 \mid H_1) = 1$, then

$$d(H_2, E) = d(H_1, E) - d(H_2 \supset H_1, E).$$

- **Theorem** implies both (i) (RDEA) $\Longleftrightarrow$ transmission and (ii) (NEA) $\Longrightarrow$ transmission, and it (iii) gives the *d-degree* to which H_2 is confirmed$_i$ by E, *whenever* $H_1 \vDash H_2$.
- This result — and its qualitative corollary — *depends on how we choose to measure degree of confirmation$_i$*. Specifically, here are 4 other measures of degree of confirmation$_i$ [11, 4].

$$r(H,E) \overset{\text{def}}{=} \frac{\Pr(H|E)}{\Pr(H)} \doteq \frac{\Pr(H|E)+\Pr(H)}{\Pr(H|E)-\Pr(H)}$$

$$l(H,E) \overset{\text{def}}{=} \frac{\Pr(E|H)}{\Pr(E|\neg H)} \doteq \frac{\Pr(E|H)+\Pr(E|\neg H)}{\Pr(E|H)-\Pr(E|\neg H)}$$

$$s(H,E) \overset{\text{def}}{=} \Pr(H \mid E) - \Pr(H \mid \neg E)$$

$$z(H,E) \overset{\text{def}}{=} \begin{cases} \frac{d(H,E)}{\Pr(\neg H)} & \text{if } d(H,E) \geq 0 \\ \frac{d(H,E)}{\Pr(H)} & \text{if } d(H,E) < 0 \end{cases}$$

- Here's a summary of which measures $\mathfrak{c}(H,E)$ of the degree to which E confirms$_i$ H imply sufficiency/necessity of

  $(\text{RDEA}_\mathfrak{c})$ $\mathfrak{c}(H_1, E) > \mathfrak{c}(H_2 \supset H_1, E)$

  for confirmation$_i$-transmission.

$\mathfrak{c}$	Is (RDEA$_\mathfrak{c}$) Sufficient?	Is (RDEA$_\mathfrak{c}$) Necessary?
d	Yes	Yes
r	No	Yes
z	Yes	No
l	No	No
s	Yes	Yes

- The s-measure [3, 9, 8] also satisfies our quantitative **Theorem** (see Extras 13). See Extras 17–18 for probability models establishing the four "No"s in the above table.

  [Our proofs of the "Yes"s for (RDEA$_r$)/(RDEA$_z$) are complex (omitted).]

- To prove our results, we'll use the following algebraic representation, and the approach described in [12].

State ($\mathfrak{s}_i$)	H_1	H_2	E	$\Pr(\mathfrak{s}_i)$
$\mathfrak{s}_1$	T	T	T	$\Pr(\mathfrak{s}_1) = \mathfrak{a}_1$
$\mathfrak{s}_2$	T	T	F	$\Pr(\mathfrak{s}_2) = \mathfrak{a}_2$
$\mathfrak{s}_3$	T	F	T	$\Pr(\mathfrak{s}_3) = \mathfrak{a}_3$
$\mathfrak{s}_4$	T	F	F	$\Pr(\mathfrak{s}_4) = \mathfrak{a}_4$
$\mathfrak{s}_5$	F	T	T	$\Pr(\mathfrak{s}_5) = \mathfrak{a}_5$
$\mathfrak{s}_6$	F	T	F	$\Pr(\mathfrak{s}_6) = \mathfrak{a}_6$
$\mathfrak{s}_7$	F	F	T	$\Pr(\mathfrak{s}_7) = \mathfrak{a}_7$
$\mathfrak{s}_8$	F	F	F	$\Pr(\mathfrak{s}_8) = \mathfrak{a}_8$

- This involves (a) translating the desired result into algebra, and (b) showing it corresponds to a theorem of algebra (or that it does *not*), assuming $\mathfrak{a}_i \in [0,1]$ and $\sum_i \mathfrak{a}_i = 1$.

- If $\Pr(H_2 \mid H_1) = 1$, then $\alpha_3 = \alpha_4 = 0$. And, we have:

$$d(H_2, E) = \frac{\alpha_1 + \alpha_5}{\alpha_1 + \alpha_5 + \alpha_7} - (\alpha_1 + \alpha_2 + \alpha_5 + \alpha_6).$$

$$d(H_1, E) = \frac{\alpha_1}{\alpha_1 + \alpha_5 + \alpha_7} - (\alpha_1 + \alpha_2).$$

$$d(H_2 \supset H_1, E) = \frac{\alpha_1 + \alpha_7}{\alpha_1 + \alpha_5 + \alpha_7} - (1 - (\alpha_5 + \alpha_6))$$

- Then, the following reasoning establishes our **Theorem**:

$$\begin{aligned} d(H_1, E) - d(H_2 \supset H_1, E) &= \left[1 - \frac{\alpha_7}{\alpha_1 + \alpha_5 + \alpha_7}\right] - (\alpha_1 + \alpha_2 + \alpha_5 + \alpha_6) \\ &= [1 - \Pr(\neg H_1 \,\&\, \neg H_2 \mid E)] - \Pr(H_2) \\ &= \Pr(H_1 \vee H_2 \mid E) - \Pr(H_2) \\ &= \frac{\alpha_1 + \alpha_5}{\alpha_1 + \alpha_5 + \alpha_7} - (\alpha_1 + \alpha_2 + \alpha_5 + \alpha_6) \\ &= d(H_2, E) \quad \square \end{aligned}$$

- If $\Pr(H_2 \mid H_1) = 1$, then $\alpha_3 = \alpha_4 = 0$. And, we have:

$$s(H_2, E) = \frac{\alpha_1 + \alpha_5}{\alpha_1 + \alpha_5 + \alpha_7} - \frac{\alpha_2 + \alpha_6}{\alpha_2 + \alpha_6 + \alpha_8}.$$

$$s(H_1, E) = \frac{\alpha_1}{\alpha_1 + \alpha_5 + \alpha_7} - \frac{\alpha_2}{\alpha_2 + \alpha_6 + \alpha_8}.$$

$$s(H_2 \supset H_1, E) = \frac{\alpha_1 + \alpha_7}{\alpha_1 + \alpha_5 + \alpha_7} - \frac{\alpha_2 + \alpha_8}{\alpha_2 + \alpha_6 + \alpha_8}$$

- Then, the following establishes the s-version of **Theorem.**

$$\begin{aligned} s(H_1, E) - s(H_2 \supset H_1, E) &= \frac{\alpha_1}{\alpha_1 + \alpha_5 + \alpha_7} - \frac{\alpha_2}{\alpha_2 + \alpha_6 + \alpha_8} - \frac{\alpha_1 + \alpha_7}{\alpha_1 + \alpha_5 + \alpha_7} + \frac{\alpha_2 + \alpha_8}{\alpha_2 + \alpha_6 + \alpha_8} \\ &= \frac{\alpha_8}{\alpha_2 + \alpha_6 + \alpha_8} - \frac{\alpha_7}{\alpha_1 + \alpha_5 + \alpha_7} \\ &= \Pr(\neg H_1 \,\&\, \neg H_2 \mid \neg E) - \Pr(\neg H_1 \,\&\, \neg H_2 \mid E) \\ &= [1 - \Pr(H_1 \vee H_2 \mid \neg E)] - [1 - \Pr(H_1 \vee H_2 \mid E)] \\ &= \frac{\alpha_1 + \alpha_5}{\alpha_1 + \alpha_5 + \alpha_7} - \frac{\alpha_2 + \alpha_6}{\alpha_2 + \alpha_6 + \alpha_8} \\ &= s(H_2, E) \quad \square \end{aligned}$$

- Here is a model (all models were found with PrSAT [12]) on which (S_1), $d(H_1, E) = 0.49$ and $d(H_2, E) = -0.49$. This is about as extreme a failure of (SCC) as possible (see *fn.* 1).

State ($\mathfrak{s}_i$)	H_1	H_2	E	$\Pr(\mathfrak{s}_i)$
$\mathfrak{s}_1$	T	T	T	$\Pr(\mathfrak{s}_1) = \frac{450}{57600}$
$\mathfrak{s}_2$	T	T	F	$\Pr(\mathfrak{s}_2) = \frac{126}{57600}$
$\mathfrak{s}_3$	T	F	T	$\Pr(\mathfrak{s}_3) = 0$
$\mathfrak{s}_4$	T	F	F	$\Pr(\mathfrak{s}_4) = 0$
$\mathfrak{s}_5$	F	T	T	$\Pr(\mathfrak{s}_5) = \frac{1}{57600}$
$\mathfrak{s}_6$	F	T	F	$\Pr(\mathfrak{s}_6) = \frac{56511}{57600}$
$\mathfrak{s}_7$	F	F	T	$\Pr(\mathfrak{s}_7) = \frac{449}{57600}$
$\mathfrak{s}_8$	F	F	F	$\Pr(\mathfrak{s}_8) = \frac{63}{57600}$

- Here is a probability model on which (S_1), (S_2), and (NEA) are true, but **Dragging** is false (this shows NEA $\not\vDash$ **Dragging**).

State ($\mathfrak{s}_i$)	H_1	H_2	E	$\Pr(\mathfrak{s}_i)$
$\mathfrak{s}_1$	T	T	T	$\Pr(\mathfrak{s}_1) = \frac{256}{512}$
$\mathfrak{s}_2$	T	T	F	$\Pr(\mathfrak{s}_2) = \frac{28}{512}$
$\mathfrak{s}_3$	T	F	T	$\Pr(\mathfrak{s}_3) = 0$
$\mathfrak{s}_4$	T	F	F	$\Pr(\mathfrak{s}_4) = 0$
$\mathfrak{s}_5$	F	T	T	$\Pr(\mathfrak{s}_5) = \frac{64}{512}$
$\mathfrak{s}_6$	F	T	F	$\Pr(\mathfrak{s}_6) = \frac{5}{512}$
$\mathfrak{s}_7$	F	F	T	$\Pr(\mathfrak{s}_7) = \frac{128}{512}$
$\mathfrak{s}_8$	F	F	F	$\Pr(\mathfrak{s}_8) = \frac{31}{512}$

- Here is a probability model on which (S_1), (S_2), and **Dragging** are true, but (NEA) is false (this shows **Dragging** $\nvDash$ NEA).

State ($\mathfrak{s}_i$)	H_1	H_2	E	$\Pr(\mathfrak{s}_i)$
$\mathfrak{s}_1$	T	T	T	$\Pr(\mathfrak{s}_1) = \frac{128}{256}$
$\mathfrak{s}_2$	T	T	F	$\Pr(\mathfrak{s}_2) = \frac{5}{256}$
$\mathfrak{s}_3$	T	F	T	$\Pr(\mathfrak{s}_3) = 0$
$\mathfrak{s}_4$	T	F	F	$\Pr(\mathfrak{s}_4) = 0$
$\mathfrak{s}_5$	F	T	T	$\Pr(\mathfrak{s}_5) = \frac{12}{256}$
$\mathfrak{s}_6$	F	T	F	$\Pr(\mathfrak{s}_6) = \frac{10}{256}$
$\mathfrak{s}_7$	F	F	T	$\Pr(\mathfrak{s}_7) = \frac{64}{256}$
$\mathfrak{s}_8$	F	F	F	$\Pr(\mathfrak{s}_8) = \frac{37}{256}$

- Here is a probability model on which (S_1), (S_2), (RDEA_r), (RDEA_l), and (S_4) are all true. This shows that *neither* (RDEA_r) *nor* (RDEA_l) is sufficient for transmission.

State ($\mathfrak{s}_i$)	H_1	H_2	E	$\Pr(\mathfrak{s}_i)$
$\mathfrak{s}_1$	T	T	T	$\Pr(\mathfrak{s}_1) = \frac{64}{512}$
$\mathfrak{s}_2$	T	T	F	$\Pr(\mathfrak{s}_2) = \frac{5}{512}$
$\mathfrak{s}_3$	T	F	T	$\Pr(\mathfrak{s}_3) = 0$
$\mathfrak{s}_4$	T	F	F	$\Pr(\mathfrak{s}_4) = 0$
$\mathfrak{s}_5$	F	T	T	$\Pr(\mathfrak{s}_5) = \frac{256}{512}$
$\mathfrak{s}_6$	F	T	F	$\Pr(\mathfrak{s}_6) = \frac{45}{512}$
$\mathfrak{s}_7$	F	F	T	$\Pr(\mathfrak{s}_7) = \frac{128}{512}$
$\mathfrak{s}_8$	F	F	F	$\Pr(\mathfrak{s}_8) = \frac{14}{512}$

- Here is a probability model on which (S_1), (S_2), $\neg(\text{RDEA}_l)$, $\neg(\text{RDEA}_z)$, and $\neg(S_4)$ are all true. This shows that *neither* (RDEA_l) *nor* (RDEA_z) is necessary for transmission.

State ($\mathfrak{s}_i$)	H_1	H_2	E	$\Pr(\mathfrak{s}_i)$
$\mathfrak{s}_1$	T	T	T	$\Pr(\mathfrak{s}_1) = \frac{14}{32}$
$\mathfrak{s}_2$	T	T	F	$\Pr(\mathfrak{s}_2) = \frac{2}{32}$
$\mathfrak{s}_3$	T	F	T	$\Pr(\mathfrak{s}_3) = 0$
$\mathfrak{s}_4$	T	F	F	$\Pr(\mathfrak{s}_4) = 0$
$\mathfrak{s}_5$	F	T	T	$\Pr(\mathfrak{s}_5) = \frac{8}{32}$
$\mathfrak{s}_6$	F	T	F	$\Pr(\mathfrak{s}_6) = \frac{3}{32}$
$\mathfrak{s}_7$	F	F	T	$\Pr(\mathfrak{s}_7) = \frac{4}{32}$
$\mathfrak{s}_8$	F	F	F	$\Pr(\mathfrak{s}_8) = \frac{1}{32}$

[1] R. Carnap, *Logical Foundations of Probability* (2ed), 1962.

[2] J. Chandler, “Contrastive confirmation: some competing accounts,” 2013.

[3] D. Christensen, “Measuring Confirmation,” 1999.

[4] V. Crupi and K. Tentori, “Confirmation Theory,” 2016.

[5] V. Crupi, B. Fitelson and K. Tentori, “Probability, Confirmation, and the Conjunction Fallacy,” 2008.

[6] F. Dretske, “Is Knowledge Closed Under Known Entailment?,” 2005.

[7] ——, “Epistemic Operators,” 1970.

[8] E. Eells and B. Fitelson, “Symmetries and Asymmetries in Evidential Support,” 2002.

[9] ——, “Measuring Confirmation and Evidence,” 2000.

[10] B. Fitelson, “Confirmation, Causation, and Simpson’s Paradox,” 2017.

[11] ——, “Contrastive Bayesianism,” 2012.

[12] ——, “A Decision Procedure for Probability Calculus with Applications,” 2008.

[13] C. Hempel, “Studies in the Logic of Confirmation,” 1945.

[14] M. Kotzen, “Dragging and confirming,” 2012.

[15] W. Salmon, “Confirmation and relevance,” 1983.