Counterfactuals, Probabilities, and Information: Response to Critics

Abstract

In earlier work we proposed an account of information grounded in counterfactual conditionals rather than probabilities, and argued that it might serve philosophical needs that more familiar probabilistic alternatives do not. Demir (2008) and Scarantino (2008) criticize the counterfactual approach by contending that its alleged advantages are illusory and that it fails to secure attractive desiderata. In this paper we defend the counterfactual account from these criticisms, and suggest that it remains a useful account of information.

1 Background

In an earlier paper (C&M), we offered the following theory of information grounded in counterfactual conditionals:

(S*) Information relations are constituted by the non-vacuous truth of counterfactuals connecting the informational relata. Thus, x's being F carries information about y's being G if and only if the counterfactual conditional ⌜if y were not G, then x would not have been F⌝ is non-vacuously true.

We compared this view against more familiar probabilistic understandings of information (Dretske, 1981; Shannon, 1948), and argued that the counterfactual account might meet some needs that probabilistic accounts do not.

In a pair of recent papers, Demir (2008) and Scarantino (2008) criticize the counterfactual approach by contending that its alleged advantages are illusory and that it fails to secure important desiderata.

Before we assess these criticisms, we want to address a potential misreading. Demir and Scarantino present us as advocating the counterfactual theory over the probabilistic theories, but this misrepresents our intentions.[1] We do not claim that the counterfactual account is the only adequate theory of information; rather, we think it is a genuine alternative to the probabilistic account

---

[1] Demir's footnote 1, added after discussion with us, notes this possible misrepresentation. We hope the present comments clarify our intentions.

that may prove useful for certain philosophical purposes. Indeed, Demir's and Scarantino's critical observations contribute to our project by pointing to further features of the account that will aid in assessing its merits and drawbacks.

In the remainder of the paper we'll consider and respond to specific objections raised by our critics. We hope to show that the counterfactual approach remains a potentially useful conception of information.

2   The Conjunction Principle

Perhaps the most important criticism Demir raises against the counterfactual theory of information turns on his allegation that the view does not satisfy the Conjunction Principle — the claim that "if a signal r carries the information B and if it also carries the information that C, then it has to carry the information B and C ...." He writes:

> ... assume that A carries the information that B and A also carries the information that C. According to their definition, A carries the information that B if the counterfactual ⌜if B were not the case, then A would not have been the case⌝ is non-vacuously true. When this definition is applied to two assumptions, one gets the following counterfactual claims:
>
> **(1)** ⌜If B were not the case, then A would not have been the case⌝ is true.
>
> **(2)** ⌜If C were not the case, then A would not have been the case⌝ is true.
>
> Now, the question is whether these two necessarily imply the following: ⌜If B were not the case and C were not the case, then A would not have been the case⌝ is true (52).

How serious a flaw is it for a theory of information to violate the Conjunction Principle? While there is surely some intuitive support for the principle, we don't believe that its vindication is mandatory. Because 'information' is a technical term, intuitions about (presumably) non-technical understandings of information are not appropriate for assessing proposals.[2] We think a more sound methodology would require would-be users of such theories to say which intuitions they want captured (for specific purposes), and then to ask which (if any) of the proffered candidate theories will do the job.[3] It is in this spirit that we put forward the counterfactual account.

---

[2]Moreover, the usual problems about the authority of intuitions (e.g., as pressed by Swain *et al.* (2008)) apply here in spades.

[3]Don't we argue against probabilistic theories on intuitive grounds — viz., that they don't limit application of the Xerox Principle and, consequently, require (counterintuitively) that the probabilities underwriting information relations are unity? In principle we're flexible here, too. If you accept the requirement of probability one, then the drawback mentioned is no longer a reason to disfavor probabilistic approaches. Our claim is conditional: if the requirement of probability one is unreasonable in relevant theoretical contexts (as many philosophers have held) then probabilistic

That said, we claim that the counterfactual account respects the Conjunction Principle; we suspect Demir has convinced himself otherwise by asking about the wrong counterfactual. Given (S*), the question relevant to the assessment of the Conjunction Principle is whether (1) and (2) necessarily imply:

(3) $\ulcorner$If $(B$ and $C)$ were not the case, then A would not have been the case$\urcorner$ is true.

But (3) does indeed follow from (1) and (2). Proof: $\ulcorner\neg(B\&C)\urcorner$ is equivalent to $\ulcorner\neg B$ or $\neg C\urcorner$, so A carries information that $B\&C$ iff $\ulcorner$if not-B or not-C were true then A would not be the case$\urcorner$ is true.[4] Q.E.D. Thus, Demir's main complaint against the counterfactual account is misplaced.

3 Inverse Probabilities

Demir argues that one of the motivations we offered in support of the counterfactual account — its independence on inverse conditional probabilities to which the usual philosophical accounts of probability are inapplicable (Loewer, 1983; Humphreys, 1985) — is unpersuasive. For, he claims, the probabilities in question can be recovered by the probabilistic theory of Shannon (1948) (contrary to our footnote 6). On Shannon's view, the mutual information between a source s and a receiver r, $I(s,r) = -\sum_{i=1}^{n} P(s_i) \log_2 P(s_i) + \sum_{j=1}^{k} P(r_j) \sum_{i=1}^{n} P(s_i|r_j) \log_2 P(s_i|r_j)$.[5] So defined, mutual information is symmetric: $I(s,r) = I(r,s)$. Consequently, Shannon can provide a value for $I(r,s)$ without relying on inverse conditional probabilities.

We believe Demir is correct in all of this, and that we were too quick in attempting to generalize the Loewer-inspired worry to all probabilistic accounts.

However, this may not save probabilistic theories of information after all. This is because the very feature of Shannon's account that Demir appeals to — its symmetry — presents a challenge to its use in many of the purposes to which philosophers have put the notion of information. Arguably, that is, we don't *want* a symmetric notion of information to figure in our explanation of mental content, knowledge, and the like, in so far as these target notions are non-symmetric (if my thought is about cows, it shouldn't follow that cows are about my thought; likewise for knowledge, etc.).[6] (Although it may be possible to generate adequate accounts of such notions by appending

---

theories won't suffice. What such considerations amount to is not a debate about what could count as a theory of information, but an open-eyed assessment of the advantages and costs of theoretical alternatives.

[4]The world Demir considers (the closest not-B and not-C world) is relevant to the evaluation of his counterfactual, but not to (3).

[5]Here $s_1, \ldots, s_n$ are discrete alternative states of s with probabilities $P(s_1), \ldots, P(s_n)$ respectively, and $r_1, \ldots, r_k$ are discrete alternative states of r with probabilities $P(r_1), \ldots, P(r_k)$ respectively.

[6]Arguably we don't want an asymmetric notion of information either, in so far as two distinct states can (in certain circumstances) carry information about each other. The information relation is plausibly (not anti-, but) non-symmetric.

some non-symmetric condition(s) to a symmetric information relation, some standard approaches to building content and contentful states out of information would imply inappropriate symmetries if Shannon's account were used (Dretske, 1981).) For this very reason, Dretske's replacement of Shannon's symmetric by his own non-symmetric definition of information is crucial to securing the applicability of the notion to the cases of interest. If so, then the ability of Shannon's formulation to sidestep the problem we discussed does not make it preferable for standard philosophical purposes.

4   Ontological Economy

Demir worries that another claimed motivation for the counterfactual view — its ontological economy — evaporates upon reflection. For while (S*) avoids reference to laws, it makes central use of counterfactual conditionals whose truth conditions are standardly understood to involve appeal to possible worlds.[7] But since commitment to possible worlds is often taken as a serious burden, Demir concludes that (S*) is not less ontologically committed than accounts that appeal to laws. As he tallies the score, then, considerations of ontological economy are neutral between the theories.

We think the score-keeping is more complicated than this makes out. To begin, possible worlds have proven useful in a large variety of philosophical projects; consequently, many will regard the appeal to possible worlds in a theory of information as previously justified by their other fruits. In any case, commitment to possible worlds is only ontologically costly given inflationary understandings of possible worlds (e.g., Lewis, 1986) — and, of course, there are deflationary (e.g., fictionalist, abstractionist, instrumentalist) alternatives available (e.g., Rosen, 1990; Armstrong, 1989; van Inwagen, 1986; van Fraassen, 1980; Forbes, 1983; Yablo, 2001). Hence, we don't see that the counterfactual account, by being committed to possible worlds, carries onerous special burdens.

By the same token, and as C&M pointed out, a commitment to laws by a probabilistic theory of information can be rendered innocuous by a deflationary understanding of lawhood — e.g., a version of the Mill-Ramsey-Lewis approach augmented to account for probabilities, (Loewer, 1996, 2004). However, for those skeptical about laws of nature (e.g., Cartwright, 1983; van Fraassen, 1989; Ward, 2002; Mumford, 2004), the counterfactual theory of information

---

[7]We are officially agnostic about the semantics for counterfactuals, so (as Demir anticipates) we might respond by rejecting the standard (possible worlds involving) semantics for counterfactuals. Demir responds that this would (i) make the counterfactual account incomplete, and (ii) make the account equivalent to that proposed by Loewer (1983).

We remain undaunted. As for (i), we never aspired to explain every piece of technical apparatus used by the theory we proposed (a good thing, too, since filling in the semantics for counterfactuals wouldn't do that either). Regarding (ii), (S*) would only become equivalent to Loewer's theory given the adoption of something like Loewer's (controversial) Humean understandings of counterfactuals and laws. Even granting all this, we see no reason for alarm; as we said in C&M, "we are happy to think of our proposal as a notational variant of that suggested by Loewer (we are delighted to have philosophical company)" (338).

may be preferable to competitors in avoiding commitment to laws. (For purposes of this paper we are agnostic about laws.)

5   Information and Intentionality

Finally, Demir is unimpressed by our attempt to render information non-doxastic so that it might underwrite reductive understandings of the intentional. C&M complained that Dretske's view of information makes essential reference to the knowledge of the recipient of a message, while he also defines knowledge in terms of information; consequently, it appears that his theory fails to break out of the circle of the intentional. As we noted, Dretske has a response to this concern — a response that urges us to continually reapply the analyses of knowledge and information until "we reach the point where the information carried does not depend on any prior knowledge about the source, and it is this fact that enables our equation to avoid circularity" (Dretske, 1981, 86). But we suggested this response fails because there is no reason to believe that repeated application of Dretske's procedure will result in a case of knowledge-independent information. Indeed, Dretske himself suggested that the procedure could result in an undending circle of mutual dependence.

Demir responds on Dretske's behalf that such mutual dependence is unproblematic if the two pieces of information in question nomically or metaphysically nest one another (Dretske, 1981, 71); for then any signal that carries one of the two will also carry the other. He concludes that "[C&M's] example of mutually dependent information does not provide a counterexample to Dretske's claim that the backwards iteration in his recursive definition of informational content can eliminate reference to doxastic states" (57).

Once again, we are unconvinced.

First, this solution is incompatible with Dretske's commitments. Suppose that K's knowledge that s is F depends on K's knowledge that s is G. And suppose we grant, per Demir, that the information that s if F and the information that s is G are mutually nested. Because Dretske holds that a signal S has the fact that t is F as its semantic content only if "S carries no other information, r is G, which is such that the information that t is F is nested (nomically or analytically) in r's being G" (Dretske, 1981, 185), it follows that neither the fact that s is F nor the fact that s is G can be the semantic content of any mental state. But Dretske (plausibly) wants it to turn out that the fact that s is F is the semantic content of K's knowledge that s is F, and that the fact that s is G is the semantic content of K's knowledge that s is G. It would seem, then, that Demir's proposal would preclude Dretske from assigning the semantic contents to states that he (and everyone else) wants. This seems an awfully high price to pay.

Second, there is, as far as we can see, nothing to ensure that whenever two doxastic states stand in a relation of mutual dependence they will nomically or metaphysically nest one another. But if they do not, then Demir's proposed solution is inapplicable. And if this occurs in even a single case, our worry

arises once again that this will disqualify Dretske's theory from serving as a reductive account of content (as per his explicit aim).

Finally, even if mutually dependent doxastic states always mutually nested one another, this still wouldn't secure the metaphysically reductive aims that many (Dretske included) have wanted from a theory of information. Nesting *would* help in explaining how, as a matter of ontogenesis, an individual organism gets itself into the situation of having a state that bears in an informational relation to the world. But this fails to resolve our worry that the sort of mutual dependence we described amounts to a kind of *metaphysical* dependence of one piece of information on the other. Such mutual dependence would mean that obtaining of a first doxastic/informational relation makes it the case that a second doxastic/informational relation obtains, while the obtaining of the second relation makes it the case that the first obtains. If this occurs (and there is nothing in Dretske's system to prevent it) then the information relations are never fully grounded in anything independent of the doxastic. To say this is just to say that — nesting or no nesting — the account fails to break out of the circle of the intentional. Thus, we suggest, those whose interests in information require that the notion be reductive will have to search elsewhere.

6 Indeterminacy

Scarantino claims that the counterfactual theory of information is unacceptably indeterminate. He comes to this conclusion by examining our discussion of a case originally described by Dretske:

> ... suppose there are four shells and a peanut is located under one of them. ... I turn over shells 1 and 2 and discover them to be empty. At this point you arrive on the scene and join the investigation. You are not told about my previous discoveries. We turn over shell 3 and find it empty Dretske (1981, 78).

On the counterfactual account, whether the examination of shell 3 carries information depends the truth of this counterfactual:

(C) ⌜if the peanut had not been under shell 4, then the result of examining shell 3 would have been different⌝.

In C&M we claimed that the truth of (C) — hence the information carried by the examination of shell 3 — depends on what possibility/probability distribution holds (346).[8] But Scarantino (2008, ms) complains that this leaves unexplained the source of the possibility/probability distributions or which distributions matter to the assessment of (C). Moreover, he argues that the needed supplementation can be bought only at the cost of either (i) giving up on the non-doxastic character of the counterfactual account, or (ii) relying on the usual

---

[8]As C&M point out, this won't result in a non-doxastic theory of information unless the distributions in question are themselves independent of the doxastic.

semantics for counterfactuals (Lewis, 1973) in a way that leads to counterintuitive results about the presence or absence of information.[9]

We are not inclined to accept (i) since this would mean surrendering one of the most distinctive features of our account. Hence, if forced to choose, we prefer Scarantino's (ii).

Scarantino claims that the standard semantics of counterfactuals implies that (C) is false, hence, that, if we accept (ii), the examination of shell 3 does not carry the information about the peanuts being under shell 4. We would be prepared to accept this result. The examination of shell 3 clearly carries information about the peanuts not being under shell 3. Moreover, the counterfactual account implies that the emptiness of shells 1 through 3 carries information about the peanut being under shell 4 (since ⌜If the peanut had not been under shell 4, then the result of examining shells 1–3 would have been different⌝ is true). Thus, the counterfactual theory finds significant information in the examination of the shells even if it does not accord perfectly with every intuition about information.

More importantly, Scarantino's case does not show that (C) must come out false on the standard semantics; rather, it shows that (C) comes out false on that semantics *given further assumptions about the case that fix the possibility/probability distribution*. It is indeed plausible that (C) is false in Scarantino's enriched scenario, in which an agent always places the peanut under even-numbered shells (Scarantino, 2008, ms). On the other hand, given other enrichments (e.g., in a scenario in which an agent always places the peanut under shells 3 or 4), the standard semantics appears to make (C) true. This seems to vindicate our claim that the information carried by the examination of shell 3 depends on the possibility/probability distribution in place.

7   Explanation

Scarantino's more significant complaint is that the counterfactual account fails to underwrite the "central explanatory role" of information as a link between learning and subsequent behavior (Scarantino, 2008, ms).

Now, it is plausible that what can be learned by an individual depends on what that individual already knows; but since the counterfactual account does not relativize information to the knowledge of a recipient, it cannot identify the information a signal carries with what can by learned from it by agents in different epistemic positions. Consequently, the differential behavior of such agents will not be fully explained by the information they receive.

But we deny that this drains information of explanatory significance. For example, in the shell game described above, the fact that the first player received the information that the peanut was not under shell 3 is a natural explainer of her beliefs and betting dispositions. After all, it is plausible that

---

[9]We ignore solutions that depend on a non-standard semantics for counterfactuals.

without having received that signal these would have been different. Something else, then, must be bothering Scarantino.

Here is a conjectured diagnosis. Scarantino holds that a subject's behavioral response to a signal is under the (exclusive) control of the information it carries. The counterfactual account implies that information is under the control of the (subject-independent) environment. Combining these two views results in the behaviorist claim that a subject's behavior is under the control of the environment. But behaviorism is mistaken. Scarantino favors responding by denying that information is under the control of the environment, and instead holds the view that information is partially determined by the background knowledge of recipients. While this response does avoid the unwanted behaviorist conclusion, there is another response left open. We suggest that a better alternative is to deny that behavior is under the exclusive control of information. Information is relevant to behavior, but it is not the only relevant factor. If we say this, then we can join Scarantino in holding that a full explanation of an agent's behavior must appeal both to the information it receives and her prior doxastic states. Who (barring behaviorists) would have thought otherwise?

References

Armstrong, D. M. (1989). *A Combinatorial Theory of Possibility*. Cambridge University Press, Cambridge.

Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford University Press, New York.

Demir, H. (2008). Counterfactuals vs. conditional probabilities: A critical analysis of the counterfactual theory of information. *Australasian Journal of Philosophy*, **86**(1), 45–60.

Dretske, F. I. (1981). *Knowledge and the Flow of Information*. MIT Press, Cambridge, Massachusetts.

Forbes, G. (1983). Physicalism, instrumentalism and the semantics of modal logic. *Journal of Philosohpical Logic*, **12**, 271–298.

van Fraassen, B. C. (1980). *The Scientific Image*. Clarendon Press, Oxford.

van Fraassen, B. C. (1989). *Laws and Symmetry*. Clarendon, Oxford.

Humphreys, P. (1985). Why propensities cannot be probabilities. *The Philosophical Review*, **94**(4), 557–570.

van Inwagen, P. (1986). Two concepts of possible worlds. *Midwest Studies in Philosophy*, **XI**, 185–213.

Lewis, D. (1973). *Counterfactuals*. Basil Blackwell, Oxford.

Lewis, D. (1986). *On the Plurality of Worlds*. Basil Blackwell, Oxford.

Loewer, B. (1983). Information and belief. *Behavioral and Brain Sciences*, **6**, 75–76.

Loewer, B. (1996). Humean supervenience. *Philosophical Topics*, **24**(1), 101–127.

Loewer, B. (2004). David Lewis' Humean theory of objective chance. *Philosophy of Science*, **71**, 1115–1125.

Mumford, S. (2004). *Laws in Nature*. Routledge, London and New York.

Rosen, G. (1990). Modal fictionalism. *Mind*, **99**(395), 327–354.

Scarantino, A. (2008). Shell games, information, and counterfactuals. *Australasian Journal of Philosophy*. in press.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423, 623–656.

Swain, S., Alexander, J., and Weinberg, J. M. (2008). The instability of philosophical intuitions: Running hot and cold on Truetemp. *Philosophy and Phenomenological Research*, **76**(1), 138–155.

Ward, B. (2002). Humeanism without Humean supervenience: A projectivist account of laws and possibilities. *Philosophical Studies*, **107**, 191–218.

Yablo, S. (2001). Go figure: A path through fictionalism. *Midwest Studies in Philosophy*, **25**(1), 72–102.