

An Objective Counterfactual Theory of Information*

Jonathan Cohen[†] and Aaron Meskin[‡]

Philosophers have appealed to information (as understood by [Shannon, 1948] and introduced to philosophers largely by [Dretske, 1981]) in a wide variety of contexts; information has been proffered in the service of understanding knowledge, justification, and mental content, *inter alia*. While information has been put to diverse philosophical uses, there has been much less diversity in the understanding of information itself. In this paper we'll offer a novel theory of information that differs from traditional accounts in two main (and orthogonal) respects: (i) it explains information in terms of counterfactuals rather than conditional probabilities, and (ii) it does not make essential reference to doxastic states of subjects, and consequently allows for the sort of objective, reductive explanations of notions in epistemology and philosophy of mind that many have wanted from an account of information.

We'll first present our counterfactual account of information (§1), and show how it sidesteps a problem that has been raised for its traditional, probabilistic competitors (§2). Next we'll compare the counterfactual account against that proposed by Dretske (§3), highlighting the differences between the two. After that, we'll turn to questions about objectivity: we'll bring out a conflict between the essentially doxastic character of traditional theories of information and the reductive purposes philosophers have had in mind in appealing to information (§4), and we'll show how the account of §1 can be formulated in non-doxastic terms. Finally, we'll consider objections against the proposed account (§5). Ultimately, we'll suggest, the objective counterfactual account of information should be taken as a serious contender to more traditional rivals.

1 Counterfactuals and Information

1.1 A Crude Counterfactual Account

It is a plausible and widely accepted view that informational relations require not only simple (actual) covariation, but also counterfactual variation. For ex-

*This work is fully collaborative; the authors are listed in alphabetical order.

[†]Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

[‡]Department of Philosophy, Texas Tech University, Box 43092, Lubbock, TX 79409, aaron.meskin@ttu.edu

ample, even if your room and my room are exactly alike in ambient temperature, the thermometer in my room carries information about the temperature of my room and not yours insofar as the thermometer's reading supports counterfactuals about the ambient temperature of my room and not yours. This line of thought leads naturally to the view that such counterfactuals can provide a handy diagnostic test for the presence of informational relations — a weak view about the relation between counterfactuals and information that we'll call (W):

(W) x 's being F carries information about y 's being G if the counterfactual conditional \lceil if y were not G , then x would not have been F \rceil is true.¹

And this, in turn, might lead one to a stronger view about the relation between counterfactuals and information that we'll call (S):

(S) Informational relations are constituted by the holding true 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 \lceil if y were not G , then x would not have been F \rceil is true.

(W) and (S) are interesting theses; however, they are both false. This is because, in cases where y 's being G is necessary (or y is world-bound), (W) gives the wrong verdict about the presence of informational relations. This shows that that (W) is inapplicable as a diagnostic test in at least these cases, and consequently, that (S) must go wrong as well.² To see this, suppose it is necessary that y is G (or that y is world-bound), and ask whether x 's being F carries information about y 's being G . According to (W), such a relation holds only if the counterfactual \lceil if y were not G , then x would not have been F \rceil is true. By hypothesis the antecedent of that counterfactual is false, so the counterfactual comes out (vacuously) true on the standard semantics for counterfactuals; consequently, (W) yields the verdict that there is information in this case.

But, if we follow Shannon's understanding of information, then this is the wrong verdict. According to Shannon, information amounts to a reduction in entropy, or uncertainty. As such, if some outcome is certain in one of the ways described, there is no entropy about that outcome that can be reduced, so nothing can carry information about that outcome. That is, if a variable y has only a single outcome G with non-zero probability, then the entropy for that variable is already zero, so it can't be reduced; it follows that the maximum mutual information between y 's being G and any other variable must be 0. Hence, nothing can carry information about y 's being G .³

¹ Here and in what follows, x 's being F and y 's being G are construed as actual events; it follows that (on (W)), one event carries information about a second only if the second actually occurs.

²Our interest in (W), and why it goes wrong, is partly autobiographical; see [Meskin and Cohen, 2006].

³The same result follows from Dretske's slightly different setup (cf. §2). If there is no possible alternative outcome for p , then the conditional probability of p given k alone is unity; hence, on this definition, too, nothing can carry the information that p ; see ([Dretske, 1981], 12).

This shows that, in many cases, (W) will find an information relation where there shouldn't be one. As such, (W) is false; *a fortiori* (S) is false as well.

1.2 A Better Counterfactual Account

If (W) and (S) fail because of cases where the counterfactuals come out vacuously true, a solution that suggests itself is merely to rule the troublesome cases out by fiat. Thus, consider the following two views about the relationship between counterfactuals and information:

(W*) x 's being F carries information about y 's being G if the counterfactual conditional \lceil if y were not G , then x would not have been F \rceil is non-vacuously true.

(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 \lceil if y were not G , then x would not have been F \rceil is non-vacuously true.

(W*) expresses the view that certain counterfactuals can provide a handy diagnostic test for the existence of an informational relation. (S*) expresses the view that informational relations are constituted by certain counterfactual relations.⁴

Since (S*) has it that informational relations are constituted by the truth of certain counterfactuals, it is natural to think of it as providing a counterfactual account of information.⁵ In what follows, we shall explore the viability of such a counterfactual account.

2 Probabilities and Information

We think the counterfactual approach is worthy of consideration if only because it provides a new and interesting way of thinking about information. But another motivation for caring about it comes from consideration of a problem raised by [Loewer, 1983] about the interpretation of probabilities essential to standard accounts of information.

The main account of information in philosophical currency, due to [Dretske, 1981], holds that a signal r carries the information that p just in case

⁴This formulation leaves it open that events can carry information about themselves. One could avoid this outcome (if desired) by stipulating that informational relations only obtain between distinct events. We take no stand as to the advisability of this stipulation.

⁵Like Dretske's and Shannon's understandings of information (cf. §1.1), (S*) entails that no information can be carried about necessary events. On the other hand, there are two ways in which this assumption might be relaxed without giving up on the spirit of the present proposal. The first option would be to simply stick with (S) as opposed to (S*). A drawback of this option is that (S) entails that *all* signals carry information about every necessary event; but presumably anyone interested in making room for the possibility of information-carrying about necessary events would find this entailment intolerable. A second alternative would be to use a non-standard account of the semantics of counterfactuals according to which not all counterlegals are vacuously true. Thanks to Fred Dretske for encouraging us to think about this issue.

the conditional probability of p , given r , (and k , the knowledge of the receiver of r), is 1 (but given k alone, less than 1); or, in familiar notation, r carries the information that p iff $P(p|r, k) = 1$ and $P(p|k) < 1$ ([Dretske, 1981], 65). In his commentary on Dretske, Loewer pointed out that the usual candidates for understanding probabilities are ill-suited for Dretske's purposes.⁶ An understanding of probability as degree of belief is inappropriate not only because it would render the resulting notion of information objectionably epistemic, but also because it would undercut the possibility of using information to provide a naturalistic, reductive understanding of belief and other doxastic states (a possibility dear to the hearts of Dretske and many other philosophers who are interested in information). A frequentist understanding of probability also seems unsuitable because such understandings are inapplicable to unrepeatable event tokens; but, presumably we want a notion of information that extends to singleton events. Propensity interpretations of probability are generally thought to avoid this last difficulty, but they, too, fail in the present context:

The propensity of a chance setup C producing outcome e is usually explained . . . as a measure of the causal tendency of C producing $[e]$. But Dretske is after the converse probability, the probability that r was produced by a chance setup C . This probability is usually not meaningful on a propensity (or for that matter a frequency) interpretation. The point is that $P(r|C)$ may be meaningful but not $P(C|r)$, since there may be no propensity $P(C)$ ([Loewer, 1983], 76).

We take Loewer to be making two points here: first, that the apparatus in terms of which the propensity theorist unpacks conditional probabilities does not deliver the sort of conditional probability Dretske needs, $P(C|r)$ — instead, it delivers the converse probability $P(r|C)$; and second, that the propensity theorist cannot make use of the probability $P(r|C)$ in order to derive the needed probability $P(C|r)$ (by Bayes's Theorem) without knowing a value for the propensity $P(C)$, which is in general not defined.⁷

⁶The remarks that follow are applicable to other accounts of information (both semantic and quantitative) that are grounded in conditional probabilities. Most saliently, consider the setup of [Shannon, 1948]: let $\{s_1, \dots, s_n\}$ be discrete alternative states of a source s with probabilities $\{P(s_1), \dots, P(s_n)\}$ respectively, and let $\{r_1, \dots, r_k\}$ be discrete alternative states of a receiver r with probabilities $\{P(r_1), \dots, P(r_k)\}$ respectively; assume that $P(s_i) > 0$ for all i , that $P(r_j) > 0$ for all j , and that $\sum_{i=1}^n P(s_i) = \sum_{j=1}^k P(r_j) = 1$. Shannon defines the *mutual information between s and r* as follows: $I(s, r) = -\sum_{i=1}^n P(s_i) \log_2 P(s_i) + \sum_{j=1}^k P(r_j) \sum_{i=1}^n P(s_i|r_j) \log_2 P(s_i|r_j)$. So defined, mutual information makes ineliminable reference to the same sorts of inverse conditional probabilities as Dretske's theory, and so is vulnerable to the concerns we raise about the interpretation of those probabilities.

Considerations of convenience and specificity suggest that we should restrict attention to a single probabilistic account in what follows; given that it is already familiar to so many philosophers, it seems reasonable to choose Dretske's account as our focus.

⁷The classic statement of the problem that propensity theories have with 'inverse' probabilities (often called 'Humphreys's paradox') occurs in [Humphreys, 1985]. For reviews of the literature on Humphreys's paradox, see [Gillies, 2000] and [Humphreys, 2004].

We are left, therefore, without a suitable way of understanding the probabilities that Dretske uses to underpin his theory of information. One response to this situation would be to hope for some new account of probability that avoids these difficulties. For those who, like us, are too impatient to wait for that outcome, the counterfactual account of information will seem attractive, insofar as it sidesteps the problems about probabilities altogether.

3 Counterfactuals and Probabilities: A Comparison

For better or worse, the account of information in [Dretske, 1981] is the most prominent and influential in circulation among philosophers. We consider it a benefit of the counterfactual account that it agrees closely with Dretske's account in its verdicts about many cases — partly because we are impressed with the power of Dretske's account, and partly because this suggests that our proposed view merits consideration as a theory of information.⁸ However, we are also interested in bringing out differences between the counterfactual and the probabilistic accounts, and showing that, where they differ, there is reason to prefer the counterfactual view; that will be the burden of the present section.

⁸Another reason for thinking that the account presented here deserves to be thought of as a theory of information comes from consideration of the conditions (derived from the quantitative framework developed in [Dretske, 1981], chapters 1–2) that, according to Dretske, an adequate theory of information must respect (63–66). While we lack the space to rehearse Dretske's quantitative framework and motivate his conditions here, we wish to point out that the counterfactual account satisfies two of his three conditions, and provides principled reasons for rejecting the third. (That the counterfactual account should follow Dretske's account reasonably closely on these matters is not so surprising, in view of the intimate relationship between probabilities and counterfactuals.)

The counterfactual account respects condition B (i.e., that a signal carries the information that y is G only if y is G) (64) because we have restricted informational relations to holding between actual events (note 1); consequently, we may infer that if x 's being F carries information about y 's being G (as understood by (S*)), then y is G . The counterfactual account also respects condition C (i.e., that a signal carries the information that y is G only if the quantity of information the signal carries about y is (or includes) that quantity generated by y 's being G) (64). For the satisfaction of the counterfactual conditional in (S*) guarantees that, no matter what other information is carried by x 's being F , that state of affairs includes the information associated with y 's being G — even if x 's being F is correlated in the actual world with various actual states of affairs, we know that the one correlation that will persist as we move between worlds is that between x 's being F and y 's being G . However, the counterfactual account does not secure condition A (i.e., that a signal carries the information that y is G only if it carries as much information about y as would be generated by y 's being G) (63). This is because the non-vacuous truth of the counterfactual \lceil if x were not F , then y would not have been G \rceil allows that the signal (viz., x 's being F) might have a non-zero equivocation (i.e., it leaves it open that $P(y\text{'s being } G|x\text{'s being } F) < 1$). We shall argue in §3.2 that this result is unobjectionable — indeed, advantageous — for a theory of information (cf. [Scarantino, 2005]). Consequently, *pace* Dretske, we neither regard condition A as a necessary condition for any theory of information worthy of the name, nor do we regard the failure of the counterfactual account to secure this condition as reason for thinking that it should not count as a theory of information.

3.1 Information and Laws

A first difference between the two accounts concerns the extent to which they rely on natural laws to explain informational relations.

For Dretske, nomicity figures in the theory as a way of respecting the insight — an insight lying at the heart of the counterfactual account of this paper as well — that information relations require counterfactual dependencies and not mere correlation. While Dretske does not make explicit reference to laws or nomic dependencies in his definition of informational content (see §2), the appeal to probabilities of 1 is meant to indicate that such laws underwrite the requisite counterfactual (and, hence, informational) relations.⁹

In contrast, it is much less clear whether a counterfactual account requires an (implicit) appeal to laws or nomic regularities. While some have argued that counterfactual relations are essentially dependent on the existence of laws or lawlike relations (e.g., [Goodman, 1954]), this is not untendentious. For those who think it is a mistake to characterize counterfactuals as essentially dependent on laws the counterfactual account will be preferred to the probabilistic account. And even those who (like us) are agnostic about that issue may find the counterfactual approach preferable since it is less ontologically committed than the probabilistic theory.

In fact, in order to get around the worries he raises about Dretske’s appeal to conditional inverse probabilities, Barry Loewer proposes a simple nomic account of information that also eschews any appeal to probability. Here is (a modified version of) Loewer’s suggestion for capturing Dretske’s key idea: x ’s being F carries the information that y is G iff (i) it is contingently true that y is G , and (ii) it is nomically necessary that if x is F then y is G .¹⁰ Loewer suggests that the truth conditions of the conditional appealed to in this account involve the existence of laws that, along with certain background conditions (and x ’s being F), imply that y is G . It is not obvious how to choose between the counterfactual and Loewer’s nomic account (particularly if you think counterfactuals and laws are interdefinable, as per [Goodman, 1954], or if you think of probabilities, counterfactuals, and laws as all emerging on an equal footing from the best systematization of the Humean mosaic, as per [Lewis, 1994], [Loewer, 2001], [Loewer, 2005]).

As we’ve noted, the counterfactual account might be preferred on grounds of ontological economy, given the (tendentious) assumption that the counterfactual account doesn’t advert to laws. Otherwise, we are happy to think of our proposal as a notational variant of that suggested by Loewer (we are delighted to have philosophical company).

⁹“The transmission of information requires, not simply a set of de facto correlations, but a network of nomic dependencies between the condition at the source and the properties of the signal” (76–77). Cf. also the quotation in note 24.

¹⁰We have modified the proposal in ([Loewer, 1983], 76), inserting the insistence that the conditional’s consequent be contingent, as a way of blocking objections about vacuous satisfaction — objections that are parallel to those that motivated the transition from (S) to (S*) in §1.

3.2 Information and Communication

A second difference between Dretske’s account and the counterfactual account concerns the way the two understand the possibility of communication and the flow of information.

Dretske argues that, in order to allow for communication chains, any semantic theory of information has to make sense of the following principle, which he calls the ‘xerox principle’: “If A carries the information that B, and B carries the information that C, then A carries the information that C” (57). But the counterfactual account of information does not support the xerox principle. For, as David Lewis points out, the following inference schema is invalid ([Lewis, 1973b], 32-34):

$$\begin{aligned} P \Box \rightarrow Q \\ Q \Box \rightarrow R \\ \therefore P \Box \rightarrow R \end{aligned}$$

Since, according to the theory on offer, informational relations are grounded in counterfactual dependencies, the non-transitivity of the latter entails the non-transitivity of the former.

Given that it fails to secure the xerox principle, does the counterfactual account preclude communication chains? We think not. For, crucially, while that account implies that the information-carrying relation is non-transitive, it does not imply that the information-carrying relation is intransitive. Our account denies that information-carrying is transitive *tout court*, but it allows that in many (but not all) cases information may flow from one event to another along a chain of communication.

Indeed, we want to suggest that it is in fact an advantage of the counterfactual account that it fails to license the xerox principle, insofar as rejecting it (for principled reasons, as explained above) provides a way of avoiding a commitment that many have found objectionable.

One of the most tendentious points in Dretske’s theory is his insistence that the information relation requires a conditional probability of p on r (and k) of unity.¹¹ Like many others (e.g., [Suppes, 1983], 81–82), we worry that this value is too demanding. However, while Dretske’s critics have often objected to his insisting on this value, for the most part they have not responded to the argument he offers in support of it (one exception is [Scarantino, 2005]). Dretske argues (58–60) that, if the xerox principle is to be respected, the conditional probability required for the information relation cannot be less than unity. For consider what happens if the probability required were some positive number $n < 1$.¹² Suppose that r carries the information that p in virtue of the fact that $P(p|r, k) = x$ (where $0 < x \leq n$), and suppose that s carries the information that r in virtue of the fact that $P(r|s, k) = y$ (where $0 < y \leq n$). Now it is consistent

¹¹For ease of exposition, here we use single letters as variables for events, although these should officially be understood as shorthand for predicational structures of the form ‘ x is F ’.

¹²Thanks to Stephen Meskin for help thinking about the conditional probabilities here.

with these probabilistic and information relations that $P(p|s, k)$ is smaller than any choice of n (in fact, it may equal 0); and, if so, no informational relation between s and p would hold, contrary to the xerox principle. Furthermore, it is absurd to think that r could carry information about p when $P(p|r, k) = 0$ since that would entail that the occurrence of r (along with k) precludes the occurrence of p (cf. [Dretske, 1981], 59). Luckily, the counterfactual account shows us exactly what to say here: it explains why the xerox principle is false (as a general claim), and therefore allows us to escape the conclusion that Dretske draws from it; at the same time, it allows for the possibility of communication chains in particular cases.¹³

4 Information and the Doxastic

In §3 we brought out some differences between our counterfactual account and Dretske’s probabilistic account of information. However, several of the most controversial elements of Dretske’s theory are orthogonal to the question of whether information is grounded in probabilistic or counterfactual relations, and instead concern the relation between information and doxastic states. The relation between the two is especially important for those who (like Dretske) hope to use information as part of a reductive analysis of epistemic and intentional notions. In this section we’ll argue that, given those sorts of reductive aspirations, Dretske’s account construes information in more doxastic terms than we should want, and we’ll show how to correct for these difficulties to arrive at a more objective understanding of information.

4.1 Information and Knowledge

We have seen that, on Dretske’s view, the information that is carried by a signal is relative to the knowledge of the prospective recipient of that signal. This relativization of information to knowledge can be motivated by its success in capturing Shannon’s central insight that information should reduce entropy or uncertainty. The note telling you that the ambient temperature is 70°F reduces uncertainty if you don’t already know this by other means, but not if you do. By making the information carrying capacity of a signal contextually sensitive to the knowledge of the recipient, Dretske brings our verdicts about information into line with those about the reduction of entropy. (Moreover, the plausible thought that two different individuals can learn something different from the same signal falls out as a consequence.)¹⁴

¹³As a referee points out, Dretske claims another benefit from the idea that the information relation requires a conditional probability of p on r (and k) of 1; namely, since he understands knowledge in terms of information (see §4.1), this idea provides for him a response to the lottery paradox ([Dretske, 1981], 99–102). In not making this requirement, then, we are also giving up Dretske’s response to the paradox. We’re not inclined to worry too much about this (even if, as per Dretske, knowledge is construed in terms of information); for there are a number of alternative responses to the paradox on offer (cf. [Vogel, 1992], [Hawthorne, 2004]).

¹⁴At one point Dretske claims that the relativization at issue is supported by ordinary ways of thinking: “I propose to follow Dennett in relativizing the information contained in a signal

But there are significant costs associated with this way of accommodating the intuition. The primary worry is that the approach looks to undercut the possibility of appealing to information in the construction of a naturalistic and reductive account of epistemic and doxastic notions (which, as we have pointed out, is central to Dretske’s project).¹⁵ This worry comes out clearly for Dretske, since his account of information makes essential reference to the knowledge of the recipient of a message, while he also defines knowledge in terms of information: “*K* knows that *s* is *F* = *K*’s belief that *s* is *F* is caused (or causally sustained) by the information that *s* is *F*” ([Dretske, 1981], 86). Unsurprisingly, Dretske has a response to the charge of circularity. He suggests that the account is not in fact circular, but recursive: “Whether a person can learn that *s* is *F* may depend on what else he knows about *s*, but it does not depend, and is not said to depend, on his knowing that *s* is *F*” ([Dretske, 1981], 86). The thought is that if we want to determine whether an agent knows a proposition we can 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” (86).

But we remain unconvinced by this reply. We see no reason to believe that repeated application of the analysis will always terminate in a case where the information carried does not depend on any prior knowledge about the source. Indeed, Dretske’s own examples suggest one obvious way in which the analysis could fail to terminate — namely, that it could result in an unending circle of mutual dependence. To see this, let it be that, as Dretske claims, *K*’s knowledge that *s* is *F* depends on the information that *s* is *F* and therefore (because of the role prior knowledge plays in his analysis of information) also on some other bit of knowledge *K* has about *s* (e.g., that *s* is *G*). For the same reasons, it seems entirely possible that *K*’s knowledge that *s* is *G* depends on some further bit of knowledge *K* has about *s*. But nothing in Dretske’s account rules out the possibility that this further bit of knowledge is in fact *K*’s knowledge that *s* is *F*; on the assumption that the dependencies under discussion are transitive, an immediate regress ensues.

Whether or not Dretske has further apparatus that could block this sort of circularity, the broader point is that Dretske has given us no reasons for believing that his analysis of information ever breaks out of the intentional/doxastic circle. The upshot is that Dretske’s account of information, as it stands, can’t be used to provide reductive and naturalistic accounts of epistemic (and doxastic) notions. What to do?

[to the knowledge of the recipients] because I think this accurately reflects our ordinary ways of thinking about such matters” ([Dretske, 1981], 79). In response, we suggest: (1) it is not at all clear to us that ordinary ways of thinking do support relativization, and (2) even if relativization did reflect ordinary ways of thinking, we are unconvinced that the technical notion of information should be beholden to such thinking.

¹⁵Dretske himself raises this sort of issue with Kenneth Sayre’s information theoretic account of recognition: “One cannot use the concept of information in the analysis of such epistemic concepts as knowledge, recognition, and memory if one turns around and uses epistemic concepts (such as learn, identify, or correctly interpreting) in the analysis of information. This makes the entire program circular” ([Dretske, 1981], 244).

One response to this situation would be to give up on the reductive aspirations that motivate Dretske and others. Alternatively, one might hope to construct an alternative theory of information that does not make essential use of doxastic analysanda.

Can the counterfactual theory of information be of service here? On the one hand, it makes no explicit use of doxastic notions in attempting to tie information to the reduction of entropy. On the other hand, many have thought that counterfactuals depend tacitly on the doxastic; e.g., Lewis's semantics for counterfactuals is based on an overall similarity metric between worlds that, he says, depends crucially on our interests ([Lewis, 1973b], 8–9). But if a non-doxastic theory of information is desired, then one could construct such a theory by adopting the counterfactual theory of information, while insisting (contra Lewis) that there are non-doxastic facts of the matter about which counterfactuals are true. The resulting theory, then, will be an *objective* counterfactual theory of information.

Likewise, one could achieve an objective probabilistic theory by combining a probabilistic account of information with the insistence that there are mind-independent facts about (i) what the options are over which probabilities are distributed, and (ii) the values of the probabilities over those options. Once again, this marks a departure from standard treatments of probability; and once again, one's willingness to make such a departure will depend on the strength of one's desire for a non-doxastic theory of information.

In the rest of this section we'll explore what such an objective theory of information would look like (concentrating on counterfactual rather than probabilistic formulations for brevity) and what it will say about certain cases. We hasten to add, however, that there are costs to insisting on the sort of objectivity at issue: in particular, it requires the heterodox view that the truth values of the counterfactuals (/values of the probabilities) supervene on the non-doxastic facts. We have no official recommendation about whether the benefits of objectivity justify its costs; we aim only to provide an honest assessment of the costs and benefits in order to facilitate an informed decision.

4.2 Information and Context

We argued in §4.1 that Dretske's relativization of information to knowledge undercuts the reductive aspirations in whose service Dretske and others have appealed to information. Suppose, then, that one is inclined to give up on Dretske's theory of information, and to hope for an alternative theory of information that does without relativization to knowledge. One needs to say something, then, about the cases that initially made the relativization seem plausible.

Consider first a case described in ([Dretske, 1981], 14).

Case 1: A group of eight employees has to choose one of their members to perform an unpleasant task. They devise a fair procedure to do this and agree to write the name of the person down on a message to send to their

boss. However, the group has agreed that if Shirley is selected, they will still write Herman's name on the message. Now suppose the group selects Herman, and his name is sent to the boss who is (prior to receiving the message) ignorant of the outcome of the selection process.

In case 1, both (i) a probabilistic account and (ii) a counterfactual account entail that the note does not carry the information that Herman was selected. For (i) the probability that Herman was selected conditional on the note reading 'Herman' is only .5 (or, at any rate, less than 1),¹⁶ and (ii) it is not the case that if Herman had not been selected that the note would have failed to read 'Herman'. But both accounts also entail that the note carries the information that Herman or Shirley was selected. (For (i) the probability that Herman or Shirley were selected conditional on the note reading 'Herman' is unity, and (ii) if Herman or Shirley had not been selected then the note would have failed to read 'Herman'.)

Now consider a variation on case 1:

Case 2: As in case 1, with the additional suppositions that the choice procedure involves a series of coin flips, and that the final flip is one that decides between the only two remaining candidates, Herman and Bill. The final flip selects Herman, and the note reading 'Herman' is sent off to an unknowing boss.

Case 2 is, in one clear sense, not substantively different from case 1; it differs only in that we have specified more about the process leading to the common outcome (Herman is selected, 'Herman' is inscribed on a note). On the other hand, the presentation of the very same sequence of events under the heading 'case 2' invites us to think of the selection process as having involved earlier stages in which the element that would otherwise have destroyed the information-carrying capacity of the note (viz., the possibility that Shirley would be selected) has been excluded. This difference is significant — it tells us more about the events, and in particular about the distribution of various alternative possibilities and probabilities among alternative outcomes, than we knew when we considered case 1. As such, it seems plausible to hold that the exclusion of putatively information-destroying elements at earlier stages of the process restores the informational relation that was absent in case 1 — viz., that Shirley's having been ruled out by earlier coin flips makes it the case that (i) the probability that Herman is selected conditional on the note reading 'Herman' is unity, and that (ii) if Herman had not been selected that the note would have failed to read 'Herman'.

We take this contrast to show that our assessments of informational relations are sensitive to some factor that distinguishes the context of case 2 from that of case 1. Upon reflection, however, this is unsurprising. It is a familiar lesson that our assessments of counterfactuals and probabilities depend on the background range of accessible possibilities and the probability distribution over al-

¹⁶Here and below we assume (for ease of exposition) that there is some coherent way of making sense of the conditional probabilities.

ternatives (i.e., the proximity relations among and probability distribution over worlds); consequently, if information is construed in terms of either counterfactuals or probabilities, our assessments of informational relations will be sensitive to the same factors.¹⁷ But it is easy to see that the alterations that distinguish case 2 from case 1 affect our beliefs about just these factors. The aggregative presentation labeled ‘case 1’ invites us to think that some ‘Herman’-note/non-Herman-selection worlds are as close to the actual world as any non-‘Herman’-note/non-Herman-selection worlds (perhaps because of some sort of principle of indifference), which, on a standard semantics for counterfactuals, entails that ‘if Herman had not been selected then the note would have failed to read ‘Herman’ ’ is false. Whereas, the articulated presentation labeled ‘case 2’ tells us that our prior assessment of the possibilities was incorrect: here we learn that (contrary to indifference assumptions) no ‘Herman’-note/non-Herman-selection world is as close to the actual world as any non-‘Herman’-note/non-Herman-selection worlds (because of selections made at earlier stages), which, on a standard semantics for counterfactuals, entails that ‘if Herman had not been selected then the note would have failed to read ‘Herman’ ’ is true. Alternatively, in terms of probabilities, the case 1 presentation suggests that some Herman-selected scenarios have a non-zero probability conditional on the note reading ‘Herman’ (again, possibly because of some sort of principle of indifference), which means thinking that the probability that Herman is selected conditional on the note reading ‘Herman’ is less than one. Whereas what we learn from the case 2 presentation is that, in fact (and contrary to indifference assumptions), no Herman-selected scenarios have a non-zero probability conditional on the note reading ‘Herman’, which means that the probability that Herman is selected conditional on the note reading ‘Herman’ is unity.

Thus, just as the evaluation of counterfactuals and assignments of probabilities depends on background assumptions about what is possible or probable, informational relations (construed on either a counterfactual or probabilistic or theory of information) are dependent on the same factors. Now, if one wants to sustain the non-doxastic character of information, one can only accept this conclusion if one is committed to the existence of non-doxastic facts of the matter about the distribution of possibilities and probabilities. This thought can perhaps be motivated by the description we’ve given already: this might explain why we would say that the additional information in the presentation of case 2 allowed for a correction of the erroneous assessment first reached about case 1. On this view, then, the initial verdict reached about information in case 1 was

¹⁷Raymond Martin draws similar conclusions in [Martin, 1975] and [Martin, 1983], and then uses them to argue against the counterfactual accounts of knowledge of [Dretske, 1971] and [Nozick, 1981]. We agree with Martin about the context-sensitivity of our evaluations of counterfactuals; however, this is compatible with holding that there are non-doxastic facts of the matter about which counterfactuals are true (despite the context-sensitivity of our assessments). If information is construed as depending on which counterfactuals are true rather than which counterfactuals we take to be true, Martin’s point is not damaging to the account of information we are exploring in the way Martin alleges it is damaging to counterfactual theories of knowledge. (We thank Dretske for bringing the Martin papers to our attention.)

erroneous (although it may have been justified), and was corrected only when we came into possession of more of the relevant facts of the matter about the distribution of possibilities and probabilities.

Likewise, consider Dretske’s original motivation for relativizing information to knowledge — his view that the note reading ‘70°F’ carries information if you haven’t already read the thermometer, but not if you have. On the contrary, a non-doxastic account should entail that the informational relation (or lack thereof) between the note and the ambient temperature is independent of the doxastic state of the recipient — rather, a non-doxastic theory of information will say that there is such a relation just in case the counterfactuals/probabilities are as required by that theory. Of course, saying this leaves room for thinking that the note may fail to be novel or significant to the recipient if she has already received another signal that carries the same information; but a proponent of a non-doxastic account will regard this as a reason to regard the note as redundant, not as failing to carry information.

4.3 Information and the Reduction of Entropy

What can a proponent of an objective theory of information say about the crucial (Shannon-inspired) point that information is a matter of “the reduction in uncertainty, the elimination of possibilities” ([Dretske, 1981], 4)? Once again, assuming one agrees with Dretske’s reductive ambitions, one can only respect this insight by construing these uses of ‘uncertainty’ and ‘possibilities’ in metaphysically robust, non-doxastic terms.¹⁸

To see how this approach differs from Dretske’s account, consider Dretske’s case of the shell game and the investigators:

...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).

(It is an unmentioned rule of the shell game that shells are put back down after having been examined.)

Because Dretske thinks of the elimination of possibilities in doxastic terms, he is committed to saying that the examination of shell 3 not only provides the two investigators with different information, but eliminates different possibilities for the two investigators. Given the first investigator’s prior knowledge, there are only two possibilities for the location of the peanut before the examination of shell 3. Turning it over eliminates one of those possibilities; hence, it provides the investigator with the information that the peanut is under shell 4. But, given the second investigator’s epistemic situation, there are four possibilities prior to

¹⁸For reasons of brevity, we’ll compare Dretske’s doxastic theory only against an objective counterfactual theory of information in this section. As far as we can see, what we say about this comparison holds (*mutatis mutandis*) about the comparison between Dretske’s view and an objective probabilistic theory of information as well.

the examination of shell 3. So turning it over only reduces the possibilities to three and, hence, does not provide the information that the peanut is under shell 4.

As we have emphasized, an objective counterfactual account will not relativize information to the doxastic states of recipients, and so can only preserve the connection between information and the elimination of possibilities by regarding the latter non-doxastically as well. Exactly what information, on this account, is provided by the examination of shell 3? As in the cases considered in §4.2, the answer to this question will depend on the distribution of background possibilities (which the objective theorist will want to treat non-doxastically). Suppose, as in the above example, that the peanut is in fact under shell 4. On certain possibility distributions, the counterfactual ‘if the peanut had not been under shell 4, then the result of examining shell 3 would have been different’ would be true; on these distributions, then, the examination of shell 3 provides the information that the peanut was under shell 4.¹⁹ The proponent of the objective counterfactual account will say, then, that there were (independent of the doxastic states of the investigators) two possibilities prior to the examination and only one possibility afterwards. But suppose the distribution of possibilities is different; in particular, suppose the peanut would have been under shell 2 if it hadn’t been under shell 4. If so, the relevant counterfactual would be false; hence, the examination of shell 3 could not provide the information that the peanut was under shell 4. Thus, given this alternative distribution of possibilities, the proponent of the objective counterfactual account will deny that the examination of shell 3 reduces two possibilities to one, and will instead maintain that it reduces three possibilities (i.e., the peanut could have been under 2, 3, or 4) to two (2 or 4).

Either way, irrespective of the background distribution of possibilities, the view will say that the examination of shell 3 does reduce possibilities — it eliminates the possibility that the peanut is under shell 3. Hence, the examination of shell 3 provides one piece of information (namely, the information that the peanut is not under shell 3) irrespective of the background distribution of possibilities.

¹⁹It is plausible that, in such a case, the first investigator would have come to learn (and, in fact, know) that the peanut is under shell 4, even though the same does not appear plausible for the second investigator (even though both investigators received the information that the peanut was under shell 4). We take this to show that any non-doxastic information-based account of knowledge will have to locate the difference between the epistemic status of investigator one and investigator two in some further fact — consequently, on this sort of theory, the relationship between knowledge and information will be less straightforward than Dretske proposes ([Dretske, 1981], 86; cf. §4.1. For suppose both investigators come to believe that the peanut is under shell 4 on the basis of the examination of shell 3. Dretske’s (unmodified) account of knowledge would seem to entail that both investigators have come to know the location of the peanut — and this verdict seems implausible in the case of investigator 2. But the objective counterfactual account does not face insoluble problems, since there appear to be the resources to explain the different epistemic status of the two investigators: namely, it is plausible that the first investigator’s knowledge of the location of the peanut depends (in part) on having received *additional* information from her prior examination of shells 1 and 2.

5 Troubles About Counterfactuals?

In this section we'll consider objections against the (objective) counterfactual account of information.

5.1 Counterfactuals Are Mysterious

In §2, we attempted to motivate the counterfactual account by arguing that probabilistic accounts depend on conditional probabilities to which all of the standard interpretations of probability are inapplicable. But one might worry that the counterfactuals to which the counterfactual account appeals are, in fact, in no better shape than the problematic probabilities.²⁰

In fact, despite the fact that there is a standard account of their semantics (see [Lewis, 1973b] for the canonical account) there are longstanding worries about counterfactuals.²¹ But we believe that our appeal to counterfactuals is on much firmer ground than Dretske's invocation of probabilities. There is, we believe, no serious reason to think the notion of counterfactuals is incoherent (notwithstanding verificationist worries, which we do not believe should be taken seriously). Of course, there is no serious reason to think that the whole notion of probability is incoherent either. The crucial comparison for our purpose has to do with whether extant theories of counterfactuals and probabilities are applicable in the contexts in which they are invoked. And our contention is that the counterfactuals that we appeal to do not face any specific problem (i.e. there is no reason to think that the general semantics for counterfactuals will not apply in their case), while the probabilities that Dretske appeals to do face a distinct problem (as we argued in §2).

There is one caveat. As we mentioned §3.1, there is a Humean approach that sees laws, counterfactuals, and probabilities as on equal footing. On this approach, the counterfactual account offers no gain in parsimony or theoretical commitment over a probabilistic account. That said, the reasons we've given for preferring an objective account that eschews doxastic notions (§4) remain in place.

5.2 Preemption

The next objection concerns preemption. Philosophers have tried to build many things out of counterfactuals, including, famously, accounts of causation (e.g., [Lewis, 1973a]).²² It is well known that that program runs into some of its most serious difficulties when it comes to preemption cases: putting it crudely, the trouble is that causation seems not to track the counterfactual relationships

²⁰We owe this objection to Craig Callender.

²¹See [Fine, 1975] for one of the early objections to Lewis's semantics for counterfactuals, and [Tooley, 2003] and [Schaffer, 2004] for recent discussion of the problems with the standard account.

²²Lewis's counterfactual account of causation is especially relevant in the present context since that account, like the counterfactual account of information, rules out vacuous counterfactuals by fiat.

in such cases. Consequently, fans of this approach to causation have needed to appeal to further apparatus, above and beyond the counterfactuals, to deal with preemption of various kinds. The obvious objection for the present counterfactual theory of information, then, is that it will run up against analogous cases of preemption, and that we'll need to appeal to further apparatus, above and beyond the counterfactuals, to deal with preemption of various kinds. Our response to this worry will be that the analogy doesn't go through: we'll suggest that, unlike causation, information does seem to track the counterfactual relationships in such cases.

The preemption problem for counterfactual theories of causation arises this way. In certain circumstances, it may be the case that c actually causes e , but if c were not to occur then some other event d would bring about e . A simple counterfactual account (or test) of causation cannot capture the causal relation between c and e , since the counterfactual 'if c were not to occur then e would not occur' is false. While there have been many attempts to deal with this problem for the counterfactual account of causation, it has proved surprisingly difficult to handle all preemption cases.²³

In contrast, a counterfactual account of information seems not to face a similar preemption problem at all; for where there is preemption, there is no information relation either. To see this, consider the following two cases.

Case 3: The room temperature is 70°F. The thermometer reads '70°F'. If the temperature had been different, the reading on the thermometer would have been different.

Case 4: The room temperature is 70°F. The thermometer reads '70°F'. But had the room temperature been different, a mischievous angel would have moved the pointer on the thermometer so that it would have read '70°F' anyway.

According to (S*), the thermometer's reading '70°F' in case 3 carries information about the room temperature's being 70°F because the relevant counterfactual is true. But according to (S*), the thermometer's reading '70°F' in case 4 does not carry information about the room temperature's being 70°F, since the counterfactual linking the two is false. This seems to be the right verdict about the case, given what we (following Shannon) want our notion of information to do: the thermometer reading in case 4 does not reduce uncertainty about the room temperature, hence no information about the room temperature is car-

²³One standard answer to traditional preemption problems is appeal to the existence of chains of counterfactually related intermediary events (i.e., intervening causal processes or causal chains) between causes and effects that are "cut off" in the case of preempted would-be-causes and effects [Lewis, 1973b]. While this move might be thought to handle some sorts of preemption, some authors have pointed to other sorts of preemption (e.g., so-called late and trumping preemption) that seem to resist such a solution ([Lewis, 1986b], [Schaffer, 2000]).

ried by the thermometer reading in that case.²⁴ ²⁵ Hence, the counterfactual account of information is not susceptible to preemption counterexamples.²⁶

5.3 Shope Cases

[Shope, 1978] provides the ingredients for constructing counterexamples to counterfactual analyses quite generally speaking (cf. [Martin, 1994]). However, we believe that putative counterexamples made according to Shope’s recipes are unsuccessful against the counterfactual account of information.

The cases to which Shope points our attention are those in which the counterfactual satisfaction of the antecedent of a counterfactual conditional appealed to in some analysis has unanticipated results on the target analysandum — forcing either its unexpected presence or absence, as the case may be. Shope uses this idea to show how an analysandum can be present when the counterfactual proposed in its analysis is in fact false, and alternatively that the counterfactual can be true when the analysandum is absent.

Since the counterfactual account is proposed as a general theory — one intended to hold for arbitrary properties F and G , a critic can raise problems for it by showing that there is a difficulty with particular properties. Consider, then, the following two potential counterexamples to the counterfactual account of information based on Shope’s recipes.

Case 5: Let it be that Gy iff $(\exists H)(x$ ’s being F carries information about y ’s being H). Then we have

$$\begin{aligned} \neg Gy &\leftrightarrow \neg(\exists H)(x\text{'s being } F \text{ carries information about } y\text{'s being } H). \\ &\leftrightarrow (H)(x\text{'s being } F \text{ fails to carry information about } y\text{'s being } H). \\ &\rightarrow x\text{'s being } F \text{ fails to carry information about } y\text{'s being } G. \end{aligned}$$

Thus, on this choice of G , the counterfactual supposition that $\neg Gy$ entails that there is not the information relation that the counterfactual account aims to analyze. Nonetheless, the critic might insist, our understanding of G seems to leave it open whether $\lceil Fx \rceil$ is true or false in various worlds; therefore, the critic can just stipulate that $\lceil \neg Fx \rceil$ holds in all the closest $\neg Gy$ -worlds. But if that is so, then \lceil if y were not G , then x would not have been $F \rceil$ is true. Furthermore, since nothing appears to prevent us from stipulating that there *are* some worlds in which $\lceil \neg Gy \rceil$ is true, it

²⁴ Cf. [Dretske, 1981]: “In saying that the conditional probability (given r) of s ’s being F is 1, I mean to be saying that there is a nomic (lawful) regularity between these two event types, a regularity which precludes r ’s occurrence when s is not F ” (245). In preemption cases, r ’s occurrence is not precluded when s is not F , since some other event might bring about r .

²⁵On the other hand, (S*) allows that the thermometer in case 4 might carry the disjunctive information that either the room is 70° F or that an angel moved the thermometer to read 70° F. We’ll ignore this complication.

²⁶For this reason, the proponent of the counterfactual account is likely to agree with Dretske’s argument that the information relation is not a causal relation ([Dretske, 1981], 26–39).

would seem that we have described a case in which the counterfactual is non-vacuously true despite the lack of an informational relation.

Case 6: Let it be that Gy iff $(\exists H)(x$'s being F doesn't carry information about y 's being H). Then we have

$$\begin{aligned} \neg Gy &\leftrightarrow \neg(\exists H)(x\text{'s being } F \text{ fails to carry information about } y\text{'s being } H). \\ &\leftrightarrow (H)(x\text{'s being } F \text{ carries information about } y\text{'s being } H). \\ &\rightarrow x\text{'s being } F \text{ carries information about } y\text{'s being } G. \end{aligned}$$

Here, the counterfactual supposition that $\neg Gy$ entails that x 's being F carries information about y 's being G . And, as before, it can seem that the critic is within her rights to stipulate that $\lceil Fx \rceil$ holds in all the closest $\neg Gy$ -worlds. If so, then the counterfactual \lceil if y were not G , then x would not have been $F \rceil$ is false (hence not non-vacuously true), even though the informational relation under discussion is present.

These cases purport to show that the counterfactual analysans proposed here is neither necessary (case 6) nor sufficient (case 5) for information. But we remain unconvinced: we can find no reason for accepting the stipulations on which the cases crucially depend.

To see why, consider case 5 first. We agree that, with G understood as suggested, the counterfactual supposition that $\neg Gy$ entails that x 's being F fails to carry information about y 's being G . But is it true, as the critic claims, that this allows for the stipulation that $\lceil \neg Fx \rceil$ holds in all the closest $\neg Gy$ -worlds? Not if the counterfactual view of information is correct.²⁷ On the other hand, the critic has provided no reason for believing that the stipulation is allowable, and it's hard to see that we could resolve this question on theory-neutral grounds absent firm pre-theoretical intuitions about what is and is not consistent with the presence of particular informational relations. Unfortunately, and unsurprisingly given that 'information' (in the sense at issue) is a technical neologism, we doubt that there are such firm intuitions.

Likewise, consider case 6. Once again, we agree that, on the critic's suggested understanding of G , the counterfactual supposition that $\neg Gy$ entails that x 's being F carries information about y 's being G . But, once again, we see no reason to accept the critic's further stipulation that $\lceil Fx \rceil$ holds in all the closest $\neg Gy$ -worlds — a stipulation that is inconsistent with the account we are defending.²⁸ As before, the critic has failed to provide any independent motivation for this stipulation, and it would seem that the sort of independent motivation that we

²⁷For, on that view, of course, the presence of the informational relation is tantamount to the non-vacuous truth of the counterfactual \lceil if y were not G , then x would not be $F \rceil$. Consequently, the counterfactual supposition that $\neg G$, conjoined with the counterfactual account, entails that \lceil if y were not G , then x would not be $F \rceil$ is *false* — a result that is inconsistent with the critic's stipulation that $\lceil \neg Fx \rceil$ holds in all the closest $\neg Gy$ -worlds.

²⁸For if the counterfactual supposition entails that the informational relation is present, then this, together with the counterfactual account, entails that $\lceil Fx \rceil$ *fails* in all the closest $\neg Gy$ -worlds.

are after could only be supplied by much firmer intuitions about information than we actually have.

Indeed, the difficulty at hand seems inevitable in any attempt to construct Shope-type counterexamples to the present proposal; hence, our response is not limited to the particular cases considered. Shope's recipe is supposed to apply to proposals of the form $\lceil p \text{ is analyzed by } (q \square \rightarrow r) \rceil$. His recipe was to find a particular q such that either $q \leftrightarrow \neg p$, while ensuring that r in all the nearest q -worlds, or else $q \leftrightarrow p$, while ensuring that $\neg r$ in all the nearest q -worlds. In the context of the present proposal, $p = \lceil x \text{'s being } F \text{ carries information about } y \text{'s being } G \rceil$, $q = \lceil \neg Gy \rceil$, and $r = \lceil \neg Fx \rceil$; thus, case 5 is an instance of the first recipe, while case 6 is an instance of the second. This means that, to carry out the recipes, we need to choose a property G such that the counterfactual satisfaction of $\lceil \neg Gy \rceil$ entails either that there is not the requisite informational relation (as in case 5) or that there is (as in case 6). However, unlike several of the targets against which Shope deploys his recipes in his original paper (e.g., epistemic justification, moral goodness) information is a defined technical notion, so we can't appeal to intuition to decide whether various cases fall under its extension. Thus, the only way to ensure the entailments needed for the recipe is to build various facts about informational relations directly into the specification of the property G (as we have done in constructing cases 5 and 6). Fair enough — but, depending on the metaphysics of informational relations, stipulations made about q ($\lceil \neg Gy \rceil$) made in the first step of the recipe may well impose constraints on what we can stipulate about r ($\lceil \neg Fx \rceil$) at the second. Our (quite general) complaint against Shope-type counterexamples to the counterfactual account of information is that the account does in fact impose such constraints, that these constraints prevent the execution of the Shope recipes against the account, and that these constraints can only be overcome by begging the question against the account.²⁹

It seems to us, then, that the Shope recipes are not sufficient, by themselves, to impugn the counterfactual account of information we are proposing. These recipes show how to construct would-be counterexamples to the account, but there is no reason to believe they can be carried out, and so they provide no reason for abandoning the counterfactual account of information.

6 Conclusion

Information has become a centrally important notion in contemporary philosophy. We have proposed understanding information in terms of counterfactual relationships, rather than (as on standard accounts) probabilities. One motivation for our proposal is just that the novel construal of information may prove useful for various theoretical purposes. Another is that the counterfactual account sidesteps difficulties about the interpretation of the probabilities in probabilistic accounts. We have also noted that Dretske's theory, insofar as

²⁹As far as we can see, this response to Shope-style counterexamples is generalizable to other proposed counterfactual analyses of technical (i.e., non-folk) notions.

it makes crucial appeal to unreduced epistemic elements, can't be enlisted to do the reductive work that Dretske and others have hoped it will. Moreover, we have explained how the counterfactual account can be elaborated in a way that avoids this difficulty. The result, then, would be an objective counterfactual account of information.

While we have not claimed that this account is the only acceptable theory of information, we have argued that the approach avoids problems that threaten the standard account, and gets us the sorts of things we want from a theory of information. Moreover, it does not appear to be susceptible to the most obvious sorts of criticism that might be lodged against it. For these reasons, we believe that the objective counterfactual approach is worth taking seriously as a theory of information.³⁰

References

- [Dretske, 1971] Dretske, F. (1971). Conclusive reasons. *Australasian Journal of Philosophy*, 49:1–22.
- [Dretske, 1981] Dretske, F. I. (1981). *Knowledge and the Flow of Information*. MIT Press, Cambridge, Massachusetts.
- [Fine, 1975] Fine, K. (1975). Critical Notice: Counterfactuals. *Mind*, 84:451–458.
- [Gillies, 2000] Gillies, D. (2000). Varieties of propensity. *British Journal for the Philosophy of Science*, 51:807–835.
- [Goodman, 1954] Goodman, N. (1954). *Fact, Fiction, and Forecast*. Harvard University Press, Cambridge, Massachusetts.
- [Hawthorne, 2004] Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford University Press, Oxford.
- [Humphreys, 1985] Humphreys, P. (1985). Why propensities cannot be probabilities. *The Philosophical Review*, 94(4):557–570.
- [Humphreys, 2004] Humphreys, P. (2004). Some considerations on conditional chances. *British Journal of the Philosophy of Science*, 55:667–680.
- [Lewis, 1973a] Lewis, D. (1973a). Causation. *The Journal of Philosophy*, 70:556–567. Reprinted in [Lewis, 1986a], 159–172.
- [Lewis, 1973b] Lewis, D. (1973b). *Counterfactuals*. Basil Blackwell, Oxford.
- [Lewis, 1986a] Lewis, D. (1986a). *Philosophical Papers, Volume II*. Oxford University Press, New York.

³⁰We are grateful to Craig Callender, Fred Dretske, Agustín Rayo, and Rob Rupert for helpful discussion of these matters and comments on early drafts.

- [Lewis, 1986b] Lewis, D. (1986b). Postscripts to “Causation”. In Lewis, D., editor, *Philosophical Papers, Volume II*, pages 172–213. Oxford University Press, New York.
- [Lewis, 1994] Lewis, D. (1994). Humean supervenience debugged. *Mind*, 103(412):473–490.
- [Loewer, 1983] Loewer, B. (1983). Information and belief. *Behavioral and Brain Sciences*, 6:75–76.
- [Loewer, 2001] Loewer, B. (2001). Determinism and chance. *Studies in the History and Philosophy of Science Part B: Studies in the Philosophy of Modern Physics*, 32(4):609–620.
- [Loewer, 2005] Loewer, B. (2005). David Lewis’ Humean theory of objective chance. *Philosophy of Science*. In press.
- [Martin, 1994] Martin, C. B. (1994). Dispositions and conditionals. *Philosophical Quarterly*, 44(174):1–8.
- [Martin, 1975] Martin, R. (1975). Empirically conclusive reasons and scepticism. *Philosophical Studies*, 28:215–217.
- [Martin, 1983] Martin, R. (1983). Tracking Nozick’s sceptic: A better method. *Analysis*, 43:28–33.
- [Meskin and Cohen, 2006] Meskin, A. and Cohen, J. (2006). Spatially agnostic informants and the epistemic status of photography. In Walden, S., editor, *Blackwell Guide to Photographs*. Blackwell, New York. In press, available at <http://aardvark.ucsd.edu/~joncohen/perception/agnosticism.pdf>.
- [Nozick, 1981] Nozick, R. (1981). *Philosophical Explanations*. Cambridge University Press, Cambridge.
- [Scarantino, 2005] Scarantino, A. (2005). Did Dretske learn the right lesson from Shannon’s theory of information? Ms., University of Pittsburgh.
- [Schaffer, 2000] Schaffer, J. (2000). Trumping preemption. *The Journal of Philosophy*, 97:165–181.
- [Schaffer, 2004] Schaffer, J. (2004). Counterfactuals, causal independence, and conceptual circularity. *Analysis*, 64(4):299–309.
- [Shannon, 1948] Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656.
- [Shope, 1978] Shope, R. K. (1978). The conditional fallacy in contemporary philosophy. *The Journal of Philosophy*, LXXV(8):397–413.
- [Suppes, 1983] Suppes, P. (1983). Probability and information. *Behavioral and Brain Sciences*, 6:81–82.

- [Tooley, 2003] Tooley, M. (2003). The Stalnaker-Lewis approach to counterfactuals. *The Journal of Philosophy*, 100:371–377.
- [Vogel, 1992] Vogel, J. (1992). lottery paradox. In Dancy, J. and Sosa, E., editors, *A Companion to Epistemology*. Blackwell, Oxford.