Shell games, information, and counterfactuals
Andrea Scarantino

To cite this Article

To link to this Article: DOI: 10.1080/00048400802215448
URL: http://dx.doi.org/10.1080/00048400802215448
Cohen and Meskin [2006] have recently proposed a novel counterfactual account of information. I argue that it is a step down from its intended target, namely Dretske’s [1981] theory of information. Thinking of the information carried by signals in terms of counterfactuals leads to falsely diagnosing *bona fide* instances of information transmission as not being instances of information transmission at all, with major loss of explanatory power.

Under what conditions does x’s being F carry information about y’s being G, where ‘x being F’ and ‘y being G’ designate events? Cohen and Meskin [2006: 335] have recently proposed the following answer: x’s being F carries information about y’s being G iff the counterfactual ‘if y were not G then x would not be F’ is non-vacuously true. This is presented as an improvement over Dretske’s influential theory of information, according to which x’s being F carries the information that y is G iff p(y is G | x is F & background knowledge k) = 1 and p(y is G | background knowledge k) < 1 [1981: 65].

The purpose of this paper is to compare the two theories with respect to a simple example about a shell game, and argue for the superiority of the Dretskean account on technical and principled grounds. Here is the example:

[S]uppose that there are four shells and a peanut is located under one of them. In attempting to find under which shell the peanut is located, I [Recipient A] turn over shells 1 and 2 and discover them to be empty. At this point you [Recipient B] 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].

---

1I want to thank Eddy Nahmias, Gualtiero Piccinini, Ben Sheredos and two anonymous referees of this journal for their very useful feedback on previous drafts of this paper. I also want to thank an anonymous referee of The Philosophical Quarterly for excellent comments on an expanded version of this paper that have helped with this shorter version as well.

2In the absence of the non-vacuity proviso, anything would carry information about y’s being G whenever it is necessary that y is G. In such case, ‘if y were not G then x would not be F’ would have an impossible antecedent, and counterpossibles are commonly assumed to be vacuously true (Lewis [1973]; but see Brogaard and Salerno [2007] for a dissenting opinion).

3Dretske [1981: 245] also formulated an alternative theory of information, according to which x’s being F carries the information that y is G iff there is a nomic regularity such that, given that x is F, y must be G by virtue of a law of nature or logic. I will not discuss this second formulation in what follows, but I want to emphasize that the virtues of the Dretskean account extolled in this paper do not apply to the nomic regularity version.

4There are two assumptions here. The first is that the two recipients know that the peanut is under one of the four shells. The second is that shells 1 and 2 are put back down after having been turned over, without changing the peanut’s location.

---

Andrea Scarantino
Let $k_A$ and $k_B$ be the knowledge states of recipients A and B before shell 3 is turned over. By assumption, $k_A$ includes the knowledge that shells 1 and 2 are empty, whereas $k_B$ comprises only the knowledge that the peanut could be under any one of the four shells. Let the signal be the emptiness of shell 3. What information does it carry?

Consider the answers provided by Dretske’s [1981] theory. According to it, signals do not carry information simpliciter, but relative to background knowledge. Relative to $k_A$, the signal carries the information that the peanut is under shell 4, because $p(\text{peanut is under shell 4}|\text{shell 3 is empty } \& \ k_A) = 1$ and $p(\text{peanut is under shell 4}|k_A) = 1/2 < 1$. Relative to $k_B$, the same signal carries the information that the peanut is either under shell 1 or under shell 2 or under shell 4, because $p(\text{peanut is under shell 4 or under shell 1 or under shell 2}|\text{shell 3 is empty } \& \ k_B) = 1$ and $p(\text{peanut is under shell 4 or under shell 1 or under shell 2}|k_B) = 3/4 < 1$.

Finally, the emptiness of shell 3 carries the information that the peanut is not under shell 3 relative to both $k_A$ and $k_B$, because $p(\text{peanut is not under shell 3}|\text{shell 3 is empty } \& \ k_A) = p(\text{peanut is not under shell 3}|k_A) = 1/2 < 1$ and $p(\text{peanut is not under shell 3}|k_B) = 3/4 < 1$.

Consider now Cohen and Meskin’s [2006] theory. According to it, the emptiness of shell 3 carries information about any event ‘y is G’ such that ‘if y were not G, shell 3 would not be empty’ is non-vacuously true (call this the information counterfactual). Under this new theoretical framework, the signal continues to carry information about the peanut’s not being under shell 3. This is because it is non-vacuously true that, if the peanut were under shell 3, shell 3 would not be empty. An entailment relation underwrites the truth of this counterfactual: the fact that shell 3 is empty entails that the peanut is not under shell 3. To generalize, whenever x’s being F necessitates y’s being G by virtue of a nomological connection between x’s being F and y’s being G, the information counterfactual comes out true, and Cohen and Meskin’s [2006] account correctly diagnoses the transmission of information. Trouble starts when nomological connections do not underwrite information counterfactuals. Consider the following counterfactual:

$$C = \text{If the peanut had not been under shell 4, shell 3 would not have been empty.}$$

The emptiness of shell 3 carries information about the peanut’s being under shell 4 just in case counterfactual C is (non-vacuously) true. Now, is C true or false? According to Cohen and Meskin [2006], the answer is: it depends. They write:

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 [i.e. shell 3 would not have been empty]’ would be true; on these distributions, then, the examination of shell 3 provides the information that the peanut was under shell 4... 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.

[Cohen and Meskin 2006: 346]

The problem with this line of reasoning is that, in the absence of (1) a criterion to establish what generates these mysterious possibility distributions, and (2) a criterion to establish which of the various distributions matter to compute the truth value of C, Cohen and Meskin’s [2006] theory leaves the question whether the emptiness of shell 3 carries information about the peanut’s being under shell 4 without a determinate answer.5

Is there a satisfactory way to resolve this indeterminacy? One option would be to allow the distribution of possibilities to vary contingently on what the recipients already know about the possibilities at the source. This option, however, is unavailable to Cohen and Meskin, who emphasize that the distribution of possibilities relevant for information transmission should be independent of background knowledge (non-doxastic): ‘[I]f one wants to sustain the non-doxastic character of information, one... is committed to the existence of non-doxastic facts of the matter about the distribution of possibilities’ [2006: 344]. This non-doxastic commitment grounds Cohen and Meskin’s claim that their theory is a naturalistic improvement over Dretske’s [1981], where the information transmitted by a signal is relativized to background knowledge k.6

Another option to resolve the problem of indeterminacy would be to rely on the standard semantics for counterfactuals [Lewis 1973], in the hope that it delivers the right verdicts concerning the truth-value of information counterfactuals. But this option would not work either, because C is arguably false under Lewis’s [1973] semantics.7 Leaving out a few details, the standard semantics for counterfactuals tells us that ‘if y were not G, then x would not be F’ is true at our world just in case, if there are ‘y is not G’-worlds, all the ‘y is not G’-worlds closest to our world are ‘x is not F’-worlds.

Question: Are all closest possible worlds in which the peanut is not under shell 4 worlds in which shell 3 is not empty? To put it otherwise, are all peanut not under shell 4-worlds without gratuitous departure from actuality also shell 3 is not empty-worlds? The answer is negative, because there are possible worlds relevantly similar to the actual world in which the peanut is not under shell 4, but shell 3 is still empty. If this is the case, it is false that if the peanut had not been under shell 4, shell 3 would not have been empty, and true that it might still have been empty through an alternative causal route, for instance if the peanut had been put under shell 2.

5The problem here is not merely that we are not told whether the counterfactual is ultimately true or false in the circumstances described in the Shell Game, but that we are not told what exactly would make the counterfactual true or false and how we could go about figuring out its truth-value.

6According to Cohen and Meskin [2006], another advantage of their account is that it does not make use of conditional inverse probabilities. Their use in the context of Dretske’s [1981] project has been criticized on the grounds that neither a frequentist nor a propensity nor a degree of belief interpretation of such probabilities appears to be compatible with the overarching objectives of Dretske’s project [Loewer 1983].

7An additional problem, acknowledged by Cohen and Meskin [2006: 342] but not considered fatal, is that it is unclear whether the doxastic element is entirely absent from Lewis’s [1973] semantics for counterfactuals. This is because ‘the relative importance of respects of comparison that underlie the comparative similarity of worlds...[is]...a highly volatile matter, varying with every shift of context and interest’ [1973: 92]. I will not discuss this further threat to the applicability of Lewis’s semantics to Cohen and Meskin’s [2006] project.
Suppose the peanut ended up under shell 4 because Tony the Thimble Rigger placed it under it after a fast sequence of reshuffles in the course of a Shell Game. Recipients A and B are players, and their job is to try to figure out the location of the peanut and bet money on it. Tony follows the policy of, say, placing the peanut randomly under even-numbered shells. Consider now possible antecedent-worlds in which, by virtue of a different set of neurons firing in his brain, Tony does not place the peanut under shell 4. In order for ‘if the peanut had not been under shell 4, shell 3 would not have been empty’ to be true, it would have to be the case that in all closest possible worlds in which the peanut is not under shell 4, shell 3 is not empty. But in some such worlds the peanut would end up under shell 2 (the other even-numbered shell), thereby leaving shell 3 empty. If this is true, there is a peanut not under shell 4/shell 3 empty world as close to the actual world as a peanut not under shell 4/shell 3 not empty world, which suffices to conclude that information counterfactual C is false.

If all of this is right, Cohen and Meskin’s [2006] theory fails to deliver on the promise of offering a non-doxastic theory of information that improves upon Dretske’s [1981] influential account. In the absence of a clear non-doxastic account of what determines the relevant possibility distributions and how to pick among them, the counterfactual account is at best incomplete, and on this ground alone unable to compete with Dretske’s [1981] theory.

On the other hand, if we rely on Lewis’s standard similarity-based semantics for counterfactuals, Cohen and Meskin’s [2006] account leads to counterintuitive results. Whereas according to Dretske’s [1981] theory the emptiness of shell 3 carries information about the peanut’s being under shell 4 to both A and B, and differentiates the information it carries to each of them, under a standard semantics for counterfactuals, Cohen and Meskin’s [2006] theory entails that the emptiness of shell 3 carries no information about the peanut’s being under shell 4.

Should Cohen and Meskin [2006] try to solve the indeterminacy problem by exploring other options for developing a satisfactory non-doxastic counterfactual account? So far I have called into question the technical details of their theory. I want to conclude by sketching a principled reason why informational semanticists should be suspicious of non-doxastic accounts of information. Cohen and Meskin [2006: 340] have suggested that the relativization of information to background knowledge is at best a reflection of our ‘ordinary ways of thinking about information’, to which the ‘technical notion of information’ should not be beholden in its attempt to satisfy reductive aspirations.

If what is at stake were only the ability of a theory of information to accommodate ordinary intuitions, this line of thought would be unexceptionable, and seemingly counterintuitive results could be shrugged off. What
is at stake, however, is something more significant, namely the extent to which a ‘technical notion’ of information can play a causal role in the explanation of behaviour. The assumption that information is explanatorily relevant underlies the central role information plays in biology, neuroscience and cognitive science. I am convinced that any satisfactory reductivist theory of information should make sense of this central explanatory role, a view Dretske [1983] explicitly defended.

But information does not affect behavior just by being ‘out there’ in the world. The way in which information acquires causal efficacy is by being consumed by information recipients. In turn, the primary form of consumption one can make of any piece of information is to learn it.\textsuperscript{10}

Crucially, learning is recipient-dependent, in the sense that it depends on what the recipient already knows. Different recipients will learn different things from the same signal, contingently on their background knowledge.

If the information a signal carries is identified with what can be learned from it, as proposed by Dretske [1981], information will track the behavioural impact of the signal in its entirety. If the information a signal carries is not identified with what can be learned from it, as proposed by Cohen and Meskin [2006], information will track the behavioural impact of the signal only partially, namely only with respect to the subset of learnable things that qualify as information. There will be things one can learn from a signal which go beyond the information the signal carries, but have a causal impact on behaviour.

Consider the Shell Game once again. Player A, who has already observed that shells 1 and 2 are empty, can learn from the emptiness of shell 3 that the peanut is under shell 4, whereas player B can learn only that it is either under shell 1, 2 or 4, each with probability 1/3. Finally, both players can learn that the peanut is not under shell 3. If player A bets a large sum with Tony the Thimble Rigger that the peanut is under shell 4, whereas B refrains from doing so, we can on Dretske’s [1981] theory explain this behavioural difference by citing the different information they received from the emptiness of shell 3.

This is not possible on Cohen and Meskin’s [2006] theory, according to which the emptiness of shell 3 carries the same information relative to A and B. More specifically, I have argued that under a standard semantics for counterfactuals the emptiness of shell 3 carries information about the peanut’s not being under shell 3, but not information about the peanut’s being under shell 4. Under these premises, we can invoke information \textit{sensu} Cohen and Meskin [2006] to explain why neither A nor B bets money on the peanut’s being under shell 3. But we will be unable to explain much beyond that.

For instance, we won’t be able to invoke the different information carried by the signal to the two players to explain why A, unlike B, made a bet that the peanut is under shell 4. The explanation of all differences in the behavioural impact of a signal owing to differences in background

\textsuperscript{10}It is primary because anything else information users do with information presupposes that such information is learned.
knowledge will have to be handed out to a separate and yet to be specified non-informational theory connecting information, learning and behaviour.

The Shell Game example is representative of many cases of information transmission where what recipients learn from signals is contingent on background knowledge. For instance, it is in the light of background knowledge that bees learn about nectar location from bee dances, rats learn about incoming electric shocks from the sounds to which they have been fear-conditioned, vervet monkeys learn about dangers from predator calls, tourists learn about unavailable rooms from ‘no vacancy’ signs, doctors learn about diseases from X-rays, and so on.

In all these cases, non-doxastic theories of information in the style of Cohen and Meskin [2006] will have a hard time tracking the behavioural impact of signals in the light of the information they carry. This will lead to a major loss of explanatory power, and the consequent inability to make sense of the central role information plays in the sciences of mind, where information is primarily invoked to explain behaviour.

This leads me to conclude that, as things stand, Dretske’s [1981] theory is more complete, more intuitively compelling and more explanatorily powerful than Cohen and Meskin’s [2006] theory. Whether Dretske’s [1981] theory is overall satisfactory as a naturalistic theory of information is another question, and one I won’t be able to tackle in this paper.11

Georgia State University

References


11I argue in Scarantino [2007] that Dretske’s [1981] theory ultimately fails as a naturalistic theory of information, but for reasons different from those suggested by Cohen and Meskin [2006].