

# Believing For a Reason

John Turri  
University of Waterloo  
[john.turri@gmail.com](mailto:john.turri@gmail.com)

*Abstract:* This paper explains what it is to believe something for a reason. I argue that you believe something for a reason just in case the reason non-deviantly causes your belief. In the course of arguing for my thesis, I present a new argument that reasons are causes and offer an informative account of causal non-deviance.

## 1. Preliminaries

Imagine a juror with a true belief that the defendant is guilty. Having paid close attention throughout the trial, she has impeccable reasons for thinking so. But she disregards these good reasons and instead believes he's guilty because the quarter turned up heads! (Heads he's guilty, tails he's not.) Our juror has good reasons for her belief. But she believes for a bad reason. Believing for a good reason is a valuable state, more valuable than merely having a good reason.

A complete epistemology requires a theory of believing for a reason. Maybe we know some things despite lacking reasons. Call such

knowledge *baseless*. Even if baseless knowledge is possible, surely not all knowledge is baseless. At least some knowledge is reason-based. Inferential knowledge is like this. You have inferential knowledge only if your inferential belief is held for a good reason.

We not only believe for reasons; we act for reasons too. There is a presumption in favor of a unified account of believing and acting for reasons. If *believing for a reason* is a causal relation, then *acting for a reason* probably is too.

So we have at least three motivations to better understand the epistemic basing relation (i.e. believing for a reason): it's a source of value, it must figure in a complete epistemology, and it affects how we should think about action.

Many epistemologists treat the causal theory of the basing relation as *the default position* (compare Plantinga 1993a: 69, Pollock 1986: 37, Huemer 2001: 56, Mittag 2002). You might think it owes its default status to one or more compelling arguments — and you'd be wrong. We can piece together various motivations offered here and there. But noticeably lacking is a clear, explicit argument.<sup>1</sup> This is sur-

---

<sup>1</sup> Swain 1985: 73–74 (see also Swain 1981: 81–82) motivates the view by appealing to the fact that it helps explain the role that non-belief states (esp. perceptual experiences) play in acquiring perceptual knowledge. But this this doesn't distinguish the causal theory from its competition. Its competitors could easily explain the relevance of perceptual experiences by pointing out that they typically provide an adequate basis of perceptual beliefs. In other words, even non-causal theorists can agree that perceptual beliefs are *based on* perceptual experiences, and thereby accommodate the commonsense view that they feature centrally in the acquisition of perceptual knowledge. Audi 1983, 1986 perhaps comes closest to offering something like an argument for the causal theory, but one is challenged to say just how the argument supposedly goes.

prising. And it gets worse. The causal theory also suffers from a serious and widely recognized outstanding liability: the deviance problem. Supported by little or no argumentation and hampered by a serious liability, how could the causal theory enjoy default status? Doubtless some will credit an unflattering source: philosophical fad.

I aim to change all that. In what follows I present a clear, explicit and intuitive argument for the causal theory. I also solve the deviance problem by presenting an informative account of causal non-deviance.

Several points are in order before proceeding. First, ‘R’ names a reason and ‘B’ a belief. Second, causation should be understood broadly to include overdetermination.<sup>2</sup> Third, R need not be *the* cause (or *the* causal sustainer) of B. It is enough that R is *a* cause, or *part of* the cause.<sup>3</sup> This occurs all the time: we believe many things for multiple reasons. Fourth, save for one detail, I say little about causation. We share a robust enough conception of causation to meaningfully discuss my proposal. In any case, a theory of causation falls beyond the scope of this paper. Finally, I pass freely between ‘believing for a

---

<sup>2</sup> Later you might wonder how well this comports with my claim that causes are difference-makers. If an event is causally overdetermined, then does any one of the overdetermining causes really make a difference? I fail to have clear intuitions about cases of overdetermination, though I recognize that some will answer ‘no’. The issues involved in sorting this out are legion and I cannot responsibly address them in this paper. Schaffer 2003 ably defends a view of overdetermination helpful to my cause; see also Loeb 1974: esp. 527–528. Thanks to John Greco for discussion on this point.

<sup>3</sup> It’s possible that we ought to think of basing as *a matter of degree*, so that the more central a reason is to the causation of a belief, the more the belief is based on a reason. Or perhaps we ought to think of basing as involving a *threshold*, so that a reason must make some minimal causal contribution to a belief in order for the latter to be based on the former at all. I set these potential complications aside in the main text.

reason' and 'based on a reason'.

Here's the plan for the paper. Section 2 argues that causation is necessary for basing. Section 3 presents my full analysis. Section 4 solves the triviality problem by presenting an informative account of causal non-deviance. Sections 5–7 respond to common concerns. Section 8 briefly sums up.

## 2. Causation is Necessary

(NC) R is among your reasons for believing Q (at time  $t$ ) only if R causes or causally sustains your belief (at  $t$ ). (I will subsequently suppress the parenthetical time-indexing.)

Here is the argument for NC.

*First*, reasons for believing are difference-makers. Suppose you thought the eyewitness testimony was the juror's reason for believing the defendant guilty. You then learn that the testimony *made no difference* to the juror's belief. It makes no difference to whether or how strongly the juror believes as she does.<sup>4</sup> You would rightly conclude

---

<sup>4</sup> As will become clear later in my discussion of Swain's theory, making a difference in the relevant sense requires more than mere counterfactual dependence. In a trivial sense, *everything* makes a difference to everything else. For any two things,  $x$  and  $y$ ,  $x$ 's presence makes at least the following difference to  $y$ :  $y$  is such that it co-exists with  $x$ , and thus  $y$  is such that it would have been different — insofar as it would have lacked relational properties that it actually has — had  $x$  not existed. Similar remarks apply to every one of  $x$ 's properties. We immediately recognize this relation as irrelevant (which is why this point is relegated to a footnote), though it's difficult to say precisely why. Suffice it to say that if you believe something for a reason, the reason certainly makes more than just this trivial difference.

that the testimony was not the juror's reason for believing.

*Second*, basing is not a brute relation. When a belief is based on a reason, they are related in some *further* way that accounts for it. Butch is a master butcher. Just by looking at a slab of meat he can tell within a pound how much it weighs. Butch knows he has this special ability, and that when he exercises it he's as reliable as a digital scale. A slab of meat gets wheeled in and placed on the scale. Butch directs his gaze thither and sees (i) that the scale reads '25 +/- 1, lbs.', and in virtue of his special butcher's ability, (ii) that the slab of meat weighs 25 pounds, give or take a pound. As it turns out, Butch forms a belief about the slab's weight for *one but not both* of these reasons. Something explains why just one of them is his reason. It's not just a brute fact.

To accept the first but not the second point is to embrace **fundamentalism**, the view that basing is a fundamental difference-making relation, on a par with causation and mereology.<sup>5</sup> Fundamentalism strikes me as fundamentally misguided. (I'm unaware of any discussion, let alone defense, of it in the literature.) In Butch's case it rules that it's just a brute fact that Butch believes for one reason but not the other—*nothing* explains why just one of them makes a difference. But we should expect an explanation for that. I don't know how to argue for this. I find it obvious, but must leave you to judge for yourself.

*Third*, NC explains why basing isn't a brute relation and why reas-

---

The discussion in the main text can thus be read as attempting to characterize the margin of difference-making beyond the trivial.

<sup>5</sup> Counterfactual theorists of causation will disagree that causation is a fundamental difference-making relation. I acknowledge this disagreement, but will not pursue it here. Thanks to Josh Schechter for discussion here.

ons are difference-makers. Causation provides the metaphysical underpinning of basing, which explains why it isn't brute. And causes are difference-makers, which explains why reasons are difference-makers. That causes are difference-makers is intuitive (compare Lewis 1973: 160–161, Menzies 2001, Sartorio 2005, and Schaffer 2005). Suppose you thought the bridge's faulty structure caused it to collapse. Then Ginny, the master engineer in charge of maintaining the bridge, tells you that the structural fault made no difference to the collapse. Provided you believed Ginny, you would rightly conclude that the structural fault did not cause the collapse.

*Fourth*, only a theory incorporating NC can explain both of these things. This fourth point requires considerable defense. What if there are non-causal relations that individually or collectively can explain both things? That would obviously undermine my argument. Accordingly I will argue that no non-causal theory proposed to date is satisfactory. This doesn't rule out that some other, as yet unarticulated non-causal theory will succeed. But it's a good start.

We find two main non-causal approaches to the basing relation. First, we have the **doxastic theory**. On this view, if you believe Q, and you believe P, and you judge that Q is good evidence to believe P, then your belief that Q is thereby among your reasons for believing P.<sup>6</sup>

---

<sup>6</sup> Korcz 1997, 2002 calls them "doxastic theories," and Kvanvig 1992: chapter 2 calls them "subjective theories." Proponents of this view, or close variants, include Foley 1987, Korcz 2000, Kvanvig 2003, Lehrer 1990, Pappas 1979a, and Tolliver 1981. Some theorists might say that the evidential belief is necessary to establish a basing relation, but so long as they grant that causation is also necessary to establish a basing relation, then their theory poses no challenge to NC.

The doxastic theory faces a serious problem. It entails that it is *impossible* to judge that you have two good reasons to believe P but believe for only one of them (compare Davidson 1963). But it's not impossible. It might be *irrational*, but not impossible. Consider this example.

**(EXHAUSTED)** Martin believes that Mars contains significant amounts of water buried just below its surface (Q). He judges that this is good evidence to believe that life exists elsewhere in the universe (P). Martin also is certain that the conditions for life are overwhelmingly abundant throughout the universe (S). He judges that this too is good evidence to believe that life exists elsewhere in the universe. But Martin is utterly exhausted and despairing from several grueling and fruitless months on the academic job market, which understandably and predictably impairs his cognitive functioning, especially at the present moment. He consequently neglects his evidential judgment about the relevance of subterranean Martian water, and bases his belief that life exists elsewhere solely on his belief that the conditions for life are abundant throughout the universe.

If this is a possible case, then the doxastic theory is false. And it certainly seems possible. The job market may be bad enough to make Martin slightly irrational. But it's not bad enough to make him impossible.

What of Martin's "neglected" evidential judgment? A doxastic theorist might respond as follows.<sup>7</sup> If by 'neglect' I mean 'forgot', then the case poses no threat to the sufficiency of the doxastic theorist's condition. If by 'neglect' I mean 'reject', then again the case poses no threat. In response, by 'neglect' I mean neither 'forgot' nor 'reject'. I simply mean that Martin is unaffected by this evidential belief, in the same way that Michael Stocker's jaded politician is unaffected by some of his moral beliefs (Stocker 1979: 741). There once was a young politician who cared about the plight of suffering people worldwide. He judged that it would be good to help them, and so he did. But he became jaded as he aged. He no longer cared about anyone outside his circle of friends and family. He still believed that it would be a very good thing to help the downtrodden, and knew there was much he could do to promote that goal. But he was no longer the least bit motivated to do so. Such failure of motivation, Stocker notes, "is commonplace." We can lose motivation in many ways, including through "spiritual or physical tiredness, through accidie, through weakness of body, through illness, through general apathy, through despair, through inability to concentrate, through a feeling of uselessness or futility" (Stocker 1979: 744). Stocker's case is widely regarded as a counterexample to the thesis known as *motivational internalism* in moral psychology, which says that judgments about what is good or right are *necessarily* motivating (see Dreier 1990: 10; Svavarsdóttir 1999: 163–5). In broad outline, Martin's case is to epistemic psychology as Stocker's is to moral psychology. Both cases involve the failure

---

<sup>7</sup> As an anonymous referee suggested.



of an evaluative belief to play its typical role.<sup>8</sup> And just as the lack of motivation in the politician's case needn't indicate loss of the relevant moral belief, the lack of basing in Martin's case needn't indicate loss of the relevant evidential belief.

Second, we have Marshall Swain's **counterfactual theory** of the basing relation (Swain 1981: ch. 3, esp. 86–87, 89–92). (I present Swain's view as simply and accessibly as I can without distortion, but it remains to some extent unavoidably technical.) According to Swain, even absent an actual causal relation between R and B, B is based on R if R would easily enough have caused B.<sup>9</sup> In a case where R\* does but R does not actually cause B, R would easily enough have caused B if and only if, had R\* not caused B but you still held B, R would have caused B.<sup>10</sup> Swain calls this relation "pseudo-overdetermination." Notice the right hand side of the 'if and only if' differs dramatically from saying that had R\* not caused B, R would have. The latter, but not necessarily the former, would be falsified if you would not hold B were R\* to not cause B.

---

<sup>8</sup> Notice how many of the items on Stocker's list of conditions tend to afflict job market candidates, like Martin. Many of us can no doubt sympathize from personal experience. Those lucky enough to have avoided the fate can simply peruse the posts and comment threads on the weblog *The Philosophy Smoker*, and its ancestor, the now defunct *Philosophy Job Market Blog*.

<sup>9</sup> Swain (1979: 30, 35–37; 1981: 91) sometimes seems to suggest that pseudo-overdetermination counts as a causal relation. If so, then the counterfactual theory cannot threaten NC. But as Swain (1981: chapter 2 and p. 86) himself recognizes, it's implausible that pseudo-overdetermination is a genuine causal relation. We best interpret him as rejecting NC.

<sup>10</sup> I suppress the causal-sustainment disjunct for ease of exposition. Swain defines the relation generally, but we focus here specifically on beliefs and reasons.

The counterfactual theory does not respect the fact that reasons are difference-makers. One thing can pseudo-overdetermine another without actually making a difference to it. Consider this example. The Red Sox are playing the Yankees for the American League Pennant. Curt Schilling gets the start in game seven for the Sox. He pitches brilliantly and the Sox win 2–0. Schilling obviously helped cause the Sox victory. As sports announcers and fans are apt to say, “Schilling is a difference-maker.” Pedro Martinez sat in the clubhouse the whole game. He made no difference to this Sox victory.<sup>11</sup> But had Schilling not pitched, Pedro would have pitched and won. So Pedro pseudo-overdetermines the Sox victory, but he made no difference.<sup>12</sup>

The counterfactual theory also faces a decisive counterexample. (This kind of counterexample is originally due to Joseph Tolliver 1981: 152–155.) Suppose Mallory believes Q solely on the basis of observation O. Mallory also believes the biconditional Q if and only if P. Together these two beliefs cause Mallory to believe P. Clearly Mallory’s belief that P is based on her belief that Q, but not vice versa. Yet the counterfactual theory entails otherwise: it entails that her belief that Q is based on her belief that P, because the latter pseudo-overdetermines the former. For had she still believed Q despite O not causing her to

---

<sup>11</sup> But what if his performance earlier in the series contributed to the Yankees’ poor performance on this night, you ask? I stipulate that no such thing has happened. Pedro hasn’t pitched yet in the series due to an illness, from which he finally recovers just before the start of game seven. Likewise for any other way you suggest Pedro might have had an effect on the game’s outcome.

<sup>12</sup> Note to baseball fans: this example was crafted before Pedro signed with the Mets, and even before the Sox played the Yankees in the 2004 post-season, back when this all seemed like just another fanciful philosophical thought experiment!

believe Q, her belief that P, along with her belief that Q if and only if P, would have caused her to believe Q.<sup>13</sup> The counterfactual theory falsely entails that her belief that Q is actually based on her belief that P.<sup>14</sup>

*Finally*, if only a theory incorporating NC can explain why reasons are difference-makers and why basing is not a brute relation, then NC will be part of *the best explanation* of those two things. So NC is true.<sup>15</sup>

### 3. The Complete Causal Account

A necessary condition does not a theory make. How shall we upgrade NC into a complete causal account?

The simplest proposal miscarries:

(#1) R is among your reasons for believing Q if and only

<sup>13</sup> It didn't *have* to turn out this way; this counterfactual isn't necessarily true. But it is true in the present case.

<sup>14</sup> My defense of the fourth point in the argument is incomplete in at least one respect. Causation is not the only difference-making relation. Consider mereological relationships. Molecules arranged in a certain way make it the case that there is a desk here and that it has certain features, but not by causing it to be here or have those features. Perhaps we can make sense of mereological relationships among beliefs: maybe your belief that P and your belief that Q are *parts* of your belief that P and Q. I doubt this strategy holds out much hope. But maybe an enterprising opponent can make something of it.

<sup>15</sup> Some epistemologists tout "gypsy-lawyer" cases as counterexamples to NC (e.g. Lehrer 1971, Harman 1973: 31–32, Lehrer 1990: 169–71, Korcz 2000, Kvanvig 2003). They're called "gypsy-lawyer" cases after Lehrer's original, which featured a "gypsy-lawyer." But these cases have failed to impress causal theorists. For example, Goldman (1979: 352: n. 8) says, "I find this example unconvincing." Pollock (1986: 81, n. 9) says, "I do not find [Lehrer's] counterexample persuasive." Swain (1981: 91) says, "I see no ground for claiming that the gypsy lawyer has knowledge." I agree with Goldman, Pollock and Swain: it has always seemed clearly false to me that the lawyer knows. But I've yet to find a plausible way to argue for this claim without simply begging the question. (Thanks to an anonymous referee for some very insightful remarks in connection with this.)

if R causes your belief.

Counterexamples abound. Al believes that he sees Sylvia, which causes him to get very nervous, which causes him to spill his tea on his leg, which in turn causes him to believe that he is in pain. Al's belief that he sees Sylvia causes his belief that he is in pain, but the former is clearly not his reason for holding the latter (Plantinga 1993a: 69, n. 8). Joe believes he's late to class, which causes him to quicken his pace, which causes him to slip and fall on his back, which causes to him to see the birds in the tree, which causes him to believe there are birds in the tree. Joe's belief that he's late for class causes him to believe there are birds in the tree, but the former is clearly not his reason for holding the latter (Pollock and Cruz 1999: 35–36). We need a way to rule out such cases.

We could bolster the biconditional's right side:

(#2) R is among your reasons for believing Q if and only if R is a *proximate* cause of your belief.

This handles Al's and Joe's cases. But it threatens to rule out far too much. Even if it turns out that sensory experience is never the proximate cause of our beliefs about our environment, even if a myriad of electrical and chemical events always intervene, surely sensory experiences would still be among our reasons for believing things about our environment.

This naturally leads to the following proposal:

(#3) R is among your reasons for believing Q if and only if R is a proximate *mental* cause of your belief.

This handles Al's and Joe's cases without ruling out sensory experience, but it faces other problems. Some believe the basing relation is transitive and so will reject #3 (and #2) because it makes *proximate* causation a necessary condition on basing, which rules out transitivity. Another objection is that #3 faces counterexamples involving deviant proximate mental causation. Through some random quirk—the result of a neural assembly malfunctioning—Wilt's belief that the lettuce has wilted is the proximate mental cause of his belief that the Patriots will win twelve games this season. But it certainly seems false that Wilt's belief that the lettuce has wilted is his reason for believing that the Patriots will win twelve games this season.

Notice that altering #3 to accommodate the transitivity-intuition won't solve this problem. Those who want to preserve transitivity will naturally invoke the ancestral of proximate mental causation. Let's define a *proximate mental causal chain* as a sequence of mental states  $m_1, m_2, m_3, \dots, m_n$ , where  $m_1$  is a proximate mental cause of  $m_2$ ,  $m_2$  is the proximate mental cause of  $m_3$ , and so on. We then get transitivity by defining the basing relation in terms of proximate mental causal chains:

(#4) R is among your reasons for believing Q if and only if a proximate mental causal chain leads from R to your belief.

But #4 also gives the wrong verdict in Wilt's case.

Since it appears we must invoke non-deviance in any case, we might as well define our target in terms of it:

(CA) R is among your reasons for believing Q if and

only if R non-deviantly causes your belief.<sup>16</sup>

This gives the right result in Al's and Joe's and Wilt's cases. But does it do so at the cost of *trivializing* the view? Can we say anything more informative than just, "Well, R didn't cause B in the right way"? I say we can, and the next section shows how.

But first notice one way the problem is less severe than it might initially appear. The causal deviance problem infects most if not all of our causal concepts (Huemer 1998: section 1.3). Doubtless a causal account of *murder* is correct. To murder someone you must cause his death. But that's not all. You must cause his death *in the right way*. You must intend to kill him, and your intention must appropriately figure into the causal explanation of his death. What does it mean for your intention to figure *appropriately*? The deviance problem strikes again (Davidson 1973). The same goes for a theory of perception. An object must cause you to have certain sensations for you to see it. But that's not all. It must cause your sensations *in the right way*. CA has a catalog of respectable partners in crime.

Correct as far as it goes, that response does not fully satisfy me. It would be nice if we could say something more.

#### 4. The Triviality Problem Solved

Consider this pair of cases:

**(OJ)** I sat at the table feeding baby Mario his break-

---

<sup>16</sup> We could just as easily have stated CA in terms of proximate causation. I leave it open whether we want to do this.

fast. I took a sip of orange juice and unwisely set the glass down within Mario's reach. His little hand darted out to retrieve the glass and its colorful contents. Spoon in one hand, baby in the other, I helplessly watched the glass tumble down, down, down. It *broke*.

(**CARAFE**) We just finished a delicious dinner. Maria turned to say something but in the process carelessly knocked a glass carafe, sending it careening from the table in my direction. Glass is fragile, so I reached out and caught it before it hit the ceramic tile floor. It *remained intact*.

In each case the outcome obtains because the glass is fragile. Yet we all recognize an important difference: the outcomes are not due *in the same way* to fragility. In OJ the glass breaks because it is fragile, and its breaking manifests its fragility. In CARAFE the glass remains intact because it is fragile, but its remaining intact does not manifest its fragility. Neither outcome obtains *only* because of fragility — in OJ Mario and the floor help out, in CARAFE my dexterity — but that does not spoil the point.

Consider also these cases:

(**BOIL**) You place a cup of water in the microwave and press start. The magnetron generates microwaves that travel into the central compartment, penetrate the water and excite its molecules. Soon the water

*boils.*

**(FIRE)** You place a cup of water in the microwave and press start. The magnetron generates microwaves that cause an insufficiently insulated wire in the control circuit to catch fire, which fire deactivates the magnetron and spreads to the central compartment. Soon the water *boils*.

The outcome in BOIL manifests the microwave's boiling power. The outcome in FIRE does not. We have a plain way to mark the distinction: in BOIL, but not FIRE, the microwave boils the water.

The examples highlight a general distinction that we all recognize between (A) an outcome *manifesting* a disposition and (B) an outcome happening *merely because* of a disposition. Outcomes can include conditions, events, and processes.

In the present context, I treat *manifestation* as a primitive. We understand it perfectly well, as my earlier examples demonstrate. It is familiar to us from our everyday dealings and extremely useful, perhaps necessary, when planning our lives as social beings (compare Sellars 1963: 11–12). True, we may want manifestation further clarified and explained, but this is something we would have wanted in any case (compare Plantinga 1993b: 5–6).

I propose to understand non-deviance in terms of the manifestation of cognitive traits.<sup>17</sup> I offer two closely related proposals: a necessary condition and a sufficient condition. If either is true or on

---

<sup>17</sup> For other important applications of the *manifestation* relation, see Turri forthcoming a, b and c.



the right track in explaining what happens in a broad range of standard cases, then we have solved the triviality problem.

The first proposal:

**Manifestation 1** (M1): R non-deviantly causes B only if R's *causing B* manifests (at least some of) your cognitive traits.<sup>18</sup>

A cognitive trait is a disposition or habit to form (or sustain) a doxastic attitude in certain circumstances (compare Peirce 1955: chapters 8 and 9). Consider some examples of cognitive traits. We habitually take experience at face value. We habitually trust what others say. We habitually reason in patterns, including those corresponding to the formal inference rules *modus ponens* and *modus tollens*, among others. Many of us unfortunately reason on the model of the less desirable *denying the antecedent* and other fallacious inference patterns. All those are plausibly innate habits, though no doubt modified and refined through experience.

Experience plays a larger role in acquiring other habits, such as those involved in learning a craft. Many such habits are articulable only demonstratively. The carpenter believes it best to strike the nail at *that* angle because things feel *this* way. The potter believes he should add more moisture to the clay because it has *that* feel. The shepherd

---

<sup>18</sup> It is instructive to compare this to Kantian conceptions of intentional action. According to Korsgaard 1997: 221, intentional action occurs “only when [the agent’s] action is the *expression* of her own mental activity” (emphasis added). Also compare Hempel’s views (1962: section 3.2; 1963a: 291–293; 1963b: section 4) on action explanation, dispositions, and “habit patterns.” It favors my theory that it is complemented by a promising analogous theory of acting for a reason.

judges a storm is brewing because the sky looks *that* way.

M1 correctly classifies our earlier problem cases. Al's belief that he sees Sylvia causes his belief that he's in pain, but not by manifesting his cognitive traits. Al isn't disposed to trust that Sylvia's presence indicates that he is in pain (unless there's more to Al and Sylvia's relationship than we've been told about!). Joe's belief that he's late to class causes his belief that there are birds in the tree, but not by manifesting his cognitive traits. Joe isn't disposed to trust that birds being in the tree indicates that he's late for class.<sup>19</sup> Wilt's belief that the lettuce has wilted causes his belief that the Patriots will win twelve games this season, but not by manifesting Wilt's cognitive traits. A random quirk is to blame instead.

Let me ward off a potential misreading of M1, especially as it relates to the cases just mentioned. For simplicity I focus on Al's case. Consider the chain of causes leading from Al's belief that he sees Sylvia to his belief that he's in pain. Presumably at least some of Al's cognitive traits manifest themselves at some links in the chain. For instance, surely *the pain's causing him to believe that he's in pain* is one such link. One might suspect, then,<sup>20</sup> that M1 fails to rule out what it's intended to rule out. That is, one might suspect that M1 fails to rule out that Al's belief that he's in pain is based on his belief that he sees

---

<sup>19</sup> When evaluating cases, we are entitled to assume that things are normal unless otherwise specified. If Al or Joe is disposed to trust in such strange connections, then that would have to be made explicitly part of the case. The examples are due to Plantinga 1993a: 69, n. 8 and Pollock and Cruz 1999: 35–36. They do not explicitly include them. If we *do* add those details to the cases, then it becomes quite plausible that the subject's belief is indeed based on the reason in question.

<sup>20</sup> As an anonymous referee suspected.

Sylvia, because the causal chain involves a manifestation of at least one relevant trait, in which case the necessary condition is met. This suspicion, while understandable, can be overcome. Granted, the causal chain does involve the manifestation of some cognitive traits. But this isn't enough to establish that *the relevant causal relation* manifests Al's cognitive traits. It isn't enough that the chain contain a link which manifests some cognitive trait.<sup>21</sup> The causal relation itself between R and B — that is, *R's causing B* — must manifest a cognitive trait, as per M1. But Al isn't disposed to trust that Sylvia's presence indicates that he's in pain. Yet he would need to have such a trait in order for the causal relation to manifest it.

Here is my second proposal:

**Manifestation 2 (M2):** R non-deviantly causes B if *R's causing B* manifests (at least some of) your cognitive traits. (Note that the 'only if' in M1 has become an 'if' here.)

---

<sup>21</sup> It's important to note that this fits into a perfectly general pattern. For an outcome to manifest a disposition, it isn't enough that the disposition manifest itself *somewhere or other* in the outcome's causal ancestry. A couple non-epistemological examples might help. Suppose Griffey's athleticism manifests itself in a spectacular catch, which causes me to get excited about my own prospects for fielding greatness, which causes me to train and practice, which in turn causes me to make a spectacular catch of my own one day. The manifestation of Griffey's athleticism caused me to make my catch, but my catch doesn't manifest Griffey's athleticism. Or suppose my musical ability manifests itself in a rousing performance of Mozart's *Alla Turca*, which causes me to want to excel at dancing too, which causes me to exercise and train, which causes me to one day perform a lovely pirouette. The manifestation of my musical ability caused me to perform a pirouette, but the pirouette doesn't manifest my musical ability. Elsewhere I show how Gettier cases display the same structure (Turri forthcoming b).

Against the backdrop of CA, M2 explains many things. It explains why perceptual experiences are often our reasons for believing things about our environment, why my belief that P is based on my beliefs that Q and (P if Q), why so many of our beliefs are based on the acquisition of testimony, why my intuition that causes are difference-makers is my reason for believing that causes are difference-makers, and why *that look* is the shepherd's reason for believing a storm is brewing. In each case the relevant causal connection manifests the subject's cognitive traits.

Combining M1 and M2 yields:

**Manifestation 3 (M3):** R non-deviantly causes B if and only if R's causing B manifests (at least some of) your cognitive traits.

Stitching together CA and M3 eliminates mention of causal deviance, yielding roughly this:

**Causal-Manifestation Account (CMA):** R is among your reasons for believing Q if and only if R's causing your belief manifests (at least some of) your cognitive traits.<sup>22</sup>

---

<sup>22</sup> Compare Goldman 1979: 346, Alston 1995: sections IV–VI, and Alston 2005: chapter 6, esp. sections iii–v. Wedgwood 2006 proposes a similar solution to the causal deviance problem for reasoning. Elsewhere I deploy the same basic idea to help infinitists about epistemic justification respond to a potentially serious objection (Turri 2009).

## 5. Generalizing

Does this account of causal deviance generalize to solve deviance problems in other areas? I'm sympathetic (though not beholden) to the idea. Here's a sketch of the strategy.

First note that deviance arises only when we consider evaluable performances of agents or systems, where responsibility of a particular sort is at stake. It's neither deviant nor non-deviant when a rock falls from a precipice, bounces several times, takes a remarkably unlikely trajectory, strikes your windshield and cracks it. This is unexpected, unlikely, peculiar, and undoubtedly exasperating but not deviant. By contrast suppose the coffee machine malfunctions, causing a fire that, amazingly, boils the water, which drips through the filter and yields a perfect pot of coffee. Irony aside, in such a case we would not say, "Boy that coffee machine sure made a good pot of coffee!", because the machine isn't appropriately responsible for the good pot of coffee, so it doesn't redound to its credit.

Next note that agents and systems possess stable features that make them capable of producing certain results in a normal environment. Indeed this is plausibly what *makes them* agents or the systems they are to begin with. Humans are equipped with traits that cause them to believe and act in certain ways when affected by certain stimuli. Coffee machines are equipped with features to receive and heat water, to hold coffee grounds, to infuse the grounds with heated water, and to collect the liquid coffee in a pot, all of which conspire to produce pots of coffee. It is when the belief mani-

festes the agent's cognitive traits, or the pot of coffee manifests the coffee machine's stable coffee-making features, that the causal chains are non-deviant and the result redounds to the credit or discredit of the agent or machine.

Abstracting sufficiently to view agents as just a special kind of system, let's define a system's T-relevant features as those features enabling it to produce a result of type T in a normal environment. We can then define non-deviant causation as follows: system S non-deviantly causes a token result *t* of type T just in case S's causing *t* manifests (at least one of) S's T-relevant features.

## 6. Situationism

Gilbert Harman and John Doris argue that experimental results from social psychology suggest that we humans don't have *moral* character traits, understood as "broad based," "relatively long-term stable disposition[s] to act in distinctive ways," which explain our behavior (Harman 1999 and Doris 2002; the quotes are from Harman). So isn't it unwise to rest our account of *anything* on the manifestation of character traits when their relevance — indeed, their very existence — has been powerfully called into question?

Philosophers vigorously dispute the significance of the experimental results, as well as Harman and Doris's interpretation of them. Many of the criticisms appear to have merit (e.g. Sreenivasan 2002, Kamtekar 2004, and Sabini and Silver 2005). But whatever one's position on this dispute, no one thinks the results suggest that

we lack *cognitive* traits. Indeed Harman and Doris's interpretation *presupposes* that we have a battery of virtuous cognitive traits, as I will now show.

To substantiate this point, let's focus on the Princeton Seminary cases that feature prominently in Doris's discussion. The seminarians are told that they're scheduled to give a presentation on the other side of campus about the Good Samaritan parable. As each seminarian departs for his presentation, things are set up so that he encounters in a corridor a confederate pretending to need help. Common sense predicts that a generous person will stop to help. Did the seminarians stop? The strongest predictor of whether they would was *how much time they thought they had to get to their destination*. Most who thought they were ahead of schedule stopped; fewer who thought they would be precisely on time stopped; and fewer still who thought they were already late stopped. The upshot of this result is supposed to be that a situational feature, not the seminarians' supposed generosity, best predicts helping behavior. Doris takes this as evidence against positing character traits, understood as stable, broad-based dispositions towards characteristic actions in the relevant circumstances.

In response, my point is that unless we presuppose that the seminarians possess a battery of cognitive virtues — in particular, unless we presuppose that they are generally attentive, perceptive, possessed of a good memory, and ready to competently draw any needed inferences — then we can't properly conclude that they lack the moral character trait of generosity. Why? Because if they *either*

- (i) don't see the confederate pretender, either because they're inattentive or imperceptive; *or*
- (ii) see him but don't remember that people who behave *like that* need help; *or*
- (iii) remember but don't draw the obvious inference that *this person needs help*,

then they lack the beliefs needed to trigger the disposition of generosity. And if they lack the appropriate beliefs, then their behavior can't count against the presence of the disposition.

## 7. Why Bother?

I treat manifestation as a primitive and use it to answer other questions. I justify this by appealing to our robust pretheoretical understanding of it, as evinced by our ability to easily sort cases involving it. But if we're going to rely on concepts we understand well pretheoretically, then why bother giving an account of *believing for a reason* in the first place? We well understand that concept pretheoretically too. So what do we gain?

We reveal its relationship to other concepts fundamental to our way of thinking about the world, particularly *causation*, *disposition* and *manifestation*. We gain greater understanding by placing the epistemic basing relation into a more general pattern (compare Davidson 1963: 10). More generally, to properly explain something we must of course employ concepts we already understand well. Otherwise our explanation would be obscure and unhelpful.



## 8. Summary

I hope to have convinced you of two things. First, you believe something for a reason just in case the reason non-deviantly causes your belief. Second, a reason non-deviantly causes your belief just in case its causing your belief manifests your cognitive traits. Combining these results yields the view that you believe something for a reason just in case the reason's causing your belief manifests your cognitive traits.<sup>23</sup>

---

<sup>23</sup> In writing this paper, I have accumulated too many debts to be confident that I recall them all. With apologies to those I might have forgotten, I thank Jason Baehr, Ali Eslami, Ben Fiedor, John Greco, Stephen Grimm, Allan Hazlett, Adam Leite, Sharifa Mohamed, Michael Pace, Jim Pryor, Bruce Russell, Mark Schroeder, Ernest Sosa, Jerry Steinhofner, Angelo Turri, and three anonymous referees for *Erkenntnis*.

## References

- Alston, William. 1995. "How to Think about Reliability." Reprinted in Sosa and Kim, eds.
- Alston, William. 2005. *Beyond "Justification": Dimensions of Epistemic Evaluation*. Ithaca: Cornell University Press.
- Audi, Robert. 1983. "The Causal Structure of Indirect Justification." In Audi 1993.
- Audi, Robert. 1986. "Belief, Reason, and Inference." In Audi 1993.
- Audi, Robert. 1993. *The Structure of Justification*. Cambridge: Cambridge University Press.
- Collins, John, Ned Hall and L. A. Paul, eds. *Causation and Counterfactuals*. Cambridge, Mass.: MIT Press.
- Cullity, Garret and Berys Gaut, eds. 1997. *Ethics and Practical Reason*. New York: Oxford University Press.
- Davidson, Donald. 1963. "Actions, Reasons, and Causes." Reprinted in Davidson 1980.
- Davidson, Donald. 1973. "Freedom to Act." Reprinted in Davidson 1980.
- Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Doris, John. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Dreier, James. 1990. "Internalism and Speaker Relativism." *Ethics* 101.1: 6–26.
- Foley, Richard. 1987. *The Theory of Epistemic Rationality*. Cambridge, Mass.: Harvard University Press.
- Goldman, Alvin. 1979. "What Is Justified Belief?" Reprinted in Sosa and Kim, eds.
- Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.
- Harman, Gilbert.. 1999. "Moral Philosophy Meets Social Psycho-

- logy: Virtue Ethics and the Fundamental Attribution Error.” Reprinted in Harman 2000.
- Harman, Gilbert.. 2000. *Explaining Value and Other Essays in Moral Philosophy*. Oxford: Oxford University Press.
- Hempel, Carl. 1962. “Rational Action.” Reprinted in Hempel 2001.
- Hempel, Carl. 1963a. “Explanation in Science and in History.” Reprinted in Hempel 2001.
- Hempel, Carl. 1963b. “Reasons and Covering Laws in Historical Explanation.” Reprinted in Hempel 2001.
- Hempel, Carl. 2001. *The Philosophy of Carl Hempel*. Ed. James H. Fetzer. New York: Oxford University Press.
- Huemer, Michael. 1998. “A Direct Realist Account of Perceptual Awareness.” Ph.D. Dissertation, New Brunswick Rutgers, the State University of New Jersey.
- Kamtekar, Rachana. “Situationism and Virtue Ethics on the Content of Our Character.” *Ethics* 114.3: 458–491.
- Korcz, Keith. 1997. “Recent Work on the Basing Relation.” *American Philosophical Quarterly* 34.2: 171–191.
- Korcz, Keith. 2000. “The Causal-Doxastic Theory of the Basing Relation.” *Canadian Journal of Philosophy* 30.4: 525–550.
- Korcz, Keith. 2002. “The Epistemic Basing Relation.” *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. URL = <<http://plato.stanford.edu/archives/win2002/entries/basing-epistemic/>>.
- Korsgaard, Christine. 1997. “The Normativity of Instrumental Reasoning.” In Cullity & Gaut, eds.
- Kvanvig, Jonathan. 1992. *The Intellectual Virtues and the Life of the Mind*. Savage, Maryland: Rowman & Littlefield.
- Kvanvig, Jonathan. 2003. “Justification and Proper Basing.” In Olson, ed.
- Lehrer, Keith. 1971. “How Reasons Give Us Knowledge, or the Case of the Gypsy Lawyer.” *Journal of Philosophy* 68.10: 311–313.

- Lehrer, Keith. 1990. *Theory of Knowledge*. Boulder, Colorado: Westview Press.
- Lewis, David. 1973. "Causation." *Journal of Philosophy* 70: 556–567.
- Loeb, Louis. 1974. "Causal Theories and Causal Overdetermination." *Journal of Philosophy* 71: 525–544.
- Menzies, Peter. 2004. "Difference-Making in Context." In Collins, Hall, & Paul, eds.
- Mittag, Daniel. 2002. "On the Causal-Doxastic Theory of the Basing Relation." *Canadian Journal of Philosophy* 32.4: 543–560.
- Olsson, Erik, ed. 2003. *The Epistemology of Keith Lehrer*. Dordrecht: Kluwer Publishing Co.
- Pappas, George. 1979a. "Basing Relations." In Pappas, ed.
- Pappas, George ed. 1979b. *Justification and Knowledge*. Dordrecht, The Netherlands: D. Reidel Publishing Co.
- Peirce, Charles. 1955. *Philosophical Writings of Peirce*. Ed. Justus Buchler. New York: Dover.
- Plantinga, Alvin. 1993a. *Warrant: The Current Debate*. Oxford: Oxford University Press.
- Plantinga, Alvin. 1993b. *Warrant and Proper Function*. Oxford: Oxford University Press
- Pollock, John. 1986. *Contemporary Theories of Knowledge*. Totawa, New Jersey: Rowman & Littlefield.
- Pollock, John and Joseph Cruz. 1999. *Contemporary Theories of Knowledge*, 2<sup>nd</sup> ed. New York: Rowman & Littlefield.
- Sabini, John and Maury Silver. 2005. "Lack of Character? Situationism Critiqued." *Ethics* 115: 535–562.
- Sartorio, Carolina. 2005. "Causes as Difference-Makers." *Philosophical Studies* 123: 71–96.
- Schaffer, Jonathan. 2003. "Overdetermining Causes." *Philosophical Studies* 114: 23–45.
- Schaffer, Jonathan. 2005. "Contrastive Causation." *The Philosophical Review* 114: 297–328.

- Sellars, Wilfrid. 1963. *Science, Perception and Reality*. New York: Routledge and Kegan Paul.
- Sosa, Ernest and Jaegwon Kim, eds. 2000. *Epistemology: An Anthology*. Malden, Mass.: Blackwell Publishers.
- Sreenivasan, Gopal. 2002. "Errors About Errors: Virtue Theory and Trait Attribution." *Mind* 111: 47–68.
- Stocker, Michael. 1979. "Desiring the Bad: An Essay in Moral Psychology." *Journal of Philosophy* 76.12: 738–753.
- Svavarsdóttir, Sigrún. 1999. "Moral Cognitivism and Motivation." *Philosophical Review* 108.2: 161–219.
- Swain, Marshall. 1979. "Justification and the Basis of Belief." In Pappas, ed.
- Swain, Marshall. 1981. *Reasons and Knowledge*. Ithaca: Cornell University Press.
- Tolliver, Joseph. 1981. "Basing Beliefs on Reasons." *Grazer Philosophische Studien* 15: 149–161.
- Turri, John. 2009. "An Infinitist Account of Doxastic Justification." *dialectica* 63:2, 209–218.
- Turri, John. Forthcoming a. "The Express Knowledge Account of Assertion." *Australasian Journal of Philosophy*.
- Turri, John. Forthcoming b. "Manifest Failure: The Gettier Problem Solved." *Philosophers' Imprint*.
- Turri, John. Forthcoming c. "Is Knowledge Justified True Belief?" *Synthese*.
- Wedgwood, Ralph. 2006. "The Normative Force of Reasoning." *Noûs* 40:4, 660–686.