ORIGINAL RESEARCH

Doxastic justification through dispositions to cause

Julius Schönherr[1] (ORCID)

Abstract
According to the standard view, a belief is based on a reason and doxastically justi-fied—i.e., permissibly held—only if a causal relation obtains between a reason and the belief. In this paper, I argue that a belief can be doxastically justified by a reason's mere *disposition* to sustain it. Such a disposition, however, wouldn't establish a causal connection unless it were manifested. My argument is that, in the cases I have in mind, the *manifestation* of this disposition would add *no positive epistemic feature* to the belief: a belief that is justified after the manifestation of a reason's causal powers must have already been justified before their manifestation. As a result, those who adhere to the standard causal view of the basing relation face a hard choice: they should either abandon the enormously popular view that doxastic justification has a basing requirement or modify their view of the basing relation.

1 Introduction

The epistemic basing relation is the relation that holds between a belief and the reason for which it is held. Although the sheer complexity of philosophical accounts of the basing relation is quite dazzling, there has, over the years, been a steady stream of proposals according to which a belief is based on a reason only if it is causally connected to the belief (e.g., Armstrong, 1973; Goldman, 1989; Harman, 1970; Moser, 1989; Turri, 2011). This position has been called the "standard" account (e.g., Korcz, 1996; McCain, 2016). Let's call philosophers who defend this view 'causalists'. They are united in their commitment to the following thesis:

Basing Causalism. An agent A's belief B is based on a reason R only if R is part of B's causal history.

✉ Julius Schönherr
  schoenherrjulius@gmail.com

1 Institute of Foreign Philosophy, Department of Philosophy and Religious Studies, Peking University, 5 Yiheyuan Rd., Haidian Dist., Beijing 100871, China

Next, all philosophers[1] who contribute to the analysis of the basing relation, except for Silva (2015),[2] take beliefs to be doxastically justified—i.e., permissibly held—only if they are based on a reason:

> **Doxastic Justification by Reasons Entails Basing.** An agent A's belief B is doxastically justified by a reason R[3] only if B is based on R.

This paper aims to show that 'Basing Causalism' is not compatible with 'Doxastic Justification by Reasons Entails Basing'. Some beliefs are doxastically justified, although they stand in no causal relation to any reason. Thus, the causalist faces a hard choice: either to abandon the enormously popular view that doxastic justification has a basing requirement, or to modify her view of the basing relation.

To be sure, over the years, many have presented non-causal theories of the basing relation and doxastic justification. These theories are built, almost exclusively, to accommodate a controversial intuition about Keith Lehrer's 'Superstitious Lawyer' case (see Sect. 2). In this thought experiment, roughly, a lawyer forms a belief about his client's innocence because this is suggested to him by a tarot card reading. He also finds impeccable proof in favor of his client's innocence. Given the details of the case, without the card reading, he would have never found this proof, and even if he had found it, he would have been psychologically unable to trust it. One important reason why non-causal theories have failed to convince causalists is that the lawyer nevertheless exhibits *residual epistemic flaws* (e.g., unreliability, epistemic luck). Theories built to accommodate this intuition—that the lawyer is doxastically justified and that his belief is based on his reasons—are therefore bound to argue that beliefs can be justified *despite* these flaws. Causalists, in contrast, have argued that these flaws are unacceptable. The result is philosophical gridlock with little movement on either side.[4]

I suggest breaking this gridlock by presenting cases in which beliefs are justified without such residual epistemic flaws, although no causal relation between a belief and its justifying reasons is established. I will present my arguments as an extension

---

[1] For example, Alston (1989, p. 108), Bondy and Pritchard (2018, p. 3812), Comesaña and McGrath (2014), Feldman (2002, p. 46), Feldman and Conee (1985), Ichikawa and Steup (2012), Littlejohn (2012), Kvanvig (2003, pp. 43–44), Lycan (2012), Neta (2010), Pollock and Cruz (1999, pp. 35–36), Pryor (2001, p. 104), Silins (2007, p. 109), Turri (2010), and Ye (2020).

[2] Silva (2015) relies on an interesting, but controversial (see Oliveira, 2015), analogy between moral and epistemic justification to oppose this claim. A detailed assessment of his idea would lead us beyond the scope of this paper. Let me, however, emphasize that Silva's proposal is radical in a way that few of us might want to accept. Justification, according to Silva, is closely tied to epistemic *permissibility*. And epistemic permissibility, in turn, should be treated along the lines of moral permissibility. Moral permissibility, however, is simply independent of the reasons that an agent has for acting. Torturing little babies is not morally justified, whatever my reasons may be. My reasons, or the lack thereof, may bear on the question of whether I am exculpated, but this seems to *presuppose* that I have done wrong and that my action is not morally justified. An action's moral permissibility has simply nothing to do with me *having* certain reasons. In this paper, I will side-step his interesting proposal, and accept that justification$_d$ requires being related to one's reasons in the right way.

[3] Foundationalists about justification believe that some justified beliefs do not, for their epistemic status, depend on other propositions. To maintain neutrality on the question of whether there such foundational justified beliefs, I restrict my claims to beliefs that are justified by a reason.

[4] This bifurcation is identified beautifully in Wallbridge (2018).

of Evans' (2013) dispositionalist account of the basing relation. Evans argues that a belief is based on a reason if, and only if, the believer is disposed to revise her belief when she loses the reason. Expanding on this idea, we can ask which role a reason should play when it is *not* lost. While Evans favors a causal sustainer account—i.e., reasons that are not lost sustain beliefs that the agent has—we can describe these cases using the same dispositionalist framework. Thus, I shall argue that beliefs can be doxastically justified when an agent's reasons have a mere *disposition* to sustain these beliefs. Dispositions to sustain, as I shall argue, are not epistemically worse than actual sustaining; such dispositions do not introduce epistemic compromise. In the cases I have in mind, the *manifestation* of this disposition would establish a causal relation, but this manifestation would not add any positive epistemic features to the belief. This train of thought is naturally cast in terms of the language of *control*: If an agent has a disposition to revise a belief when a reason is lost and a disposition to affirm the belief as long as the reason persists, then the belief is under the control of a reason. Beliefs being under the control of reasons is what is essential to doxastic justification.

Let me use the rest of this introduction to set out the central concept of 'doxastic justification' (henceforth 'justification$_d$'), which is routinely contrasted with propositional justification. According to the standard use of the terminology, an agent is propositionally justified in believing some proposition if she *has* adequate reasons to justify a belief, i.e., if the belief is justifi*able* in light of her reasons. The agent, though, might not even have the belief. In contrast, an agent's belief is said to be doxastically justified if, and only if, she has the belief and holds it justifiably, or, as it is often said, permissibly.[5]

Crucially, justification$_d$ (of non-foundational beliefs) requires more than simply having both a belief and the evidence to support it. This can be illustrated by examples such as this one:

> **Juror**. 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. (Turri, 2011, p. 383)

Although the juror has both the belief and the corresponding evidence, her belief does not seem to be justified$_d$. The problem in cases such as 'Juror' is that the evidence is not related to the belief in the right way. The nature of this relation is the topic of the present paper.

The next two sections will serve to set up my argument. In Sect. 2, I will introduce Lehrer's 'Superstitious Lawyer' case in greater detail and explain why accounts that have tried to accommodate this case have so far failed to convince causalists. In Sect. 3, I will set out causal theories of the basing relation (and thereby of justification$_d$). In Sect. 4, I will present my argument that beliefs can be justified in light of a reason's mere disposition to sustain it.

---

[5] See, for instance Goldman (1989, p. 59), Wedgwood (2012, p. 274).

2 Lehrer's 'Superstitious Lawyer'

Over the years, a substantial minority of philosophers have developed a great variety of non-causal accounts of the basing relation. These accounts aspire, almost exclusively, to accommodate Lehrer's (1971) controversial intuition that the lawyer's belief concerning his client's innocence in the following example is based on his evidence:

> **Superstitious Lawyer.** Eight murders have occurred. The lawyer's defendant is proven to be guilty of the first seven of these murders. Initially, the lawyer believes and desires that the defendant is also guilty of the eighth murder (closely mirroring public sentiment on this issue). However, placing overwhelming trust in tarot card readings, he concludes that his client is innocent after the cards suggest to him that his client is innocent. Being surprised by this result, the lawyer re-examines the evidence and finds an impeccable, but complicated, line of reasoning decisively proving the defendant's innocence in the case of this last murder. Importantly, had the lawyer found the evidence without the prior consultation of the cards, his strong desire to declare his defendant guilty, in conjunction with public pressure, would have left him unable to appreciate the evidence correctly. He would have believed that the client is guilty despite being in possession of all the right evidence.

Some have taken this example to suggest that the lawyer's belief, after finding a proof for the client's innocence, is *based on* the evidence, that it is *justified$_d$*, and that the lawyer *knows* that his defendant is innocent,[6] although (a) the evidence is neither part of the belief's causal history (it doesn't cause or causally sustain the belief), and although (b) the evidence is no counterfactual cause of the belief (i.e., had the card reading not occurred, the lawyer would have not trusted the evidence) (e.g., Bondy & Adam Carter, 2020; Kvanvig, 2003; Korcz, 2000; Swain, 1981).

Ever since Lehrer's example entered the philosophical scene, the discussion of the basing relation has been somewhat bifurcated.[7] On the one hand, there are those, a substantive minority, who abandon the causal criterion aiming to accommodate Lehrer's intuition. In 'Superstitious Lawyer', however, the connection between the lawyer's belief and his evidence is quite loose: The evidence neither actually, nor counterfactually, causes his belief. For this reason, philosophers who are intrigued by Lehrer's example have developed weaker conceptions of the basing relation to accommodate their intuition that the lawyer's belief is justified$_d$ and based on the evidence.[8]

---

[6] In the relevant literature, these three claims (about basing, justification$_d$, and knowledge) usually go together, i.e., philosophers either deny or accept these claims wholesale. Lehrer, for instance, frames his paper as a discussion of the basing relation, and then explicitly argues that the lawyer has knowledge and is justified in believing. Bondy and Adam Carter (2020) are an exception in that they affirm that the lawyer's belief is based on the evidence and that the lawyer has justification$_d$, but they remain non-committal about the knowledge claim.

[7] Wallbridge (2018) presents a compelling analysis of this bifurcation.

[8] Swain, for instance, argues that a belief can be justified if it is "quasi-overdetermined" (Swain 1981, Chap. 3). The core of quasi-overdetermination is that a belief can be justified if in the closest possible worlds in which both the evidence and the belief exist, the evidence causes the belief.

On the other hand, some dismiss Lehrer's intuition, emphasizing the fact that the lawyer's belief seems to exhibit a myriad of *prima facie* epistemic defects: The lawyer's belief is not reliable (e.g., Audi, 1983; Turri, 2011; Wallbridge, 2018; Ye, 2020), it is formed on the basis of luck (Bondy & Pritchard, 2018, Footnote 15), and he seems to invite epistemic blame for irrationally placing trust in the tarot cards.[9] As a result, causalists have often reacted to Lehrer's thought experiment with a Gallic shrug, comfortably sticking to a causal theory of the basing relation and, thereby, justification$_d$. Goldman, for instance, simply states, "I find this example unconvincing. To the extent that I clearly imagine that the lawyer fixes his belief solely as a result of the cards, it seems intuitively wrong to say that he *knows* — or has a *justified belief* — that his client is innocent" (Goldman, 1979, p. 22).

I think this bifurcation in the debate is unavoidable as long as we attempt to provide an account of justification$_d$ that aims to accommodate Lehrer's intuition. It is quite simply true that the lawyer got somewhat lucky in holding his belief. After all, the cards could have easily yielded a different result. It is also simply true that the lawyer, placing trust in the cards, seems to lack crucial epistemic virtues, virtues that someone who trusts only in the evidence would have exemplified. Any non-causal theory of the basing relation that aims at accommodating the superstitious lawyer intuition must, therefore, maintain that the lawyer is justified *despite* these flaws. This, in turn, gives causalists reason to reject such theories.

As I have indicated above, the point of the present paper is *not* to argue that non-causal theories that aim at accommodating the superstitious lawyer intuition are wrong. Instead, I simply wish to point out that it is unsurprising that these theories have failed to convince committed causalists. If our aim is to convince causalists that justification$_d$ does not require causation, we could try to confront them with cases that do not involve such epistemic compromise. This is the strategy I adopt in this paper. Before I present this account, however, let me introduce the causal theory of the basing relation and justification$_d$ in some reasonable detail.

3 The standard causal account of the basing relation

According to causal accounts of the basing relation, not every way in which a reason can cause a belief qualifies as a case of belief basing. So-called "deviant" causal chains do not instantiate the basing relation (see Evans, 2013, p. 2950f. for a nice summary). Plantinga (1993, p. 69) provides a classic example of this phenomenon of deviant causation:

> [S]eeing Sylvia, I form the belief that I see her; as a result, I become rattled and drop my cup of tea, scalding my leg. I then form the belief that my leg hurts; but though the former belief is a (part) cause of the latter, it is not the case that I accept the latter on the evidential basis of the former.

Intuitively, my belief that I see Sylvia is not the reason for which I believe that my leg hurts. It is a cause of my belief but not the basis for it. To avoid the problem of deviant

---

[9] I have not seen this last point in print, however.

causation, causalists argue that a belief isn't based on a reason unless it is caused through a narrowly circumscribed, usually taken to be a cognitive, mechanism. Turri's (2011) theory of the basing relation exemplifies this kind of view:

> **Causal-Manifestation Account.** 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.

This formulation amalgamates several conditions. First, it is presupposed that a person *has* a reason for believing Q. Second, the account identifies a disposition, cast in terms of the workings of cognitive traits. Call this the 'disposition condition'. This condition is designed to avoid the deviance problem. In the example stated just above, for instance, the belief that I see Sylvia caused the belief that my leg hurts, but this latter belief is not caused by engaging a cognitive trait. In this paper, I shall have nothing more to say about the deviance problem, and I accept that causal mechanisms relevant for the characterization of the basing relation must be constrained to avoid causal deviance.

Third, and most importantly for this paper, Turri's account[10] specially asks for the *manifestation* of these traits. Call this the 'manifestation condition', which is supported by a powerful argument: Reasons *for which* someone holds a belief can be cited when explaining *why* this person believes as she does. This argument has most recently been stressed by Neta (2019, p. 181), but the idea can be traced at last to Davidson's (1963) discussion of reasons for action. Reasons for which we *act* can be cited in answering the question of why we act. Similarly, reasons for which we *believe* are reasons that can be cited in answering the question of why we believe. If, say, my belief that Trump lost the election is based on the CNN report that I watched earlier today, then my watching the CNN report may be cited as an answer to the question of why I believe this. Reasons why are explanatory; they explain why a person believes as she does. Reasons that are not part of a belief's actual causal history—as causal creators or sustainers—turn out not to be explanatory. They cannot be cited in answering the crucial "why?" question, even if these reasons exhibit a host of other epistemic virtues. Let me illustrate this with the help of an example from Neta (2019, p. 192):

> **Toshiro** hears the CNN report that Russian forces have bombed civilian targets in Syria. He also represents […] CNN as a trustworthy news source in cases such as this, and so he believes that the CNN report provides him with justification for believing that Russian forces have just bombed civilian targets in Syria. And furthermore, after watching the CNN report, he believes that Russian forces have bombed civilian targets in Syria. Unbeknownst to Toshiro, however, the latter belief about Russian forces was formed just a few milliseconds before he had finished parsing the sentences of the CNN broadcast. It was formed as the result of subliminal messages coming from his TV screen—messages the content of which would in fact justify his belief that Russian forces attacked civilian targets in Syria. Furthermore, the CNN broadcast itself cannot help to explain Toshiro's belief: the subliminal messages coming from Toshiro's TV screen cause Toshiro not only to believe that Russian forces have bombed civilian targets in Syria but

---

[10] See Armstrong (1973) and Sosa (2015) for similar accounts.

also to suffer from a temporary but severe aphasia that prevents him from putting together his belief about the trustworthiness of CNN with his belief about the content of its broadcast to deduce any consequences from their conjunction.

In this stylized vignette, Toshiro *has* the evidence, he *knows* that he has the evidence, and he knows that the evidence supports his belief. His evidence, however, neither causes nor causally sustains his belief. Intuitively, then, when we ask what explains Toshiro's belief, we would cite the subliminal messages he received, but not his evidence (the CNN report). In line with this intuition and in support of the causal account of the basing relation, Toshiro's belief does not seem to be based on the CNN report; instead, it seems to be based on subliminal messages.

Turri's rationale for including a manifestation condition is essentially the same. Reasons *for which* a person believes, he writes, are "difference makers." Only causes, however, make a difference. Turri (2011, p. 387f.) illustrates this point with the help of the following 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. But had Schilling not pitched, Pedro would have pitched and won. So Pedro pseudo-overdetermines the Sox victory, but he made no difference.

Pedro Martinez does not causally contribute to the Red Sox's winning that day while sitting on the bench, which is why we wouldn't cite the fact that Pedro Martinez sat on the bench as part of an explanation of why they won. Instead, we would cite Curt Schilling's impeccable performance. Martinez's performance would be explanatorily relevant for the victory only if he actually intervened (e.g., by pitching). We can cement this train of thought—from basing to explanation, and from explanation to causation—in the following argument:

The Argument from Explanation:
P1. **Basing Entails Explanation.** An agent A's belief B is based on a reason R only if R explains (at least in part)[11] why A holds B.
P2. Explanation Entails Causation. An agent A's belief B is explained by a reason R only if R is part of B's causal history.[12]

C.: Basing Causalism. An agent A's belief B is based on a reason R only if R is part of B's causal history.

---

[11] Partial explanation is to be understood widely, in particular, to encompass causal sustaining and overdetermination (see e.g., Sartorio 2013).

[12] The idea that explanation requires causation in the context at hand should not be taken to mean that all explanation is causal explanation. Most philosophers recognize non-causal explanations such as metaphysical grounding (e.g., a disjunction is true because at least one of its disjuncts is true), and deductive explanations (e.g., Tweety the sparrow is grey because all sparrows are grey). In the present context, however, in which we're looking for ways in which reasons might explain beliefs, these alternative forms of explanations are not serious contenders.

The 'Argument from Explanation', I submit, has strong intuitive appeal[13] and it explains why so many have been drawn to some version of 'Basing Causalism'.

I have presented this argument in some detail to advance my first substantive point towards the conclusion of this paper, i.e., that justification$_d$ does not require causation. Notice that the 'Argument from Explanation' has strong *intuitive* appeal when applied to the basing relation: Reasons *for which* a person holds a belief explain why she holds this belief. Explanation, in turn, is a causal relation. The same argument, it turns out, is less intuitive when applied to justification$_d$. The idea that a person's belief is justified$_d$ only if this reason explains why she holds this belief is simply not as strong a pretheoretical constraint. After all, justification$_d$ is a technical notion.

Furthermore, everyday talk of justification (without the subscript "d") is often not connected to explanation or causation at all, for instance when we talk about someone's actions being (un)justified. I may tell my friend that it was not morally justified to steal his girlfriend's jewelry, even if I recognize that he did it exclusively for honorable reasons (e.g., to pawn the jewelry and donate the money to Oxfam).[14] In this case, we speak of justification as a property of an action independent of the reasons a person has for performing this action. The upshot, then, is this: while there is an intuitive argument in favor of causal accounts of the basing relation, this argument has less intuitive appeal when applied to justification$_d$.

To be sure, it is an open question why this argument has less intuitive appeal when applied to justification$_d$. It might be, for instance, that our everyday concept tracks what philosophers call 'propositional justification' (as opposed to doxastic justification). Alternatively, everyday uses of 'justification' may track the various philosophical concepts in different contexts. I wish to remain neutral on this issue. My moderate point is just that we cannot appeal to pretheoretical linguistic intuitions about, say, the use of words such as "justification" and "is justified" to motivate the view that justification$_d$ requires an attitude's causal connection to reasons.

For this reason, even those who are already committed to a causal account of the basing relation, should at least reconsider whether justification$_d$ is likewise governed by such a causal constraint. In the next section, I will continue to present arguments with the aim of showing that justification$_d$ does not require causal connections to reasons.

4 Doxastic justification through dispositions to cause

In this section, I shall argue, in a way that should be acceptable to causalists, that justification$_d$ does not require the exercise or the manifestation of a non-deviant causal mechanism (e.g., a cognitive trait). Beliefs can be justified when such a mechanism is

---

[13] Many causalists have explicitly embraced it and those who haven't embraced it have either simply denied it (Bondy & Adam Carter, 2020), or they have not addressed it (e.g., Lehrer 1971; Leite 2004). As stated above, I don't wish to convince the committed non-causalist. Notably, Lehrer (1971, p. 312) aims to provide an answer to the question "how" someone knows a proposition. He says that "[r]easoning gives a man knowledge if and only if it is a correct answer to such a question." He doesn't, however, continue to discuss this issue further.

[14] This has been pointed out by Silva (2015).

exercis*able* or manifest*able* in the right sort of way. This idea can be developed as an extension of Evans' (2013) account of the basing relation. He argues for the following thesis:

> **DT.** *S*'s belief that *p* is based on *m* iff *S* is disposed to revise her belief that *p* when she loses *m*.

Evans' account shares crucial similarities with standard causal accounts of the basing relation set out in Sect. 3. Belief *revision* refers to a cognitive process. This provision is meant to rule out causally deviant belief change (see Sect. 3). A tumor whose removal would result in a belief change does not qualify as the basis of a belief, because this change would not count as a belief *revision* (Evans, 2013, p. 2953). Beliefs resulting from a tumor removal are not based on anything. Next, Evans casts the basing relation in terms of a *disposition* to revise a belief (as opposed to its counterfactual exercise) to account for cases of overdetermination. Imagine, for instance, that you have 10 reasons, each sufficient for *p*. Losing one of these reasons should not lead to belief revision which is why the counterfactual "if you were to lose a reason, you would revise your belief" is false, and yet, the belief was partly based on that reason. Thus, although I might not, in fact, revise a rationally overdetermined belief after losing one of the reasons that support it, I may still have a *disposition* to revise it; a disposition whose manifestation is temporarily masked by the presence of my remaining nine reasons in support of *p*.

Now, Evans' account contains explicit provisions only for the case in which a reason is *lost*. It contains no provision for the case in which a reason is *retained*. Intuitively, we also want to know what role a reason must (or may) play if it is *not* lost. Evans (2013, p. 2955) considers, and tentatively affirms, the idea that his account amounts to a form of causal sustaining:

> I am aware that some readers will find it hard to conceive of someone's being disposed to revise her belief when she loses *m* without *m*'s somehow causally sustaining her belief. If these readers are correct in thinking that dispositions to revise entail causal sustaining, then my dispositional theory is ultimately a sort of causal theory. But that is fine with me.

Why is it hard to conceive of a disposition to revise a belief that is not preceded by causal sustaining? Intuitively, the fact that a reason used to causally sustain a belief elegantly explains why someone would be disposed to revise it after losing a reason that supports it. Analogously, the fact that my dining room table was sustained by four wooden legs is part of an elegant explanation for why the table collapsed after removing one of its legs. If no alternative explanation can be found, then Evans' theory does amount to a causal account, at least in extension. On reflection, however, we can conceive of an alternative explanation in terms of *dispositions* to sustain. Imagine the following case:

> **Judy** is a juror in a trial. At the beginning of the trial, when Judy first sees the defendant, she immediately comes to believe that they must be guilty because of their strange-looking mustache. Throughout the trial, she pays close attention to all the evidence presented which decisively proves that the defendant is guilty.

Upon hearing the evidence, Judy acquires a disposition to revise her belief if she were to lose some of the evidence (e.g., if she learned that the alleged murder weapon was a toy weapon). *She also acquires a disposition to affirm her belief if she were to lose her belief's current sustainer but nevertheless kept the evidence (e.g., if the defendant shaved off their mustache, but the murder weapon did not turn out to be a toy gun).*

When Judy acquires the evidence, she acquires a disposition to affirm the belief when she loses her current sustainer.[15] Such a disposition works just like a disposition to revise with the difference that it leads to belief maintenance, as opposed to belief change. While belief *revision* is a non-deviant form of mere change, belief *affirmation* is a non-deviant form of mere sustaining.

We should note right away that it is at least conceptually possible for Judy to acquire such a disposition to affirm her belief which remains causally *in*active as long as the original cause persists. Dispositions to maintain are not *thereby* causally active. The following example illustrates this:

Thermostat. Currently, my office temperature is controlled by the new thermostat that I just installed. The thermostat is monitoring the room temperature. Were the temperature to drop, the thermostat would send signals to the central heating unit, thereby engaging it, causing the room to heat. Ever since I installed it a week ago, the thermostat has been inactive because, at this time of year, the temperature is just lovely and no heating is required.

My thermostat has a disposition to sustain the room temperature, but it does not enter the causal scene as long as it has not regulated the room's temperature. The central intuition behind cases such as 'Judy', and indirectly behind 'Thermostat', is that a belief *can be under the control* of a reason, implemented through dispositions to revise and maintain, although no causal connection between both (reason and belief) has been established. Such a connection would not be established until such control is *exercised*.[16] Being in control, however, does not *amount to* its exercise. Beliefs can be justified$_d$, as I will argue in more detail below, when they are controlled by reasons in this way. This account can be summarized as follows:

Doxastic Justification through Dispositions to Cause[17]

---

[15] To be clear, this is a wide scope disposition to 'affirm the belief when she loses her current sustainer'. It is not a narrow scope disposition to 'affirm the belief', when she loses her current sustainer.

[16] The exercise of control seems to be a straight-forwardly causal notion (see Shepherd 2014, p. 396).

[17] It is worth mentioning that this account, although stated, for ease of presentation, in terms of dispositions to cause, could likewise be stated in terms of counterfactuals. First, some philosophers take "dispositions to be reducible to counterfactuals" (Bondy, 2016, p. 550). Evans himself, notwithstanding serious reservations, is open to such a reduction (see Evans 2013, p. 2955). If such a reduction succeeds, then my dispositional account is just a counterfactual account in disguise. Let me here just give a brief sketch of such a restatement in terms of counterfactuals. Consider clause b. Instead of referring to a disposition to cause, we might require that R would non-deviantly cause B if B were to lose its current sustainer. Such an account no doubt needs further fine-tuning. For instance, we would want to allow for cases in which if B's present sustainer were lost, R would not cause B because some cause other than R would cause it (see e.g., Lemke 1986 for interesting suggestions along this line).

In some cases, S's belief B is doxastically justified in light of reason R, although R is not part of B's causal history. In these cases,

a. Disposition to Revise. S is disposed to revise B when she loses R.
b. Disposition to Affirm. S is disposed to affirm B as long as she retains R.

Now, why should we take this possibility seriously? A critic may object: "I can grant that devices such as thermostats can be in control without exercising any; but surely, reasons cannot behave this way. If Judy appreciates the evidence while also holding the belief that this evidence would justify, then her reasons must become causally active; they must causally sustain the belief. This is how brains work!" The supposed critic recognizes that it is *conceptually possible* for controlling reasons to not manifest their causal powers to maintain a belief but argues that it is *cognitively impossible.* I'd like to give two replies.

First, it is an empirical question whether beliefs are always causally active in sustaining pre-existing, possibly baseless, beliefs. This is simply not to be decided a priori.

Second, a host of empirical research suggests that this way of modeling the connection between beliefs and evidence might sometimes be correct and that the actual causal connection implemented in the brain can be very unintuitive and surprising. Recently popular predictive coding models of cognition heavily emphasize the importance of predictive processes in the way we interact with the environment (Clark, 2013; Friston & Kiebel, 2009; Hohwy, 2013; Hohwy et al., 2008; Rao & Ballard, 1999; Spratling, 2016; Westra, 2019). Crucially, on this approach, many of our cognitive representations (e.g., perceptual, proprioceptive) are not constructed from bottom-up signals coming from the environment; instead, they are constructed from statistically informed top-down expectations that try to predict what those bottom-up environmental signals will be. These top-down predictions are checked against incoming sensory signals and are modified only if an error between both (prediction and sensory input) is detected. Perceptual representations are causally constructed on the basis of top-down predictive processes. Visual stimuli, in turn, become causally relevant only if this top-down construction process turns out to be erroneous. In such cases, bottom-up visual information is relevant to "cancel" (see Friston, 2002, p. 247) and adjust the initial prediction aiming to minimize the error between both (bottom up and top down) signals.

One intuitively plausible interpretation of this approach to cognition is that visual stimuli are dispositionally relevant to visual processing. They provide "feedback [that] supervises" (Hohwy, 2007, p. 323) perceptual content. The construction of perceptual content, however, proceeds in a top-down fashion, without the causal help of visual signals. Although in the case of error-free predictions, visual information is not causally relevant to the construction of a visual representation, we should not rule out the possibility that such information is relevant for the justificatory status of these representations.

The point of this discussion is to show that the causal pathways implemented in our brain are often hard to predict on a priori grounds. It certainly *seems to us* that the

content of our perception is the result of bottom-up processes that use visual information coming from the world to construct a representation of the world. However, if predictive coding theory is correct, this assessment may just be incomplete. Although predictive coding theory proposes a novel understanding of the way representations are caused (i.e., top-down rather than bottom-up), it seems at least initially plausible to say that the visual information that serves to check and correct those representations is crucial for their justification$_d$.

The imagined critic may try to play down the relevance of predicting coding theory by pointing out that the basing relation is a relation that holds between an agent's *beliefs* and her reasons, not between her visual stimuli and perceptual representations. This line of criticism, however, can be resisted. First, although discussions of the basing relation and justification$_d$ are traditionally centered around beliefs, we may ultimately want a single theory for "all rationally determinable conditions" (Neta, 2019, p. 182) such as "a belief, a judgment, an emotion, an intention, a preference, a choice, or an action, or perception" (Neta, 2019, p. 182). Second, it is not uncommon to classify perceptual representations as beliefs (see Lyons, 2017 for a summary). Third, predictive coding theory, although originally developed in computer science as a theory of data compression (see Clark, 2013, p. 182), and most prominently applied to perception, is perceived by its proponents as a general theory of cognition (see Clark, 2013) and has been applied to a wide variety of cognitive processes including belief (e.g., Hohwy, 2013), memory (e.g., Barron et al., 2020), and social cognition (e.g., Westra, 2019).

Given that these arguments are roughly correct, we can ask whether beliefs that are causally sustained by a reason have positive epistemic value that merely dispositionally sustained beliefs lack. My answer to this question is "no". Let me consider two possible arguments in opposition of this conclusion.

In a recent paper, Bondy and Pritchard (2018) ask why beliefs that are based on evidence are epistemically better than merely propositionally justified beliefs. If a person "has a true belief, and she has good evidence for it, and there's no evidence against it, why [should] it matter if she doesn't believe on the basis of the good available evidence?" (Bondy & Pritchard, 2018, p. 3812). The fact that a belief is based on a reason, they argue in answering this question, "rules out propositional epistemic luck" (Bondy & Pritchard, 2018, p. 3812). This is the luck of having a good reason to support one's belief. "In a wide range of cases, where S believes for bad reasons but has good reasons available, S is just lucky to have good reasons available. In these cases, it could easily have been the case that S would have held the same belief in the same way but lacked good reasons for it" (Bondy & Pritchard, 2018, p. 3814). Suppose, to illustrate this using their central example, that Helen has good evidence that there is ice cream in the freezer (e.g., her trustworthy roommate told her so). Helen, however, holds her belief on the basis of an irrational fear to live in a world where her freezer contains no ice cream. If her fear is robust, then so is her belief, and her having this belief is not lucky. However, Helen was lucky that she had evidence at her disposal. After all, had her roommate not told her about the ice cream, she would still have had her belief that there is ice cream, but she would have lacked evidence to support it. Thus, she got propositionally lucky in that she could have easily failed to have the evidence.

Are beliefs that are merely controlled by a reason through a disposition to sustain (or affirm) them propositionally luckier than beliefs that are causally sustained by such a reason? This is the question of whether in close worlds, in which a person still holds the belief, she would have lacked the evidence on the control view, but not on the causal sustainer view. To evaluate this question, we should distinguish two scenarios: First, we should consider scenarios in which Helen initially has evidence, along with either a suitable disposition to cause her belief or with the evidence causally sustaining her belief, but then, ultimately, lacks it because she *lost* it. We should note that Evans' 'disposition to revise' clause, which I have adopted, has it that those who lose evidence have a disposition to revise their belief. This disposition is construed as manifesting unless there are additional reasons masking its manifestation, which is not true of Helen. Therefore, even the control view entails that those who lose evidence don't retain their belief, which is why they are not propositionally lucky. In this case, mere dispositionalist views and causal sustainer views make the same prediction.

Second, we can consider close worlds in which a person holds a belief but never receives the evidence in the first place. Consider again Helen, who has an irrational fear of being out of ice cream, which in turn causes her to believe that there is ice cream in the freezer. She also receives decisive evidence that there is ice cream in the freezer. For our purposes, we can imagine two roles that Helen's evidence might play: first, we can imagine that Helen acquires a disposition to sustain her belief through her evidence; second, we can imagine that her evidence becomes a causal sustainer, thereby causally overdetermining her belief. Both variations entail that Helen still has the belief in close worlds where she never receives the evidence. After all, if she never receives the evidence, her irrational fear would still cause her belief. Therefore, causal sustaining would not improve Helen's epistemic situation when compared to a mere disposition to sustain her belief.

At this point, the friend of the causal theory might wish to exclude causal overdetermination: if Helen's fear is sustained by her evidence *and* by her irrational fear, then she cannot be justified. A belief, the defender of the causal view may continue, cannot be justified unless her belief is caused *only* by reasons. This line of argument, however, won't work. Although reformulating the causal account in this way would introduce a relevant difference, it is not a reformulation that the causalist should pursue. After all, any belief sustained by reasons is *also* sustained by a myriad of other causes. This is the problem "overabundant sustainers" (Evans, 2013, p. 2948):

> "The continued existence of my beliefs depends, in some sense, on a host of psychological and environmental conditions. On any plausible account of causal sustaining, these things will count among the causes of the continued existence of my beliefs. But then there is trouble for sustaining accounts of causation: most of the things that sustain my beliefs are not among the bases of those beliefs. Call this the problem of overabundant sustainers."

Causalists must therefore admit that beliefs that are based on reasons are also causally sustained by many other factors.

Let's move on to the second argument. Virtue epistemologists such as Riggs or Greco might argue that an agent cannot be *credited*, or *praised*, for holding a belief unless her reasons causally bear on her belief. Thus, an agent might become eligible

for epistemic praise when a belief of hers is sustained by her reasons, but she might be thought not to be eligible for praise when her reasons have a mere disposition to sustain her belief. Consider an analogy which will help evaluate this claim:

> **Noemi** was recently appointed the head of a local co-op. Her goal is to manage the co-op responsibly (e.g., secure its commercial success, guarantee a healthy work environment). For quite a long time now, the co-op has been running like a clockwork: the employees are well-instructed, the supply chains are in place, everybody is enthusiastic and happy, revenue is burgeoning at an exponential rate, etc. For these reasons, Noemi hasn't actually done anything with regard to the co-op. Instead, she has been writing papers on the Chinese Room thought experiment. If, however, her intervention was required, she would competently and swiftly intervene to ensure the co-op's continued success.

The intuition is that Noemi may not receive credit or praise as long as she has not contributed to the store's success. Similarly, we might want to say that an agent does not deserve epistemic praise for holding a belief as long as her reasons haven't causally contributed to her belief. Intuitively speaking, I think this is false. First, it seems that our lack of praise for Noemi is a result of our epistemic limitations. We wouldn't praise Noemi because we don't know yet whether she is a good manager. However, if these limitations were removed, if we did know all the relevant facts about her, it would be appropriate to praise her. After all, we would be able to judge that she's doing exactly what she is supposed to do (right now, she's not supposed to do anything), and that she has the appropriate dispositions to intervene. Furthermore, even if it were true that Noemi couldn't claim praise unless she intervened in the store's affairs, our question in the context of belief is not whether merely controlled beliefs lack epistemic value. Instead, the question is whether establishing a causal connection would *add* such value. In cases such as 'Noemi', needlessly doing things would not make Noemi praiseworthy. A wise manager acts if necessary, she doesn't necessarily act. In fact, needlessly exerting causal force just for the sake of it would invite *criticism*. After all, her efforts would be expended needlessly.

The point of this example is to show that there are plausible scenarios in which an agent, such as Noemi, deserves praise although she has not causally contributed to the relevant positive upshot. As long as Noemi has an appropriate disposition to intervene, causal contribution is not necessary for such praise. A critic may object that there are certain elaborations of this case in which Noemi does have a disposition to intervene but in which she is nevertheless not praiseworthy. Imagine for instance that Noemi is in a drug-induced coma. If problems in the co-op were to arise, someone would wake her, and she would swiftly solve the problem at hand. Intuitively, in this elaboration of the case, Noemi is not praiseworthy for her inaction while being in a coma.[18] I can think of two ways to address this ambiguity (i.e., between cases in which Noemi is and cases in which she isn't praiseworthy). First, we might deny that comatose Noemi has a disposition to intervene. Instead, upon being woken, she *acquires* such a disposition. Such dispositions, that are acquired when they are tested for, are known as reverse finkish dispositions (see e.g., Martin, 1994). Alternatively, it

---

[18] I owe this objection to an anonymous reviewer.

might be argued that a *mere* disposition to intervene is not sufficient to be eligible for praise. Instead, Noemi must also monitor or supervise the goings-on of the co-op. This idea was mentioned above in the discussion of predictive coding. Although notions such as "supervision" or "monitoring" are in turn semi-technical and ultimately in need of further explication, the present context does not mandate further elaboration. All that needs to be agreed on is that a disposition to intervene, implemented through supervision, does not thereby establish a causal connection. And this much seems true. A complete account of justification$_d$ may require further examination of these semi-technical notions; but remember, my argumentative goal was to show that causal connections are not necessary for justification$_d$, my goal was not to provide a complete account of the kinds of disposition that are sufficient for justification$_d$.

Reasons that justify$_d$ a belief in the way I have described—i.e., by way of controlling the belief through dispositions to cause—don't cause this belief. What is more, such reasons neither seem to determine the existence of the belief in some other, non-causal, way. For instance, such reasons do not *ground* or *constitute* the existence of the belief. One might therefore wonder what, if anything, the presence of these reasons does explain. Surely, these reasons should make *some* difference. My answer is twofold. First, although such reasons do not cause the belief, they are nevertheless relevant to explain its epistemic status. Beliefs are justified$_d$ at least in part *in virtue* of the fact that they are controlled by reasons. In particular, if the arguments presented just above are correct, then such dispositions ground a variety of good-making epistemic features: they remove epistemic luck, and they can explain why an agent is epistemically praiseworthy for believing. Second, although dispositionally relevant reasons might not be causes, they might, in broader sense, still be part of a causal explanation. Let me explain. Famously, Davidson (1967) argued that events, not facts, are causes but that facts can play a role in causal explanations.

In a similar vein, Lewis argued that "to explain an event is to provide some *information* about its causal history" (my italics) (1986, p. 217), but giving information may not involve picking out a cause. For instance, one may hold that only events are causes, but nevertheless explain an event by "saying that such and such a kind occurred, rather than that some particular event occurred" (Beebee, 2004, p. 302). In principle, on Beebee's (2004) interpretation of Lewis' view, any piece of information that teaches you something about an event's causal history may become part of a causal explanation. For instance, certain absences can be explanatorily relevant for a causal explanation in that they can be provide insights in what did *not* cause some event. 'Flora's *not* watering the plants' may be relevant to explain why the plants died because such an explanation teaches us that 'Flora's watering the plants' cannot be found among the causes of the plants' death and that Flora's watering the plants might have prevented the plants from dying. Hence, in learning about this absence, we learn something interesting about the causal process that led to the plants' demise. Similarly, although dispositionally relevant reasons are not causes of a belief, they can still provide interesting information about the causal processes involved. For instance, the presence of such a disposition teaches us that reasons would have been among the causes in nearby worlds in which the actual causes are absent. What is more, the presence of such reasons explains why the existence of the belief is *robustly* dependent on the agent's reasons.

5 Conclusion

In this paper, I have argued that beliefs can be doxastically justified in virtue of the fact that a reason has a disposition to non-deviantly sustain the belief, i.e., when a reason controls a belief. In particular, in some cases, a belief that is justified by a reason's disposition to sustain it does not lack positive epistemic features that would be added if this disposition were manifested. For this reason, I conclude that beliefs that are justified in the latter case are also justified in the former case. Beyond arguing for the conceptual possibility that reasons can have a disposition to sustain an independently existing belief without thereby manifesting this disposition, I have pointed to predictive coding theory to suggest that this possibility is popular among many empirical researchers. These arguments challenge those who believe that the basing relation is a causal relation either to relinquish the view that doxastic justification has a basing requirement, or to abandon the causal view of the basing relation.

Declarations

References

Alston, W. (1989). *Epistemic justification*. Cornell University Press.

Armstrong, D. M. (1973). *Belief, truth and knowledge*. CUP Archive.

Audi, R. (1983). The causal structure of indirect justification. *The Journal of Philosophy, 80*(7), 398–415.

Barron, H. C., Auksztulewicz, R., & Friston, K. (2020). Prediction and memory: A predictive coding account. *Progress in Neurobiology, 192*, 101821.

Beebee, H. (2004). Causing and nothingness. In L. A. Paul, E. J. Hall, & J. Collins (Eds.), *Causation and counterfactuals* (pp. 291–308). MIT Press.

Bondy, P. (2016). Counterfactuals and epistemic basing relations. *Pacific Philosophical Quarterly, 97*(4), 542–569.

Bondy, P., & Adam Carter, J. (2020). The superstitious lawyer's inference. In P. Bondy & J. Adam Carter (Eds.), *Well-founded belief: New essays on the epistemic basing relation*. Routledge.

Bondy, P., & Pritchard, D. (2018). Propositional epistemic luck, epistemic risk, and epistemic justification. *Synthese, 195*, 3811–3820.

Clark, A. (2013). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Comesaña, J., & McGrath, M. (2014). Having false reasons. In *Epistemic norms: New essays on action, belief, and assertion*. Oxford University Press.

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy, 60*(23), 685–700.

Davidson, D. (1967). Causal relations. *Journal of Philosophy, 64*, 691–703. Reprinted in Davidson (1980): 149–162.

Evans, I. (2013). The problem of the basing relation. *Synthese, 190*(14), 2943–2957.

Feldman, R. (2002). *Epistemological duties*. Oxford University Press.

Feldman, R., & Conee, E. (1985). Evidentialism. *Philosophical Studies, 48*(1), 15–34.

Friston, K. (2002). Beyond phrenology: What can neuroimaging tell us about distributed circuitry? *Annual Review of Neuroscience, 25*, 221–250.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1521), 1211–1221.

Goldman, A. I. (1979). What is justified belief? In *Justification and knowledge* (pp. 1–23). Springer.

Goldman, A. (1989). *Epistemology and cognition*. Harvard University Press.

Harman, G. H. (1970). Knowledge, reasons, and causes. *The Journal of Philosophy, 67*(21), 841–855.

Hohwy, J. (2007). Functional integration and the mind. *Synthese, 159*, 315–328.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition, 108*(3), 687–701.

Ichikawa, J., & Steup, M. (2012). The analysis of knowledge. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (winter 2012 edition). Stanford University.

Korcz, K. A. (1996). *The epistemic basing relation*. Dissertation, The Ohio State University.

Korcz, K. (2000). The causal-doxastic theory of the basing relation. *Canadian Journal of Philosophy, 30*(4), 525–550.

Kvanvig, J. L. (2003). Justification and proper basing. In E. Olsson (Ed.), *The epistemology of Keith Lehrer* (pp. 43–62). Kluwer Publishing Co.

Lehrer, K. (1971). How reasons give us knowledge, or the case of the gypsy lawyer. *The Journal of Philosophy, 68*(10), 311–313.

Leite, A. (2004). Believing one's reasons are good. *Synthese, 161*(3), 419–441.

Lemke, L. (1986). Kvanvig and Swain on the basing relation. *Analysis, 46*(3), 138–144.

Lewis, D. (1986). *Philosophical papers* (Vol. II). Oxford University Press.

Littlejohn, C. (2012). *Justification and the truth-connection*. Cambridge University Press.

Lycan, W. (2012). Explanationist rebuttals (coherentism defended again). *Southern Journal of Philosophy, 50*(1), 5–20.

Lyons, J. (2017). Epistemological problems of perception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. https://plato.stanford.edu/archives/spr2017/entries/perception-episprob/

Martin, C. B. (1994). Dispositions and conditionals. *The Philosophical Quarterly, 44*, 1–8.

McCain, K. (2016). *The Nature of Scientific Knowledge: An explanatory approach*. Springer Undergraduate Texts in Philosophy. Springer.

Moser, P. K. (1989). *Knowledge and evidence*. Cambridge University Press.

Neta, R. (2010). Liberalism and conservatism in the epistemology of perceptual belief. *Australasian Journal of Philosophy, 88*(4), 685–705.

Neta, R. (2019). The basing relation. *Philosophical Review, 128*(2), 179–217.

Oliveira, L. R. (2015). Non-agential permissibility in epistemology. *Australasian Journal of Philosophy, 93*(2), 389–394.

Plantinga, A. (1993). *Warrant and proper function*. Oxford University Press.

Pollock, J., & Cruz, J. (1999). *Contemporary theories of knowledge* (2nd ed.). Rowman and Littlefield.

Pryor, J. (2001). Highlights of recent epistemology. *The British Journal for the Philosophy of Science, 52*(1), 95–124.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*, 79–87.

Sartorio, C. (2013). Making a difference in a deterministic world. *Philosophical Review, 122*(2), 189–214.

Shepherd, J. (2014). The contours of control. *Philosophical Studies, 170*, 395–411.

Silins, N. (2007). Basic justification and the Moorean response to the skeptic. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (Vol. 2, pp. 108–142). Oxford University Press.

Silva, P., Jr. (2015). Does doxastic justification have a basing requirement? *Australasian Journal of Philosophy, 93*(2), 371–387.

Sosa, E. (2015). *Judgment and agency*. Oxford University Press.

Spratling, M. W. (2016). Predictive coding as a model of cognition. *Cognitive Processing, 17*(3), 279–305.

Swain, M. (1981). *Reasons and knowledge*. Cornell University Press.

Turri, J. (2010). On the relationship between propositional and doxastic justification. *Philosophy and Phenomenological Research, 80*(2), 312–326.

Turri, J. (2011). Believing for a reason. *Erkenntnis, 74*(3), 383–397.

Wallbridge, K. (2018). The peculiar case of Lehrer's lawyer. *Synthese, 195*(4), 1615–1630.

Wedgwood, R. (2012). Justified inference. *Synthese, 189*(2), 273–295.

Westra, E. (2019). Stereotypes, theory of mind, and the action–prediction hierarchy. *Synthese, 196*(7), 2821–2846.

Ye, R. (2020). A doxastic-causal theory of epistemic basing. In P. Bondy & J. Adam Carter (Eds.), *Well-founded belief: New essays on the epistemic basing relation*. Routledge.