THE DUALIST

Stanford's Undergraduate Journal of Philosophy

This issue is dedicated to

The Kilimanjaro Darkjazz Ensemble.

THE DUALIST

Volume XVIII ▪ Spring 2013

Department of Philosophy

Stanford University

TABLE OF CONTENTS

Can External Reasons Explain Actions?

Brian Kogelmann
University of Illinois at Chicago

Abstract:
Bernard Williams claims that for a reason to actually be a reason, it must be capable of explaining an agent's actions. He also claims that what are ostensibly external reasons are not capable of performing this explanatory function, and thus are not really reasons at all. This paper endeavors to further delve into the question of whether external reasons can explain actions by attempting to understand what it would look like for an external reason to do so. In going about this task the paper proposes a new manner through which external reasons can explain actions. Although this proposed explanatory mechanism deviates from the traditional way of understanding reason-giving explanation, it is a conception of explanation that I argue is satisfactory.

1. The purpose of this paper is to answer the question posed in the title. This question – whether an agent's supposed external reason to φ can explain the agent actually φ-ing – is of importance for the current internal-external reasons discourse. This is because Bernard Williams' seminal paper on internal and external reasons[1] challenges the external reasons theorist precisely on these grounds: that for something to be considered a reason it must be capable of explaining an agent's actions, and, Williams argues, external reasons cannot

---

[1] Bernard Williams, "Internal and External Reasons."

perform this necessary function. But is Williams correct in his assessment? Possibly, though more might be said in defense of the external reasons theorist. This paper explores the relationship between external reasons and explanation. The thesis of this paper is as follows: there is an intelligible way that we can conceive of an external reason explaining an agent's actions, which involves the introduction of a new explanatory mechanism that deviates from the traditional way of understanding reason-giving explanation. In the next section I get at the heart of the problem concerning external reasons and explanation by examining Williams' two arguments against external reasons, and what I take the connection between these two arguments to be. While my interpretation of the two arguments Williams presents is a standard one within the secondary literature, the connection that I draw between these two arguments is not as common. Hopefully, the reading I offer of Williams, where I highlight this connection between the two arguments, is a convincing one. At the end of the next section it will become clear what is at issue in regards to external reasons and explanation. I will also offer an adumbration of how the paper will proceed.

2. Statements of the form "A has a reason to φ" and "there is a reason for A to φ" are subject to two sorts of interpretation: the internal interpretation and the external interpretation. On the internal interpretation, the sentence implies that "A has some motive which will be served or furthered by φ-ing."[2] As an example of this, if I have a motive to eat something sweet, the sentence "Brian has a reason to eat a piece of chocolate cake" will be true under the internal interpretation because my motive – to eat something sweet – will be served by my eating chocolate cake. Conversely, if eating a piece of chocolate cake served no motive of mine then the sentence "Brian has a reason to eat a piece of chocolate cake" would be false under the internal interpretation. On the external interpretation of sentences of this form, "there is no such condition [that a reason must satisfy an existing motive], and the reason-sentence will not be falsified by the absence of an appropriate motive."[3] As an example of an external reason statement, the sentence "Brian has a reason to give money to Oxfam" will be true (if it is true at all, that is) regardless of my motive, for, let us suppose, I have no motive to give money to charity. If the sentence "Brian has a reason to give money to Oxfam" were analyzed under the internal interpretation then the statement would be false, for, as just mentioned, giving money to charity would serve no motive of mine. Williams ends up concluding that all external reason statements are false, or, to put it another way, that there are no external reasons.

　　More needs to be said concerning the relationship between an agent's motivations and their reasons for action. It would be a mistake to think that an agent has an internal reason to φ just in case φ-ing would satisfy some immediate desire of the agent's. Indeed, Williams calls this view the "sub-Humean model," rejecting it for its simplicity. Instead, Williams claims that internal reasons need not be grounded in some immediate desire, but rather

---

[2] Ibid., 101.
[3] Ibid., 101.

grounded in an agent's subjective motivational set, what Williams labels an agent's S. An agent's S is meant to be broad and encompassing, and its borders are a bit nebulous: it "can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent."[4] By grounding internal reasons in an agent's S rather than mere desires, we allow for an agent to have internal reason to carry out long-standing pro-attitudes even if these pro-attitudes are at odds with sudden urges.

As an example, I might have a long-standing pro-attitude to donate money to charity. When Red Cross calls and asks for a donation, I might not, at that particular moment, feel like donating money. But I still have internal reason to donate money to charity because of my long-standing commitment to do so, even though the pro-attitude might currently lie in abeyance. Of course, our mercurial desires are also a part of our S. Thus, in the situation above, I also have internal reason not to donate money to Red Cross, because of my suddenly-acquired desire not do so. What remains undetermined is which internal reason is the stronger reason. The truth-condition of an internal reason statement is now as follows: an agent has an internal reason to φ just in case φ-ing would satisfy some member of the agent's S. Conversely, we can define external reasons as follows: an agent has an external reason to φ just in case φ-ing would not satisfy any member of the agent's S.

In recounting Williams' argument against external reason statements ever being veridical, I rely on the work of Elijah Milgram.[5] In keeping with Milgram and other able interpreters,[6] I hold that, when examining Williams' paper, two separate arguments present themselves. The first argument, which I shall call the *explanatory argument*, appeals to the intuitive notion that reasons must be capable of explaining actions, or, we might say, that reasons must have an explanatory component. I formulate the explanatory argument as follows:

(E1) "…reasons must figure in some correct explanation of [an agent's] action."[7]
(E2) "…nothing can explain an agent's (intentional) actions except something that motivates him to act."[8]
(E3) From (E1) and (E2) it follows that for reasons to explain actions, they must bear some relationship to an agent's motivations.
(E4) An agent's motivations are a proper subset of an agent's "subjective motivational set," what Williams labels an agent's S.

---

[4] Ibid., 105.
[5] Elijah Millgram, "Williams' Argument Against External Reasons," 198.
[6] Rachel Cohon also draws the same general divide, highlighting an argument predicated on explanation, and one predicated on motivation. See her "Are External Reasons Impossible?", 546.
[7] Bernard Williams, "Internal and External Reasons," 102.
[8] Ibid., 107.

(E5) By (E3) and (E4) it follows that reasons must bear some relationship to an agent's S.

(E6) Internal reasons are reasons that "display a relativity… to the agent's *subjective motivational set*… the agent's S."[9]

(E7) From (E5) and (E6) it follows that all reasons are internal reasons.

(E8) By the converse of (E7) it follows that no reasons are external reasons.

Thus, from the assumed premise that reasons must be capable of explaining an agent's actions, and from another assumed premise characterizing what explanation requires (namely, some connection to an agent's motivations), we conclude that there are no external reasons.

The second argument Williams offers against the external reasons theorist shall be called the *motivational argument*. It is in this argument that we see Williams' Humean sympathies come to light, as the argument relies on a variation of Hume's famous point that reason alone cannot give rise to motivation. The motivational argument goes as follows:

(M1) As a basic criterion of external reasons, "if the agent rationally deliberated, then, whatever motivations he originally had, he would come to be motivated to [carry out the external reason in question]."[10]

(M2) "… rational practical deliberation [through which one acquires new motivations] in each of its forms has as its starting point the subjective motivational set had by the agent prior to this deliberation."[11]

(M3) In the case of the agent coming to be motivated to carry out an external reason through practical deliberation, "there is no motivation for the agent to deliberate *from*, to reach this new motivation," because external reasons, by definition, are not grounded in an agent's S.[12]

(M4) By (M2) and (M3) it follows that practical deliberation cannot give rise to the motivation to carry out an external reason, for practical deliberation must be grounded in an agent's S, and the external reason, by definition, is not grounded in the agent's S.

(M5) (M4) contradicts (M1).

(M6) Since (M1) is a basic requirement that external reasons must meet, and since we are shown by (M5) that the requirement is unsatisfiable, it follows that there are no external reasons.

---

[9] Ibid., 102.

[10] Ibid., 109.

[11] Brad Hooker, "Williams' Argument against External Reasons," 42. This quote is not from Williams, but rather taken from an excellent exegetical piece by Hooker. The particular statement quoted summarizes a section wherein Hooker collects and analyzes all the statements Williams makes about practical deliberation and determines, as shown in the quoted section, that, for Williams, practical deliberation is grounded in an agent's S.

[12] Bernard Williams, "Internal and External Reasons," 109.

While the explanatory and motivational arguments seem quite distinct from one another, it is important to understand how Williams connects them if we are to show how an external reason might explain an agent's actions.

Tracking how Williams gets from the explanatory argument to the motivational argument will illustrate the connection between the two. Before beginning the explanatory argument, Williams prefaces it by saying that "no external reason could *by itself* offer an explanation of anyone's actions."[13] This is of course shown in the explanatory argument, but what this statement suggests is more important. The statement implies that there might be something – perhaps a fact about the agent's psychology – to supplement the external reason with, such that, if present, this supplement would allow for the external reason to perform its required explanatory function and thus satisfy (E1).

Right after giving the explanatory argument Williams attempts to find this supplemental feature: what if, he posits, the agent in question believed the external reason he supposedly has; this, maybe, could constitute a sufficient connection to an agent's motivations allowing for the supposed external reason to play an explanatory role. Here, Williams responds with a proof by cases: either believing the external reason constitutes being motivated to carry out the external reason or it does not. If it does not, then "we are no further on." If it does, then "an *internal* reason statement could truly be made" (i.e., the agent has an internal, not an external reason).[14] Either way, the response in question fails to show that external reasons can explain actions, disallowing these external reasons from satisfying (E1).

But Williams is not done after this failed attempt. He asserts that "it does not follow from this that there is nothing in external reason statements," suggesting that there might be some other way of supplementing external reasons so that they might be capable of explaining an agent's actions. Here, Williams offers another attempt at finding this link: "…the content of the external type of statement will have to be revealed by considering what it is to *come to believe* such a statement – it is there, if at all, that their peculiarity will have to emerge."[15] This of course leads us into premise (M1), which simply fleshes out what this coming-to-believe process consists in (where one comes to believe and be motivated by deliberation divorced from preexisting dispositions). This link between the explanatory and motivational arguments, where Williams attempts to find something to supplement external reasons with to allow them to perform their required explanatory function, suggests that Williams seems to think that *if* the external reasons theorist shows that (M1) is possible, *then* they will show that an external reason can explain an agent's actions. But (M1) cannot be satisfied – so the motivational argument goes – due to the nature of practical deliberation (M2) and the nature of external reasons (M3). Thus, external reasons cannot explain an agent's actions, but only because

---

[13] Ibid., 106.
[14] Ibid., 107.
[15] Ibid., 108.

(M1) is unsatisfiable. In some counterfactual world where this criterion is able to be met, an external reason *could* explain an agent's actions.

Admittedly, it remains unclear how satisfying (M1) will allow for an external reason to explain an agent's actions. To see this point, suppose that some agent satisfies (M1), implying the falsity of (M2). Satisfying (M1) means the following: our agent, at some point in time, has an external reason to φ. Furthermore, our agent, at this time, is not motivated to φ. Our agent deliberates in a manner detached from his S, and, through such deliberation, becomes motivated to φ. By definition, since our agent has acquired the motivation to φ, it follows that our agent now has an internal reason to φ. If and when our agent actually φ's, how will the agent's pre-deliberation external reason to φ explain the agent actually φ-ing? Indeed, it seems like the internal reasons theorist can simply say something like this: the agent, pre-deliberation, supposedly has an external reason to φ. Post-deliberation, the agent has gained the motivation to φ, and thus, by definition, an internal reason to φ. When the agent actually φ's, it is this newly acquired internal reason that will explain the φ-ing; the pre-deliberation external reason will not be involved in this explanatory process. Since this supposed external reason to φ is removed from the internal reason to φ (which explains the φ-ing), this external reason does not explain why the agent φ-ed, and thus, since it is unable to explain the agent's φ-ing, is not really a reason at all. This sort of challenge is what motivates the paper's guiding question: *even if* (M1) is satisfied, how does the satisfaction of (M1) allow for an agent's supposed external reason to φ to explain the agent's actually φ-ing?

Let me take some time here to make explicit what it is I am trying to do in this paper. There are some philosophers who think that we have external reasons for action. Derek Parfit and Thomas Nagel come to mind here. These philosophers often respond to Williams by engaging with the motivational argument. In doing so, these thinkers claim that (M2) is false: that practical deliberation *can* operate in a manner detached from an agent's dispositions, and, in doing so, can allow for an agent to acquire the motivation to carry out some action *not* grounded in prior motivation.[16] Thus, these thinkers believe they can show that (M1) is satisfiable, ostensibly meeting Williams' challenge. But there is another basic condition that Williams thinks external reasons must meet, laid out in premise (E1). Namely, external reasons must be capable of explaining actions. Therefore, those external reasons theorists who believe (M1) can be satisfied still have work to do: they must also show that (E1) is satisfiable, or they can argue that we need not hold on to (E1) in the first place.

Under my interpretation of Williams (as laid out above), I think Williams meant to say that when (M1) is satisfied (E1) will also be satisfied. If I am right, then there should be little work to do for those external reasons theorists who argue for the satisfiability of (M1). But, as I have just shown, it is unclear how the satisfaction of (M1) leads to the satisfaction of (E1). We must then ask: even when (M1) is satisfied, how is (E1) also satisfied? In other words,

---

[16] See, for instance, Derek Parfit, "Reasons and Motivation," 105. Also see Thomas Nagel, *The Possibility of Altruism*, chapter five (pp. 27-32).

even if an agent deliberates in a manner detached from their S and acquires the motivation to carry out some external reason they have, how will this external reason explain the agent's future actions? As already mentioned in the introduction, I *do* think there is an intelligible way of understanding how external reasons can explain actions if (M1) is satisfied, even though this method of explanation does indeed deviate from the traditional manner of reason-giving explanation. To argue for this conclusion I assume that (M1) is satisfiable; this assumption commits me to the falsity of (M2).

There are two responses the external reasons theorist can give here to show how satisfying (M1) allows for external reasons to explain actions. These two responses will be vetted in the next two sections of this paper. One response goes like this: the pre-deliberation external reason and the post-deliberation internal reason are, in actuality, the *same reason*. More specifically, as the agent goes through the deliberation process, the agent's relationship to the reason changes: pre-deliberation the agent is *not* motivated to carry out the reason, post-deliberation the agent *is* motivated to carry out the reason. This change in motivation by definition changes the type of the reason: pre-deliberation the reason is an external reason, post-deliberation the reason is an internal reason. But even though the agent's relationship with the reason changes which in turn changes the type of the reason, the reason still remains the same reason throughout the deliberation process. Thus, when the internal reason post-deliberation explains why the agent φ-ed, the external reason is also explaining, for it is the same reason as the explaining internal one. I think this response is off the mark, and shall show why it is incorrect in the next section. If (M1) is meant to show how external reasons might explain actions, it is not because the pre-deliberation external reason is the same reason as the post-deliberation internal one.

Here is another response as to how satisfying (M1) allows for external reasons to explain actions: while it is the acquired post-deliberation internal reason that will do the brunt of the explanation as to why the agent φ-ed, the external reason still does some explanatory work by giving an account of the *origins* of the new motivation and subsequent internal reason. In other words, the external reason explains why an agent φ-ed in so far as it gives an account of why the agent has a disposition to φ in the first place. Of course, having reasons explain in this manner deviates from the traditional way of understanding how reasons explain actions (what this traditional view is shall be explicated later on). Even so, explanation in this manner seems like a perfectly acceptable option. Moreover, if we understand (M1) as saying something like this, then we can make sense of why Williams phrases (M1) in the specific way that he does, and, furthermore, why certain responses to Williams miss the mark. This account of how external reasons can explain actions is discussed in section four.

Before examining these two responses I need to make one more point. In suggesting the two responses above as to how the satisfaction of (M1) allows for external reasons to explain actions, I am not suggesting that Williams endorsed either view. If I knew what Williams meant in suggesting (according to my interpretation of him) that the satisfaction of (M1) shows the explanatory

capacities of external reasons, then there would be no need for this paper. But since I am not sure what he meant, I need to offer potential explanations of what he could have meant and then evaluate the plausibility of these potential explanations. That is what gives this paper its guiding question: *even if* (M1) is satisfied, how does the satisfaction of (M1) allow for an agent's supposed external reason to φ to explain the agent's actually φ-ing? I shall now examine two possible answers to this question.

3. This is the view that shall be shown false in this section: when (M1) is satisfied, the pre-deliberation external reason to φ and the post-deliberation reason to φ are, in actuality, the same reason. As the agent goes through the deliberation process the agent changes his relationship with his reason to φ. Pre-deliberation the agent is not motivated to φ, post-deliberation the agent is motivated to φ. This transformation of the agent's relationship to the reason changes the type of the reason. Pre-deliberation, when our agent is not motivated to φ, the reason is by definition an external reason. Post-deliberation, when our agent is motivated to φ, the reason is by definition an internal reason. Nonetheless, the agent's reason to φ remains the same reason throughout the deliberation process, and thus, when the post-deliberation internal reason explains why our agent φ's, the pre-deliberation external reason explains the φ-ing as well. Since this pre-deliberation external reason has explained our agent's φ-ing, (E1) has been satisfied, allowing for our pre-deliberation external reason to actually be a reason.

　　　　Let us flesh out this response with an example. Suppose that some agent, agent B, has an external reason to volunteer at a soup kitchen. Since B's reason to volunteer at a soup kitchen is an external reason, it follows by definition that B has no member of his S that would be satisfied by him volunteering. B's good friend, agent C, comes along one day and tells B that he has reason to volunteer at a soup kitchen and that he really ought to do so. Recalcitrant at first, B thinks about C's coaxing and, after deliberating about the matter, B comes to believe that he indeed does have such a reason, while also acquiring the motivation to carry out the reason (here, (M1) has now been satisfied). Since B has acquired the relevant motivation, and thus now has a member of his S that would be satisfied by his volunteering at a soup kitchen, it follows by definition that B now has internal reason to volunteer at a soup kitchen. But even though B's reason has changed from an external reason to an internal reason due to the shift in B's S (which happened through B's deliberating), B's reason remains the same reason. If and when B actually volunteers at a soup kitchen, B's internal reason will explain why B did so, and, because the post-deliberation internal reason and the pre-deliberation external reason are the same reason, B's external reason has explained B's actions as well, satisfying (E1).

　　　　Before showing why this view is false I need to establish a few facts about reasons for action. We know that reasons count in favor of bringing about certain states of affairs. I might have a reason to eat less high fructose corn syrup, to learn the mandolin, to write a philosophy paper, and to eat pistachio ice

cream. But beyond our reasons counting in favor of bringing about certain states of affairs our reasons also have *content*. The content of a reason is the consideration that counts in favor of bringing about the relevant state of affairs. As an example, I never have *just* a reason to eat pistachio ice cream; I have reason to eat pistachio ice cream *because* I like pistachio ice cream. Or I have reason to eat pistachio ice cream *because* I like pistachios but have never tried the ice cream flavor, and want to see if my pro-attitude towards the shelled nut carries over to the frozen dairy treat. Or I have reason to eat pistachio ice cream *because* I am trying to impress a female companion of mine who likes pistachio ice cream, even though I do not care for it. As a final example, I might have reason to eat pistachio ice cream *because* I am severely allergic to pistachios, and want to seriously harm myself in the most appetizing manner possible.

I have said that the content of a reason is the consideration that counts in favor of bringing about the relevant state of affairs. If I have a reason to φ then the consideration that counts in favor of me φ-ing is the reason's content. This definition of content, though, is useless if we do not know what constitutes a consideration. What counts and does not count as a consideration will depend on the sorts of reasons we are dealing with. With internal reasons, it remains a commonsense response to say that the consideration counting in favor of me carrying out the internal reason will be the fact that were I to carry out the internal reason, some member of my S would be satisfied. Thus, the content of an internal reason will explicate how carrying out the internal reason satisfies some member of the relevant agent's S. As an example, if I have an internal reason to buy a certain guitar then the content of this internal reason will explain how buying the guitar will satisfy some member of my S: perhaps that I want to complete my guitar collection and the particular guitar in question will complete the collection.

When it comes to external reasons, what counts as a consideration and thus what counts as a content depends on how one grounds external reasons; that is, what their truth conditions are. If one believes that we are given external reasons from certain "irreducibly normative truths," then the consideration of an external reason will explicate what the relevant irreducibly normative truth is. An example: Parfit argues that there is an irreducibly normative truth saying that agony is objectively bad; this truth then gives us external reason to avoid future agony. This external reason to avoid future agony will thus have the following content: we have external reason to avoid future agony because agony is objectively bad. Some Kantians also believe that there are external reasons, though they do not believe that there are irreducibly normative truths. Instead, these Kantians often say that we have external reason to φ because our rational agency demands it. In such cases, the consideration counting in favor of carrying out one's external reason to φ will refer to how φ-ing is required by rational agency. An example: I have external reason not to tell a lying promise because telling a lying promise violates my rational agency by failing to pass the categorical imperative test.

That our reasons for action have content allows us to determine when one reason, say reason r, is the same reason as some other reason, say reason s.

It is an obvious mistake to say that two reasons are the same reason as one another just in case both reasons count in favor of bringing about the same state of affairs. An agent who has reason to go to church because he wants to worship his God, and another agent who has reason to go to church because he wants to rob the place and is casing the joint, do not go to church for the same reason. But when two reasons share the same content as one another we can say that they are the same reason. An example: agent D's mom is rushed to the hospital for some sudden affliction, and because she cannot tend to her affairs, she asks D to do a few things for her: to water the plants, to let the dog out, and to tell the bridge club that she will not be able to make it to that night's meeting. Agent D has reason to do all these things, and we can say that the content of all three reasons will be the same: because his mother is in the hospital and he wants to help her out. In such a case, it seems natural to say that D's reason to water the plants and his reason to let the dog out and his reason to apprise the bridge club are all the same reason: namely, because D's mom is in the hospital and he wants to help her out. D's three reasons are all the same reason because all three reasons share the same content as one another.

Let us return to our example supporting the thesis under investigation and give B's reason to volunteer at a soup kitchen content. Let us say that B has external reason to volunteer at a soup kitchen because volunteering at a soup kitchen is what a virtuous person would do. In retelling our story, when agent C admonishes agent B, he tells B that B has reason to volunteer at a soup kitchen because that is what a virtuous person would do, and, as a result, that B should do so too. Although B has, at first, no member of his S that would be satisfied by carrying out the external reason, B thinks about C's words and comes to realize that he does have reason to do so, and, in acquiring the motivation to volunteer at a soup kitchen, B now has a member of his S that would be satisfied by him doing so ((M1) has now been satisfied). Later on, B approaches C: "you were right all along," B says, "I *do* have reason to volunteer at a soup kitchen *because* that is what a virtuous person would do." If and when B actually volunteers at soup kitchen, B's newly acquired internal reason explains B's actions. Since the pre-deliberation external reason to volunteer at a soup kitchen is the same reason as the post-deliberation internal reason to volunteer at a soup kitchen, the pre-deliberation external reason has explained B's actions as well as well, satisfying (E1).

We can begin putting pressure on the view being considered by inquiring whether B's post-deliberation internal reason is actually an internal reason. Let us begin by looking at the supposed internal reason's content. If we take B's word for it from the paragraph above, the content of B's post-deliberation internal reason is the same as B's pre-deliberation external reason. Namely, B claimed that he had internal reason to volunteer at a soup kitchen because that is what a virtuous person would do. Recall that internal reasons are defined as reasons that satisfy some member of the S of the agent employing the reason. Moreover, the whole point behind the internal conception of reasons is that an agent has an internal reason *because* fulfilling the reason in question would satisfy some member of that agent's S. Given the nature of internal

reasons, it seems natural to say that the content of an internal reason should reflect the fact that internal reasons are reasons that satisfy some member of the S of the agent employing the reason. That is, the content of some agent's internal reason to φ should reflect the fact that φ-ing would satisfy some member of the agent's S, for the whole point behind the internal conception of reasons is that internal reasons do just that.

If we accept the above thesis then it is unclear whether B's internal reason to volunteer at a soup kitchen is indeed an internal reason. B claimed in the few paragraphs above that he had internal reason to volunteer at soup kitchen because volunteering at soup kitchen is what a virtuous person would do. If we take this claim as an accurate representation of the content of B's internal reason then it would be difficult to accept that this reason is actually an internal reason, for the content of this supposed internal reason does not reflect the fact that volunteering at a soup kitchen would satisfy some member of B's S. Of course, it remains a common sense response to say that if B fully expressed the content of his internal reason then we would find a clause referring to how the volunteer work would satisfy some member of B's S. It is just a fact about our language and our desire for brevity that led to B claiming that he had reason to volunteer at a soup kitchen *merely* because that is what a virtuous person would do; in other words, an ellipsis of sorts. But, if the content of B's internal reason were fully spelled out, we would get something like this: B has internal reason to volunteer at a soup kitchen because that is what a virtuous person would do and B wants to be a virtuous person.

We have determined that B's post-deliberation internal reason has a content that does refer, once fully spelled out, to the satisfaction of some member of B's S. Thus, the post-deliberation internal reason is actually an internal reason. But now we have a new problem: our fully spelled out content for B's post-deliberation internal reason no longer matches the content of B's pre-deliberation external reason. Recall that the content of B's pre-deliberation external reason to volunteer at a soup kitchen was because that is what a virtuous person would do. As we just determined, the content of B's post-deliberation internal reason is because that is what a virtuous person would do *and* B wants to be a virtuous person. Since the pre-deliberation external reason and the post-deliberation internal reason have different contents, we must conclude that they are *not* the same reason as the thesis under scrutiny claims. Now, when the post-deliberation internal reason explains why B volunteered at a soup kitchen, the pre-deliberation external reason, because it is an entirely different reason, is removed from the explanation of B's actions, unable to satisfy (E1).

Here, the defender of this thesis might respond: let us edit the content of B's pre-deliberation external reason so it can match the content of B's post-deliberation internal reason. Now, the content of the pre-deliberation external reason to volunteer at a soup kitchen is as follows: because that is what a virtuous person would do and B wants to be a virtuous person. But now we have a new problem. External reasons by definition are reasons that do not satisfy any member of the S of the relevant agent. Just as we determined that the content of

internal reasons should reflect the fact that internal reasons satisfy some member of the S of the relevant agent, it seems natural to claim the opposite for external reasons: that the content of an external reason cannot refer to the satisfaction of some member of the S of the relevant agent. For if an external reason's content *did* refer to the satisfaction of some member of the S of the agent employing the reason, how is the reason external to the agent's S? If we accept this result then it is obvious why the proposed revision fails: by changing the content of B's external reason to volunteer at a soup kitchen in the manner suggested, the external reason is no longer an external reason at all. Rather, the pre-deliberation reason is an internal reason. If this were the case then the whole point of this exercise would be lost, for there is now no longer an external reason that we can show performing an explanatory function, leaving (E1) unsatisfied.

The defender of the thesis under investigation might demur to the claim that the content of B's pre-deliberation external reason *must* change during the deliberation process, yielding a new content for B's post-deliberation internal reason. Let us suppose that the content of B's external reason does not change as (M1) is satisfied. As has already been shown, if the content of B's post-deliberation reason does not reflect the fact that volunteering at a soup kitchen would satisfy some member of B's S, then it would be difficult to consider the reason an internal reason. But there is another possibility here: if we hold that the content of B's pre-deliberation external reason does not change through the deliberation process, we can say that, post-deliberation, B's pre-deliberation external reason remains an external reason. Since B has acquired the motivation to volunteer at a soup kitchen during the deliberation process, it follows by definition that B also has, post-deliberation, an internal reason to volunteer at a soup kitchen. Now we have this picture: pre-deliberation, B has external reason to volunteer at a soup kitchen. Post-deliberation, B's external reason remains an external reason and B also has a newly acquired internal reason. Thus, there are two post-deliberation reasons to volunteer at a soup kitchen: an external reason and an internal reason.

If the above situation is the case then B's external reason to volunteer at a soup kitchen will still be unable to explain B's volunteering at a soup kitchen. As established in Williams' explanatory argument, external reasons, without supplement, are unable to explain actions due to their lack of relationship to an agent's motivations (premise (E2)). Even though B's external reason to volunteer at a soup kitchen is present post-deliberation and thus when B actually volunteers at a soup kitchen (that is, if he even does so), it will still run into the problems outlined in section two, and thus be unable to explain B's actions. Even so, B's newly acquired internal reason *will* be able to explain B's actions. This is of course no help to the external reasons theorist though: for under this scenario the internal reason is separate from the external reason, as the external reason remains constant through the deliberation process and the internal reason is newly acquired. Thus, B's external reason remains incapable of explaining B's actions, unable to satisfy (E1).

4. We have yet to establish how satisfying (M1) will show how external reasons can explain actions. We have already determined that it is not because the pre-deliberation external reason is the same reason as the post-deliberation internal reason. A different response goes like this: when an agent satisfies (M1) by deliberating themselves into being motivated to carry out an external reason of theirs, the external reason can explain the agent's actions in the sense that we can point to the external reason as the cause of the agent's newly acquired motivation and the subsequent internal reason that comes with it. The newly acquired internal reason will still perform an explanatory function, but this need not preempt any explanation the external reason might perform. The type of explanation that the external reason will perform is simply different from the type of explanation that the internal reason performs, and is indeed different from the traditional understanding of how reasons explain actions. What is this traditional conception of reason-giving explanation and how does Williams fit into this tradition?

To answer this question we can look to Donald Davidson's work on explanatory reasons. Explanatory reasons are "reasons we appeal to in attempting to explain actions and attitudes"[17] (Davidson uses the term "rationalize," so an explanatory reason allows us to rationalize an agent's actions). Explanatory reasons are different from normative reasons (what we have been talking about) in that a normative reason "is a consideration that counts in favor of or against doing something."[18] Davidson's explanatory conception of reasons looks like this:

Giving the reason why an agent did something is often a matter of naming the pro-attitude (a) or the related belief (b) or both... call this pair the *primary reason* why the agent performed the action.[19]

Constructing an explanatory reason, then, is a matter of citing two features. One must cite a pro-attitude and a belief. So if we wanted to construct a primary reason as to why some agent quit the football team we would need to cite a pro-attitude, say that our agent was afraid of getting injured, and a belief, say that our agent believed that playing football was injurious.

Unlike Davidson's conception of explanatory reasons, Williams' conception of internal reasons is normative. So how do we get from Davidson's conception of explanatory reasons to Williams' internal reasons? This is accomplished by dropping one of the necessary conditions that Davidson's reasons require while keeping the other. More specifically, the requirement that one needs both a pro-attitude and a related belief is reduced so that one needs only a pro-attitude (or, in Williams' terms, a member of the agent's S) with no belief.[20] Williams emphasizes that what gives his conception of internal reasons

---

[17] James Lenman, "Reasons for Action: Justification v. Explanation."
[18] Stephen Finlay and Mark Schroeder, "Reasons for Action: Internal v. External."
[19] Donald Davidson, "Actions, Reasons, and Causes," 686.
[20] Williams claims that "A has a reason to φ only if he could reach the conclusion to φ by a sound deliberative route from the motivations he already has [that is, φ-ing must satisfy some member of the agent's S]." Moreover, Williams thinks that this necessary condition

normative force is that it goes beyond merely describing what an agent is currently motivated to do: " 'A has a reason to φ' means more than 'A is presently disposed to φ.' "[21] The normativity comes in here as we can now tell an agent what he should or should not become disposed to do by correcting and critiquing his *beliefs* about certain situations, which might, if our revisions are correct, better lead our agent to satisfying some member of his S. In Williams' own words: "We are allowed to change – that is, improve or correct – his beliefs and his reasoning in saying what it is he has reason to do. This is already enough for the notion to be normative."[22] So in getting rid of the belief requirement we take an explanatory conception of reasons and transform it into a normative conception.

Since Williams believes that all reasons must be capable of explaining actions (premise (E1)), it follows that his internal reasons must be capable of explaining actions. With this internal conception of reasons, it is unclear if the explanation of actions is done solely by the fact that the internal reason is related to the agent's S, or whether this internal reason must be supplemented with a belief before it performs its explanatory function. There are two ways this might go: since internal reasons do not require beliefs, we can say that an internal reason to φ is sufficient for explaining why an agent φ's solely because the internal reason is grounded in a motive (as broadly construed) to φ. Moreover, the internal reason would explain the φ-ing by virtue of its content. As an example, if we are explaining why an agent went for a jog we would turn to his internal reason to go for a jog, and explain the jogging through the reason's content (for example, because the man wanted to be in good shape).

Or, Williams could say that while an agent has an internal reason to φ just when he has a a member of his S that would be satisfied by φ-ing, when the agent actually φ's, we will need to go beyond citing the content of the internal reason and also provide a belief, where the belief tells a story about how the agent thinks committing the action in question will lead to satisfying his motive to φ. So beyond citing the internal reason's content that the man wanted to stay in shape we would go beyond the internal reason and also provide a belief – perhaps that the man thought jogging was an excellent way of remaining in good health. While I am not sure what Williams would say here, let us, for the sake of simplicity, assume that Williams would follow Davidson and require that an internal reason, in order for it to explain an agent's actions, must be further supplemented with a relevant belief. Call these conceptions of reason-giving

---

"provides a sufficient condition as well." See his "Internal Reasons and the Obscurity of Blame," 35. Compare this to the two necessary conditions Davidson stipulates for his explanatory conception of reasons: one needs both a pro-attitude *and* a related belief, whereas Williams only needs the pro-attitude (or, in Williams' terms, a member of agent's S).

[21] Ibid., 36.

[22] Ibid., 36.

explanation, where one cites a desire (as broadly construed) and a belief, *desire/belief models* of explanation.[23]

I think it is pretty clear that external reasons are not able to explain actions via desire/beliefs models of explanation. To see this, consider the following: suppose that an agent has an external reason to avoid agony. What is the external reason's content? As has been established in section three, the content of an external reason cannot refer to the satisfaction of any member of the S of the agent employing the reason. In this case, let us say the content of the external reason to avoid agony is that agony is objectively bad. Suppose our agent is about to encounter a particularly agonizing situation, but, at the last minute, avoids the situation. Here, can the agent's external reason to avoid agony explain the agent actually avoiding agony? In attempting to explain the action with the external reason we would turn to the external reason's content – that agony is objectively bad – and realize that there is nothing about the supposed badness of agony that can explain why our agent avoided the agony. Maybe if the content of the external reason was that agony is objectively bad and that the agent in question wants to avoid things that are objectively bad. But in this case, the content of the supposed external reason refers to the satisfaction of some member of the S of the agent employing the reason ("the agent wants to avoid things that are objectively bad"), and thus is now no longer an external reason. External reasons, due to their lack of reference to the satisfaction of some member of the S of the agent for whom they are reasons for through the reason's content (Williams captures this point in premise (E2)), are unable to explain actions through desire/belief models of explanation, which, as has been shown, are traditionally used in reason-giving explanation.

But there might be other ways for reasons to explain actions than the traditional desire/belief models of explanation just discussed. To open the door to this possibility, consider the following discourse: "why did that child run away in terror at the sight of that pigeon?", "because when he was a little boy he was attacked by a swarm of angry pigeons in Trafalgar Square." Here, nothing has directly been said about the boy's motivations or desires. Even so, the answer to the question of why the boy did what he did was sufficiently explanatory even though nothing about the boy's pro-attitudes was cited. Take another example: "why did that man refuse to eat that appetizing steak?", "because he just watched a documentary revealing the dark side of the beef industry." Again, no desires or motivations have been directly invoked in the explanation of the agent's actions. Still, the explanation given seems just as

---

[23] I have claimed that desire/belief models of explanation are traditionally used in reason-giving explanation. Here are some other thinkers (besides Davidson and Williams) who conceive of reason-giving explanation in such a manner: Alvin Goldman, *A Theory of Human Action*, 78; Fred Dretske, *Explaining Behavior: Reasons in a World of Causes*, 109; Robert Audi, "Acting for Reasons," 146; and Daniel Dennett, *The Intentional Stance*, 49. Furthermore, for a historical background of desire/belief models of explanation, see Rüdiger Bittner, *Doing Things for Reasons*, chapters one and two (pp. 3-42).

sufficient in explaining why the man refused the steak as a more traditional explanation would, where we would cite the man's desire to avoid beef products and his belief that the steak was a beef product.

Of course, one might say that we have implicitly said something about the boy's and the man's dispositions: namely that the boy is terrified of pigeons and that the man is disgusted by beef. I think it is true that we have said something about these agents' dispositions, and that is part of the point. We have said *why* they have the fear of pigeons and disgust of steak that they have. No doubt the boy's running from the pigeon could be explained with one of the desire/belief models of explanation listed above. In following Davidson we could construct a primary reason by citing the fact that the boy has a fear of pigeons and also believes that the bird flying towards him is a pigeon. But just because we can explain the boy's actions with a desire/belief model of explanation it does not follow that the first attempt at explanation – the one in which we invoke the origins of the boy's dispositions – is any less explanatory.

Even if we accept that actions can be explained by giving an account of the origins of the motives and internal reasons we have to carry out these actions, we still have not said anything about external reasons. Indeed, the first example explains the boy's actions in terms of an event (that tragic day at Trafalgar Square) rather than a reason. This is of course fine; I think citing an event is perfectly acceptable in this non-traditional conception of explanation. But, if we can pinpoint the origin of some agent's motivation to φ on an external to φ, then this external reason will be performing the same function as the event was performing above, and thus *can* explain why the agent φ's, for it tells us why the agent is disposed to φ in the first place. There are many things that can cause one to have the dispositions and motivations that one has, and each different causal agent can be invoked to explain. If an external reason can cause an agent to be motivated to carry out an action, then it too can be explanatory. Call this sort of reason-giving explanation the *causal model* of explanation.[24]

More needs to be said about how an agent's external reason to φ can cause an agent to acquire the motivation to φ. Indeed, it would be a queer thing if I, when walking down the street one day, were suddenly caused out of the

---

[24] In one sense, some desire/belief models of explanation are causal. This is because, many action theorists argue, reasons (like Davidson's primary reasons composed of pro-attitudes and beliefs) *cause* actions. As an example, I might have a reason to drink some potato vodka, and, according to this thesis, this reason causes me to actually drink the alcoholic beverage. But when I use the technical term "causal model of explanation" I refer to a type of reason-giving explanation whereby a reason causes an agent to acquire a *motivation*, not to commit an *action*. Thus, these desire/belief models of explanation that conceive of reasons as causing actions are not causal models of explanation as I understand them. Some thinkers who believe that reasons cause actions are as follows: Davidson, Goldman, and Dretske. Of course, not all those who employ desire/belief models of explanation hold that reasons cause actions. For a good breakdown that separates causal (in terms of reasons causing actions) desire/belief models of explanation and non-causal desire/belief models of explanation, see G.F. Schueler's *Desire: Its Role in Practical Reason and the Explanation of Action*, chapters four and five (pp. 115-170).

blue to acquire the motivation to φ from some external reason of mine. Instead, the agent who has the external reason to φ will have to do some work to acquire the motivation to φ. The obvious candidate here is that the agent realizes that he has an external reason to φ, through something like moral intuition, and, from realizing this fact, deliberates his way to becoming motivated to φ (this process of deliberating one's self into being motivated to carry out an external reason is what premise (M1) requires).

Now obviously those entrenched in the Humean camp will deny this possibility, arguing that there is no faculty of deliberation that could accomplish this. But not everyone agrees with this conception of practical deliberation – that practical deliberation necessarily proceeds from an agent's preexisting dispositions. There are those who argue that deliberation *can* be a process detached from agent's S, and, even given this lack of attachment, is still capable of motivating an agent. The purpose of this paper is not to engage in a debate concerning the nature of practical deliberation and motivation. The point I am trying to make is that if an external reason to φ is to cause an agent to acquire the motivation to φ then the agent will need to engage in some sort of deliberation detached from his dispositions. So it is not the external reason all on its own that causes the acquisition in motivation, but the external reason *and* the application of the agent's deliberation. That one needs to go beyond the external reason to offer a causal model of explanation is somewhat fitting, given that Williams' conception of internal reasons, as we are interpreting it, requires the addition of a relevant belief to an internal reason in order to explain actions with a desire/belief model of explanation.

Let us see a causal model of explanation in action. Suppose that there is an external reason for E to donate money to charity. As is, this external reason will be unable to explain E's giving money to charity – if E ever decides to give money to charity – through a desire/belief model of explanation. But suppose that E, through her faculty of moral intuition, realizes that she has an external reason to give money to charity because giving money to charity would maximize societal happiness (this, we can note, is the content of E's external reason). From here E deliberates her way into being motivated to carry out this external reason, and, in the process, also acquires an internal reason to donate money to charity (after this step, agent E has satisfied (M1)). If and when E gives money to charity, we can explain her doing so in two ways. First, with her newly acquired internal reason to give money to charity, we can explain E's actions with a desire/belief model of explanation by citing the content of E's newly acquired internal reason and a related belief of E's. We can also use a causal model of explanation to explain E's action: "why did E donate all that money to charity?" "because she realized that she had reason to donate money to charity, as donating money to charity would maximize societal happiness." Here, E's external reason explains E's actions because it explains why E has the motivation to donate money to charity in the first place. Thus, with our causal model of explanation an external reason *has* explained an agent's actions, and has thus satisfied (E1). Moreover, we have answered the paper's guiding question: *if* (M1) is satisfied, *then* an external reason to φ can explain an agent's

φ-ing through a causal model of explanation, where the external reason to φ explains the agent's φ-ing by giving an account of the origins of the agent's disposition to φ.

That the external reason to φ must be the origin of the motivation to φ in order for us to make use of a causal model of explanation helps us make sense of why Williams phrased (M1) in the specific way he did. When reading the few paragraphs where the motivational argument is presented, one is struck by how incessantly Williams thinks that the coming-to-believe and motivation-acquiring process of (M1) must happen "*in a special way*." Not because the agent is guided by something like "moving rhetoric" or some other non-rational process, but because the agent is "seeing things aright" and has rationally deliberated in a manner detached from the agent's preexisting dispositions.[25] If the agent came to believe and be motivated to carry out a supposed external reason because they were swayed by moving rhetoric, then the external reason did not cause the acquisition in motivation like we want it to, disallowing us from appealing to our causal model of explanation. Something else has caused this acquisition (like convincing and poetic speech), which will then do the explaining, not the external reason.

Here is an example of this: suppose agent F has an external reason to give money to charity. Agent F acquires the motivation to give money to charity because she just watched a moving speech by Barack Obama, in which the president tried to inspire the nation to do their civic duty and help those less fortunate. If F donates money to charity, and we want to explain F's donating with a causal model of explanation, then we will get something like this: "why did F donate that money to charity?" "because Barack Obama inspired her to in a speech." The external reason to donate money to charity is *not* explaining F's actions because the external reason did not cause F to acquire the disposition to give money to charity; a rhetoric-filled speech did.[26] Instead, F's donating to charity, if we were to explain it with a causal model of explanation, would be explained by Barack Obama's moving speech rather than the external reason. But if F acquires the motivation to donate to charity *because* of the external reason to donate money to charity, which would happen with the help practical deliberation divorced from preexisting dispositions (what is spelled out in (M1)), then we can say that the external reason to donate to charity itself caused F's motivation to donate money to come about, and thus can be cited as an explanation of F's actions, satisfying (E1).

---

[25] Bernard Williams, "Internal and External Reasons," 108-109.
[26] Of course, it could be the case that the rhetoric-filled speech causes F to think about what it is she has reason to do. From here, F realizes that she has external reason to donate money to charity, and, in doing so, becomes motivated to donate. In this case the external *does* cause the acquisition in motivation, even though the deliberation was inspired by a speech. But in the case I am presenting above, I am assuming that the rhetoric-filled speech inspires F in a semi-hypnotic fashion, where the acquired motivation does not arise from deliberation inspired by the speech, but rather from the speech itself.

As a final note, we can examine how this interpretation of (M1) affects some of the criticisms launched against Williams' argument against external reasons. Since we are interpreting (M1) in such a way that allows for an external reason to explain actions via our new causal model of explanation, any attempt at editing the content of (M1) would thus disallow external reasons from performing their required explanatory function. As an example, John McDowell criticizes Williams on the grounds that the process of coming to believe and acquiring the motivation to carry out an external reason need not happen in such a narrowly defined manner. He asks: "why must the external reasons theorist envisage this transition to considering the matter aright as being effected by *correct deliberation*?"[27] From here, McDowell suggests that the transition from not being motivated to carry out an external reason to being motivated to carry out the external reason can happen in some non-deliberative manner, his example being conversion.[28] But if we allow (M1) to be edited in such a way, then there is nothing to guarantee that the external reason will be the cause of the acquisition in motivation, which is required if it is to explain via our causal model of explanation. In the case of the agent who undergoes conversion and is then motivated to carry out their supposed external reason, we can certainly say that the external reason was *not* the cause of the acquisition in motivation, and thus will be unable to explain in the manner that we want it to.[29] Because of this, McDowell's response to Williams is rendered impotent under our new interpretation of how external reasons can explain actions through the satisfaction of (M1).

5. To end this paper I will sum up what has been established. First, I have shown that there is a link between Williams' explanatory and motivational arguments, such that if one satisfies (M1) then one will show how an external reason can explain an agent's actions. Even so, it remained unclear how satisfying (M1) could allow for external reasons to explain actions. On one interpretation, the pre-deliberation external reason and the post-deliberation internal reason are the same reason, so that when the internal reason explains the agent's actions in the traditional manner of reason-giving explanation (desire/belief model), the external reason does too. I showed that this view is mistaken. From here I took a look at another way we can understand (M1). With our new causal model of explanation, external reasons can explain actions if they can be cited as the origins of a motivation or desire and the subsequent internal reason that comes along with it. This is certainly a deviation from traditional desire/belief models of explanation, but one, I have argued, that is satisfactory.

---

[27] John McDowell, "Might There Be External Reasons?", 99-100.

[28] Ibid., 102.

[29] John Brunero comes to a similar conclusion in his assessment of McDowell's response to Williams. He argues that what rules McDowell's response to Williams out is Williams' commitment to explanation, which McDowell ignores. See his "McDowell on External Reasons," particularly section three (pp. 26-29).

It is also important to note what has not been established in this paper. I have not shown that there are external reasons. I have only established that *if* one shows that (M1) is possible *then* there is an intelligible way we can conceive of external reasons explaining actions via a causal model of reason-giving explanation, such that external reasons can now satisfy (E1). But the point of the motivational argument is that (M1) is not possible due to the nature of practical deliberation (M2) and the nature of external reasons (M3). Thus, if the external reasons theorist wants to allow for external reasons to explain actions using the causal method of explanation that I have proposed, then they will have to argue against Hume's forceful thesis that reason alone cannot give rise to motivation, to show that an external reason can cause an acquisition in motivation with the aid of deliberation divorced from dispositions (that is, the external reasons theorist must argue that (M2) is false). To be sure, there are already those who *do* argue for this thesis, some in a quite convincing manner. Of course, there is another route the external reasons theorist might take. They could reject (E1) all together, arguing that normative reasons need not be caught up in the business of explanation. This, I believe, would be a worthy research question to pursue and, if answered, a significant contribution to the internal-external reasons discourse. The question then remains: must reasons explain?

Works Cited:

1.   Audi, Robert. "Acting for Reasons" in *Action, Intentions, and Reason*. Ithaca: Cornell University Press, 1993.

2.   Bittner, Rüdiger. *Doing Things for Reasons*. Oxford: Oxford University Press, 2001.

3.   Brunero, John. "McDowell on External Reasons." *European Journal of Philosophy* 16.1 (2008): 22-42.

4.   Cohon, Rachel. "Are External Reasons Impossible?" *Ethics* 96.3 (1986): 545-556.

5.   Davidson, Donald. "Actions, Reasons, and Causes." *The Journal of Philosophy* 60.23 (1963): 685-700.

6.   Dennett, Daniel. *The Intentional Stance*. Cambridge: MIT Press, 1987.

7.   Dretske, Fred. *Explaining Behavior: Reasons in a World of Causes*. Cambridge: MIT Press, 1988.

8.   Finlay, Stephen and Mark Schroder. "Reasons for Action: Internal vs. External." *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.),

<http://plato.stanford.edu/archives/fall2008/entries/reasons-internal-external/>.

9. Goldman, Alvin. *A Theory of Human Action*. Englewood Cliffs: Prentice Hall Inc., 1970.

10. Hooker, Brad. "Williams' Argument against External Reasons." *Analysis* 47.1 (1987): 42-44. 1996.

11. Lenman, James. "Reasons for Action: Justification vs. Explanation." *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2011/entries/reasons-just-vs-expl/>.

12. McDowell, John. "Might There Be External Reasons?" in *Mind, Value, and Reality*. Cambridge: Harvard University Press, 1998.

13. Millgram, Elijah. "Williams' Argument Against External Reasons." *Noûs* 30.2 (1996): 197-220.

14. Nagel, Thomas. *The Possibility of Altruism*. Princeton: Princeton University Press, 1970.

15. Parfit, Derek. "Reasons and Motivation." *Proceedings of the Aristotelian Society, Supplementary Volumes* 71 (1997): 99-146.

16. Schueler, G.F. *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge: MIT Press, 1995.

17. Williams, Bernard. "Internal and External Reasons" in *Moral Luck*. Cambridge: Cambridge University Press, 1981.

18. Williams, Bernard. "Internal Reasons and the Obscurity of Blame" in *Making Sense of Humanity*. Cambridge: Cambridge University Press, 1995.

Humean Constructivism, Minimalist Constitutivism, and Naturalism

Kevin Dorst
Washington University in St. Louis

Abstract:
In recent years, metaethical debates have increasingly discussed constructivism – the view that normativity is something that we construct out of our own activity as practical reasoners. By grounding this process of construction in certain constitutive norms of agency, constructivists have sought to reconcile the objectivity of practical normativity with a naturalistic worldview. Though ambitious forms of constructivism appear to face insurmountable difficulties, Sharon Street proposes a Humean version of constructivism that avoids some of these problems. Despite this success, I will argue that her account is deeply problematic due to its reliance on a minimal version of constitutivism. That, is, she invokes certain (putative) constitutive norms of the attitude of making a normative judgment, but the evidence that can be marshalled in favor of such norms is inadequate. I argue that she must justify such norms by appealing to facts about our phenomenology, and yet these facts can be perfectly well explained without invoking such constitutive norms. The upshot is that Street (and constitutivists more generally) will be hard-pressed to justify their claims to the inevitability of certain constitutive norms.

Sharon Street (2008) proposes a form of Humean metaethical constructivism that aims to reconcile practical normativity with a naturalistic understanding of the world. The basic thought is that after we've rejected normative realism (due to epistemological concerns[1]), the best way to understand facts about value is as a *constructed* out of what individuals *judge* to be valuable. Street's account is incredibly rich and nuanced, so I cannot treat it fully here. Instead I will focus on its ability to furnish us with an account of *normative error*, which rests on an appeal to a (minimal) version of constitutivism. Constitutivism (familiar from Korsgaard 1996 and Velleman 2009) is the project of attempting to ground normativity on what is constitutive of agency – or, in Street's case, what is *constitutively entailed* from our own normative judgments. Though constitutivism in general faces powerful objections (Enoch 2006; 2011), Street maintains that her minimalist version is able to avoid these worries (2010, 374). I think Street is plausibly correct in this claim, but I will use Enoch's objection as a springboard for presenting a more specific critique of her account. In short, Street's explanation of normative error rests on a claim that we as humans are *inevitably bound* by certain constitutive norms, and I will argue that this claim is unjustified and implausible.

In the first section I will lay out the main outlines of Street's constructivism, including how it extracts normative error out of the (putative) constitutive norms that govern the attitude of a normative judgment. In the second section I argue that in order to respond to the Enochean objection, Street must rest her claim to constitutive norms on an appeal to certain phenomenological facts about humans. In the third section I attempt to undercut this connection by offering alternative explanations of these phenomenological facts that fail to generate normative error. In addition to opening up possibilities that Street has not addressed, I contend that one of these explanations is actually *more* plausible than her constitutive norms, since it better coheres with her own naturalism. As such, Street's account of normative error is left needing modification or further justification. Finally, in the conclusion I sketch the implications of my arguments for constitutivism more generally.

I. Humean Constructivism about Practical Reason

Street's metaethical constructivism is "premised on a rejection of realism" (Street 2008, 220), where realism is understood as the thesis that normative facts are *mind-independent*, or true independent of our evaluative attitudes. The brunt of Street's argument against such realist theories goes as follows (see her 2006). Evolutionary pressures have strongly influenced the content of our normative judgments, as there are substantial selective pressures in favor of certain value judgments (such as valuing one's own survival, the survival of one's offspring, etc.). Since realism claims that normative facts are true independent of our judgments about them, then it must claim that either (i) there is *no relation* between these evolutionary pressures and the normative facts, or (ii) these pressures somehow *tracked* the independent normative facts. If we go with

---

[1] See Street 2006 for a powerful presentation of these worries.

option (i), then in all likelihood our normative judgments are hopelessly confused and pervasively false. On the other hand, option (ii) utterly fails as an explanation since there is a much simpler alternative: there *are no* independent facts to track. Instead of being selected to have *true* judgments, creatures were simply selected to have *adaptive* judgments. The upshot of this argument is (purportedly) that normative realism is indicted by naturalism, since a naturalistic understanding of the development of our own normative judgments forces realism into an untenable epistemology. From this argument Street concludes that we must accept some sort of anti-realist account of normativity in which normative facts are mind-*de*pendent.

What should such an anti-realist account be based upon? Street points out that "[e]ven if we aren't sure what *value* is, we *do* understand the attitude of *valuing*…." (2010, 366; my emphasis). This is "the attitude of *normative judgment* or of *taking something to be a reason*" that we are all familiar with (Street 2008, 230).[2] For instance, Street would describe the attitude of valuing your own survival as 'taking yourself to have reason to survive.' This attitude is not only familiar, but also naturalistically understandable: the point of having normative judgments is to *motivate us* to respond to our environment.[3] Combining these thoughts with its proximity to the notion of value, the attitude of *normative judgment* may be a good place to begin. But how precisely should we characterize this attitude? Street says there is a sense in which we must take it as a primitive that it is picked out phenomenologically: taking yourself to have a reason is a familiar "conscious experience… of one thing's seeming to *demand* or *call for* or *count in favor of* something else" (2008, 240). For instance, if you see a car heading straight toward you while driving, there is an unmistakable motivational 'push' in favor of swerving out of the way. However, the raw motivational push of normative judgments is importantly different from that of simple desires, as the former (but not the latter) follow the patterns of certain *constitutive norms*.[4]

To illustrate the concept of a constitutive norm, Street offers an example of someone who asserts both of the following propositions: (1) 'I am a parent'; and (2) 'I have no children.' The oddity here is a conceptual one: what it *means* to be a parent is to have children, so someone who (in full consciousness) asserts both (1) and (2) simply "does not understand the concept of parenthood…." (2008, 227). In other words, asserting that you are a parent *constitutively entails* asserting that you have children. Continuing Street's example, compare the above claims about parenthood to the following:

---

[2] In line with Street's usage, I use the phrases 'normative judgment,' 'taking/judging yourself to have reason,' and 'valuing' interchangeably. Likewise, all references to 'reasons' will be to *normative* reasons, not explanatory reasons (i.e. reasons that justify an agent's action, not ones that explain it).

[3] See Street 2006, 118-21 for a discussion of how the attitude of normative judgment likely evolved.

[4] Street 2008, 230-1. I will return to the distinction between normative judgments and desires below.

(3) I have conclusive reason to get to Rome immediately.
(4) Getting on a plane is the only way to do so.
(5) I have no reason to get on a plane.

Street contends that someone who asserts (3)-(5) is doing something odd in precisely the same way as someone who asserts both (1) and (2): "she isn't making a *mistake* about what reasons she has; rather, she simply doesn't count as genuinely making the first 'normative judgment' (or for that matter the second) at all. She's not doing what's constitutively involved in taking oneself to have a reason" (228). Thus one example of a constitutive norm of normative judgments is that they must exhibit means-ends consistency, at least "when attending to the matter in full awareness."[5] Further, Street is at pains to emphasize that this discussion of constitutive norms is (purportedly) a purely *formal* characterization of the attitude of valuing, and "does not itself presuppose any normative notions" (2010, 367). These claims are necessary to prevent Street's constructivist account from becoming viciously circular: if her constitutive norms invoked normative notions, then we obviously could not informatively specify an account normativity with these norms.

So given this account of constitutive norms, how do we get an explanation of normative error? According to Street, when you are in *full awareness* of your own normative judgments and the non-normative facts, then it is impossible to be mistaken about what your judgments entail. However, when you are *not* in such a state of full awareness, then you can make instrumental mistakes by failing to be motivated in accordance with your own commitments. For example, suppose that you judge yourself to have conclusive reason to Y, and yet you are unaware that Z is a necessary means to Y. On Street's proposal, if you then fail to Z, you are making a mistake by your own lights:

[S]imply by judging yourself to have reason to Y, you're *thereby* – as a constitutive matter – also judging yourself to have reason to take the means to Y, whatever those may be. So even if you don't know that Z is a means to Y, and think that you have no reason whatsoever to Z, you *do* have reason to Z – *according to you*. (2008, 229)

So there is a sense in which making normative judgments serves to "give laws to oneself" (229), in that you are committed to having reason to do what those judgments entail.
Though Street never uses it, there is an analogy with assertions that helps to draw this claim out. If John asserts '$P\&Q$', he *thereby* (as a constitutive matter)

---

[5] Street 2008, 228. I will come back to the precise formulation of this means-ends consistency norm (and its reliance on idealization) below, as this will be important in my critique. Although Street identifies several other constitutive norms of normative judgments (see 2008, 229 and footnote 38), for simplicity I will only explicitly deal with means-ends consistency. Nothing of substance hangs on this matter however, as my discussion is easily translatable to apply to the other norms as well.

asserts '*Q*' – that's simply (part of) what the conjunction *means*. So suppose that John asserts '*P&Q*', continues to go about his day, and then later asserts '¬*Q*'. Assuming there have been no relevant changes in his evidence or beliefs, there is a very clear sense in which John has made a mistake by his own commitments. This is the type of intuition that Street tries to use to proffer an account of normative error. (I admit that it offers a powerful intuition, but I will argue that the assertion case is not analogous to that of normative judgments.) Generalizing this discussion, Street contends that the *actual* (practical[6]) normative reasons an individual has are to be *constructed* out of her *judgments* about her reasons, according to the constitutive entailments of these judgments. As she puts it, "[a]ccording to *metaethical constructivism*, the fact that *X* is a reason to *Y* for agent *A* is constituted by the fact that the judgment that *X* is a reason to *Y* (for *A*) *withstands scrutiny* from the standpoint of *A*'s other judgments about reasons" (2008, 223), where this process of 'withstanding scrutiny' is cashed out in terms of the constitutive norms discussed above. Further, Street offers a *formalist* exposition of metaethical constructivism (in contrast with Korsgaard's) in which "there is nothing in particular that one must value if one values anything at all" (244). Thus what we have reason to do ultimately depends – and depends completely – on what we *take* ourselves to have reason to do.

Street's account is much more complex than I have been able to do justice to here, but I hope the basics are now clear. The general picture is one of 'constructivism all the way down,' in which practical normativity is radically subjective, and yet there are very substantive ways in which individuals can go wrong.[7] I take it to be clear that the plausibility of her entire account rests on this point about normative error: if it failed to explain how individuals could make such errors, it would be closer to an error theory or a form of nihilism than anything else. And since her explanation of these normative errors rests on her constitutivism, we can question her entire account by showing her use of these (putative) constitutive norms to be problematic. In the remaining sections I will attempt to expose such problems.

II. Shmagency and Phenomenology

My basic argument against Street's account is that (1) it must claim that her constitutive norms are the best explanation of certain phenomenological facts about humans, and yet (2) there are alternative, *better* explanations of these facts

---

[6] Street elsewhere (2011) suggests a similar constructivist account for *theoretical* reason based on the constitutive entailments of the attitude of 'belief'. See also Velleman 2000 for a similar constitutivist proposal. I believe that these accounts face analogous objections to those I will present in this paper, but I cannot pursue the possibility here.

[7] Some may think this account is clearly *too* subjective, for it seems to clearly allow that an "Ideally Coherent Caligula" will be morally obligated to torture others. Street bites the bullet here, but see her 2009 for an interesting response to this objection. For other objections – such as worries that her account is circular, self-defeating, or collapses into other metaethical theories (this last worry is raised forcefully by Enoch 2009) – see Street 2008; 2010; and ms.

that fail to generate normative error. In this section I will argue for (1) in three steps. First I point out that Street needs to claim that we as humans inevitably take ourselves to have *reasons* in the precise sense of her term; otherwise people could simply make use of other types of motivational attitudes that do not generate normative error. Second, one way to justify this inevitability claim would be to show that such alternative motivational attitudes are conceptually problematic, but in II.2 I argue against this possibility. Third, in II.3 I suggest that the best option for Street is to ground her constitutivism in an appeal to phenomenological facts about humans. I then address this appeal in section III by offering alternative explanations of these facts.

II.1: Inevitability and Shmagency
Constitutivists like to emphasize that they are merely fleshing out the meanings of our concepts – they take themselves to be engaged in "an exercise in descriptive philosophical analysis" (Street 2010, 374) when they specify what is constitutively involved in 'agency' or the attitude of 'normative judgment.' So presumably they are simply giving us tautologies: 'a normative judgment aims at means-ends consistency' is supposed to be on par with 'a bachelor is an unmarried male.' However, the first step to my critique is to show that putting forward tautologies cannot be *all* that constitutivists are doing. To illustrate this, I will simply grant Street's claims about the constitutive entailments of the attitude of taking oneself to have a reason. Thus we have (roughly):
The attitude of 'taking oneself to have reason to *X*' constitutively entails an aim of means-ends consistency.[8]

Is (A) all we need to show that individuals have made normative errors when they fail to exhibit means-ends consistency? No.
The problem can be brought out by an objection forcefully presented by Enoch.[9] Suppose we confront Jill by showing her that she takes herself to have conclusive reason to *X,* and that *Y* is a necessary means to *X*, so therefore she is making an error if she does not *Y*. If all we have from (A) is a mere tautology, then Jill can simply respond: "I must have misspoke. I don't take myself to have *reason* to *X*; I take myself to have *shmeason* to *X* (which is like a reason, only it does not entail an aim of means-ends consistency). So I have *not* made any commitments, and I make no error if I do not *Y*." The problem is that (A) is merely (part of) the formal specification of the concept NORMATIVE JUDGMENT: there is no requirement that people actually *use* the concept. Or if they do use it, whenever we go to show them their errors they can simply shift out from under us: instead of reasons they have *shm*easons, instead of normative judgments they make *shm*ormative judgments. When all we have are the tautologies of the constitutivists, we are trying to ground normativity on "a few

---

[8] I will return to a more precise formulation of this norm below, wherein I will explicitly address Street's appeals to idealization (which I allude to in (A) with the word "aim").
[9] Enoch 2006; 2011. His main targets are Korsgaard and Velleman, and he in fact does not apply his objection directly to Street. But the basic idea is easily transposable.

concepts which are entirely abstract, wholly insubstantial, and likewise floating about entirely in air" (Schopenhauer [1840] 1995, 62).

So beyond (A), constitutivists need something analogous to the following:

(B) We, as humans, *can't help* but take ourselves to have 'reasons' *in the precise, constitutively loaded sense specified by (A)*.

The point is that constitutivists cannot settle for simply proposing tautologies: they must show that we are *bound* by these tautologies; that these tautologies (inevitably) *apply to us*.[10]   Armed with (B), we can point out to Jill that she *can't* take herself to have shmeasons: she's stuck with reasons, hence stuck with the possibility of making normative errors.[11]   So if she can justify the claim to inevitability captured in (B), Street's account of normative error will go through.

II.2: The Possibility of Shmagency

How does Street justify (B)? Somewhat surprisingly, she does not directly address the need for this claim (or the shmagency worry from which it arises) other than with a few quick references in her more recent work. Her response is that if the account of constitutive norms "is thin enough to be plausible… then [the constitutivist] will be able to make it clear why questions about one's reasons to play the valuing game [i.e. use *reasons*, not shmeasons] are ill-formed…" (2010, 374). Given Street's use of the term 'ill-formed,' perhaps the thought is that talk of 'shmeasons' is conceptually flawed. This charge has some *prima facie* plausibility, as we are perfectly familiar with *reasons* (and the attendant attitudes of valuing, normative judgment, etc.), but it is not at all clear what a 'shmeason' is even supposed to be.[12]   Following this line, one way to justify (B) would be to show that *no* being could (genuinely) take itself to have shmeasons, since the very concept is incoherent.

In response I will argue that the concept is perfectly coherent and that we can very well imagine a shmagent (i.e. a being who is motivated only by shmeasons). To get clearer on the attitude of taking oneself to have a shmeason, it will help to look closer at how Street explains the attitude of taking oneself to have a *reason*:

---

[10] It is worth noting that other constitutivists openly recognize the need for claims analogous to (B), though they're defenses are not directly applicable Street's minimalist account (nor, I believe, are they effective in defending their own accounts, though I cannot enter this discussion here). See Korsgaard 1996, 236; Ferrero 2009, 304; Velleman 2009, 136-7.

[11] In fact, it may not be this simple, as Enoch has argued that even a claim to inevitability like (B) does not settle the question (see his 2006, 186-92; 2011, 215-7). However it is not entirely clear to me that his arguments are effective or avoid begging the question against the constitutivist. As such, I will focus on arguing against (B). But bear in mind that even if my arguments fail, constitutivists like Street are not necessarily home free.

[12] Ferrero raises this worry in his 2009, 311-2, though he does not think the concept is certainly incoherent. Enoch is aware of the issue (2011, 214-5), but his treatment of it is far from satisfying: he simply offers a brief analogy instead of an actual account of shmagency.

Our independent understanding of the attitude of taking something to be a reason is supplied by two main things. It's supplied first of all… by our understanding of *what it is like* to have a certain unreflective experience. And it's supplied second of all by our recognition of what is constitutively involved in the attitude of judging something to be a reason…. (2008, 241)

We touched on this "unreflective experience" above: it is the experience of a raw motivational push, or "one thing's seeming to *demand* or *call for* or *count in favor of* something else" (240). So a normative judgment is an attitude that has raw motivational push, and which entails being further motivated in accordance with certain constitutive norms. *Prima facie,* we can simply specify the concept of a shmeason by separating the motivational push from the constitutive norms of reasons: a shmeason is an attitude that has raw motivational push, and yet has *no* such constitutive entailments.

    Though this seems clear enough, Street will likely object that I have merely performed a linguistic manipulation, and have not successfully specified a new concept. The claim will be that, though we can talk about both (1) the 'raw motivational push' and (2) the 'constitutive entailments' of normative judgments, (1) and (2) are not truly separable concepts. We can illustrate this claim by examining Street's analogy with color. In describing the experience of taking something to be a reason, she contends, "Just as the experience of color cannot be adequately described except by invoking color concepts, so the type of [normative] experience in question… cannot adequately be described except by invoking normative concepts" (2008, 240). If I may extend the analogy, while we can describe the color *red* by referring to both (3) a 'raw redness phenomenology' and (4) the fact that 'red is more orange than green,' such verbal separation is not truly informative. This is because (3) and (4) are inextricably connected: we can't imagine a phenomenology of redness that is *not* more orange than green. Likewise, so the argument goes, though we can *verbally* separate out (1) the 'raw motivational push' from (2) the 'constitutive entailments' of normative judgments, (1) and (2) are in fact inseparable pieces of the same phenomenology.

    In response, I contend that the color analogy does not hold: we *can* gain independent traction on the notion of 'raw motivational push' without invoking constitutive norms. In fact, Street herself seems to commit to this separability when she distinguishes the attitude of normative judgment from that of *desire*. In so doing she agrees that both are motivational attitudes, but highlights that "normative judgments are different from desires in virtue of the kinds of constitutive involvements I've been sketching" (2008, 230). While normative judgments necessarily track means-ends consistency, it is perfectly sensible to *desire* to live, know that the only way to live is to have your limb amputated, and yet have *no* desire for such an amputation.[13] Further, this difference in

---

[13] Street 2008, 230-1. This understanding of desires is the simple one involved in (say) desiring a cookie; it does not capture the broader concept of a "pro-attitude" that is sometimes meant by the term 'desire'.

constitutive entailments is the *only* difference between desires and normative judgments that Street points out. So it appears that we *can* understand 'raw motivational push' independently of the constitutive norms of reasons: we need simply look to desires.[14] And we can now use this understanding to define a *shmormative judgment* (i.e. taking oneself to have a *shmeason*) as an attitude that has raw motivational push, and yet does not follow the robust constitutive norms that Street identifies. As one brutally obvious example, we could simply define 'shmormative judgment' as a *desire*.[15] Certainly we can coherently and informatively describe a being that is only motivated by desires; under this definition of 'shmormative judgment,' such a being would be our shmagent. Given this specification, I contend that Street cannot justify (B) by appealing to the conceptual impossibility of shmagency.

II.3: An Appeal to Phenomenology
Having established the possibility of shmagents, Street can most plausibly justify (B) with a more particular claim about *us,* as *humans*. This remains a fairly plausible line of justification: though we can perfectly well *imagine* a shmagent, I have done nothing to indicate that *we* can be shmagents. Street can still claim that as humans, we are the types of beings who inevitably make *nor*mative judgments about our *rea*sons – we simply *cannot* make shmormative judgments about shmeasons. Continuing with my above definition of shmeasons as desires, the claim is that humans are inevitably motivated by attitudes other than desires, namely by reasons. Unfortunately Street does not take an explicit stand on this issue, so it is not entirely clear how she would justify this claim. However, there is an obvious option available to her (and perhaps to constitutivists more generally): she can appeal to our *phenomenology* and the patterns that our motivations *actually* tend to follow. The claim will be that (some of) our motivations inevitably track means-ends consistency (at least when in full awareness), so we cannot make only shmormative judgments.

　　　Now in explicating this phenomenological claim we have to be careful not to beg any questions on either side (which unfortunately is a bit tricky, due to the tendency of constitutivists to turn ordinary concepts into terms of art). Clearly humans are motivated by something very similar to Street's "normative judgments," and I admit that it is plausible to say that this is an attitude distinct from desires. But we cannot explain such motivation with her vocabulary

_____

[14] Street may contend that the motivational push of desires is different from that of reasons, and thus my specification of 'shmeasons' fails. However, (i) at the very least she needs to do more work to justify this phenomenological claim, and (ii) in order to define 'shmeasons' we do not need the phenomenology to be precisely the same as that of reasons; we merely need to be able to pick out a *motivational* attitude that does not follow her constitutive entailments.

[15] I believe we could further flesh out the class of motivational attitudes by examining attitudes that combine raw motivational push with other entailments that are less robust than reasons but more robust than desires. One example of such an intermediate attitude between desires and normative judgments may be the "proto" forms of valuing felt by lower animals that Street discusses in her 2006, 119.

because that would beg the question in favor of the constitutivists (recall that the debate is over (B): whether we must inevitably take ourselves to have *reasons* as Street defines them, or just some attitude that appears quite similar). I believe we can gain traction on this issue by using an example. Take an instance of someone who forms the judgment "I have conclusive reason to go to Rome" and *isolate the phenomenology that this statement picks out* (without reading into the fact that we use the term 'reason' to convey such phenomenology). Now whatever the best way to classify this phenomenology, it certainly has a raw motivational push: a certain *counting-in-favor-ness* about it. So refer to this raw feeling as CFN.

Given this terminology, I suggest that the best way for Street to try to justify (B) is to appeal to facts about our experience of CFN. Continuing with her example: imagine a scenario in which you experience CFN for going to Rome immediately, and you find out that getting on a plane is a necessary means for doing so. It seems clear that – as long as you are fully aware of the situation – you will inevitably experience CFN for getting on a plane (or else stop feeling CFN for going to Rome). Generalizing this example, the phenomenological claim that Street is after is that humans' experience of CFN manifests itself along the patterns that are constitutively entailed by normative judgments.[16] This is a plausible claim: it appears to be a raw fact about our phenomenology that (when in full awareness of all the objects we feel CFN for and all the non-normative facts[17]) we experience the motivation of CFN in concert with means-ends consistency. I am going to grant this phenomenological claim to Street, so call it the *phenomenological fact*.

What is the significance of this phenomenological fact? It is a phenomenon that requires explanation. And *here* is where we garner support for Street's account and for (B), as she can execute an inference to the best explanation. She can contend that the best way to explain the inevitability of the phenomenological fact is by saying that we inevitably make *nor*mative judgments and feel the motivational push of *rea*sons.[18] Thus our CFN tracks means-ends consistency because we *can't help* but use these concepts, and so we can't help but make the conceptual commitments that are constitutively entailed from them. Given this IBE, the phenomenological fact appears to offer a powerful justification of (B), as it supplies us with an answer to the blunt shmagency objection. It turns out that we simply cannot be shmagents, since we inevitably experience CFN for some objects, and we cannot subsequently fail to be motivated to perform actions that are necessary to obtain these objects.

---

[16] Note that these patterns of manifestation are what separate CFN from simple desires. The question of the next section will be whether the fact that CFN follows these patterns is enough to justify Street's appeal to constitutive norms.

[17] These idealizations are obviously necessary to make this claim true, and they will be important momentarily.

[18] This has to be an IBE, not a deductive inference, because (as I will argue in the next section) the phenomenological fact is compatible with other explanations and so does not *entail* the inevitability of Street's constitutive norms.

Further, I contend that Street will be forced to make some move analogous to this IBE from our phenomenology. The possibility of shmagency requires her to appeal to some particular fact about *us*. And as the constitutive norms are supposed to govern our attitudes, it is unclear what could be appealed to other than our *experience* of these attitudes. Further, recalling that Street explains our knowledge of the concept of a reason in part by appealing to our knowledge of a "certain unreflective *experience*" (2008, 241; my emphasis) this phenomenological justification appears to cohere with her approach to the issue.

III. Alternative Explanations

Allow me to briefly lay out the dialectical situation as it currently stands. Street's account of normative error (hence the plausibility of her entire proposal) rests on her constitutivism, which itself depends on claim (B) – that humans inevitably make *nor*mative judgments, and thus make the attendant conceptual commitments. Claim (B), in turn, is justified by an IBE from the phenomenological fact about our experience of CFN. Given this structure, we can call her entire account into question by proffering *alternative* explanations of the phenomenological fact. I will put forward two such explanations, both of which fail to generate normative error.

But before going into these alternatives, we can illustrate how exactly there is room for an argument here by re-examining the analogy with assertions. Recall how Street's account of normative error is supposed to work. Feeling CFN is purportedly analogous to a situation in which a person asserts '*P&Q*', and then we merely point out that he is committed (by his own standards) to *Q*. However, now that claim (B) is justified by an appeal to phenomenology, the situation is different. It is instead as if a person is making certain noises, and then we (i) *classify* these noises as an assertion that '*P&Q*', and then (ii) point out that he is committed to *Q*. Step (ii) appears to be airtight, but there is room to cause trouble at step (i), which is analogous to the IBE from the phenomenological fact to the inevitability of normative judgments.

III.1: Underdetermination of Constitutive Norms

With this possibility in mind, turn to my first alternative explanation. I contend that even if our CFN is bound by *some* set of constitutive norms, the phenomenological fact underdetermines the correct account of these norms. It is compatible not only with Street's proposal (i.e. the norms of 'reasons'), but also with norms that fail to generate normative error. To see why this is so, recall Street's exposition of such error.

 [S]imply by judging yourself to have reason to *Y*, you're *thereby* – as a constitutive matter – also judging yourself to have reason to take the means to *Y*, whatever those may be. So even if you don't know that *Z* is a means to *Y*, and think that you have no reason whatsoever to *Z*, you *do* have reason to *Z* – *according to you.* (2008, 229)

Translating her talk of 'judging yourself to have a reason' into 'having CFN', we can see (from her first sentence) that this explanation rests on the following constitutive norm:

(C1) By having CFN for Y, you thereby have CFN for all Z that are necessary means to Y, whatever they may be.

The question I want to raise is whether (C1) can be justified by an appeal to the phenomenological fact. So suppose, *arguendo*, that some constitutive norm is the proper explanation of this fact. The phenomenological fact states that if you feel CFN for X, you will inevitably feel CFN for the necessary means to X *assuming full awareness* of one's other attitudes of CFN and all the non-normative facts. This translates directly into the following constitutive norm:

(C) By having CFN for Y, you thereby – *in conditions of full awareness* – have CFN for all Z that are necessary means to Y.

And note that (C) is *not* equivalent to (C1), since the latter makes no mention of the condition of full awareness. How can Street move from (C) to the norm she needs, i.e. (C1)? If we follow the line I am now suggesting, she must draw an inference to the best explanation from the phenomenological fact captured by (C). This explanation has some plausibility. In a situation in which an individual lacks full awareness, (C) describes what his *counterfactual* motivation would look like when given more information. And (C1) looks like a simple way to explain this counterfactual motivation: he feels CFN for Z when made aware of its connection to Y because he was aiming to feel CFN for such necessary means all along. In fact, it may appear that (C1) more or less *follows* from (C): what else could explain this counterfactual motivation?

There are other explanations, however. Consider the following norm:

(C2) By having CFN for Y, you thereby have CFN for all Z that *you take to be* necessary means to Y.

There are two important points we can draw from this alternative. First, (C2) implies (C) just like (C1) does. In fact, under conditions of full awareness (C1) and (C2) are extensionally equivalent: when you are fully aware, what you *take* to be a necessary means to Y simply *is* a necessary means to Y (that's what it means to be 'fully aware'). Second, *(C2) fails to generate normative error.* Street gets error out of (C1) by appealing to the commitments we necessarily make, whether or not we are aware of them. (C2), in contrast, only creates commitments when we are *aware* of such implications. According to (C2), we cannot make instrumental mistakes by our own commitments when we lack awareness, because in the absence of such awareness we *have no* such commitments.

Having (C2) on the table shows that even if (C) is true and the phenomenological fact is governed by *some* constitutive norms, it is not immediately clear that these are the norms of 'reasons.' (C) could be made true by either (C1) or (C2), so for Street's account of normative error to succeed, we need to be given some basis to pick the former over the latter. And it is clear

that no appeal to our phenomenology will settle this issue, as (C2) explains this phenomenology equally well. I do not wish to push this alternative too strongly: surely there are things that could be said in favor of (C1) over (C2). My point is merely that this is an argument that has to be made. So at the very least, Street's account needs further justification of its specific formulation of the constitutive norms if it is to proffer an account of normative error.

III.2: Mere Psychology

I now turn to my second, more pressing alternative to Street's constitutive norms. Recall that the phenomenological fact states that (in full awareness) we experience patterns of CFN in concert with means-ends consistency. Street's constitutivism rests on the claim that the best explanation of such patterns is that we inevitably use concepts (reasons) that entail conceptual commitments to means-ends consistency. Thus she looks at our brute phenomenological patterns and *adds* her claim to conceptual commitments to explain these patterns. My basic objection to this move is that we can offer a simpler, evolutionary-psychological explanation of these patterns that makes no reference to constitutive norms or their conceptual commitments. Thus instead of our experience of CFN being governed by constitutive norms, it follows *mere psychological regularities* that fail to generate normative error.
Recall (an instance of) the phenomenological fact we are trying to explain:
(PF1) Suppose you feel CFN for getting to Rome immediately. In such a situation, if you (in full awareness) learn that getting on a plane is a necessary means to do so, then you will inevitably feel CFN for getting on a plane.

Street's constitutivism rests on an inference from the robustness of the regularity captured in (PF1) to a claim that our CFN is conceptually tied to the aim of means-ends consistency. By feeling CFN for Rome, we have *conceptual* commitments such that we would be making a *mistake* if we did not feel CFN for getting on a plane (when lacking full awareness). With this in mind, consider another fact about our phenomenology:
(PF2) Suppose you see that a boulder is about to fall on you. In such a situation, you will inevitably feel CFN for jumping out of the way.

(PF2) is also a robust regularity, which captures our tendency to feel CFN for survival. Given this robustness, can we infer from (PF2) that our[19] CFN is *conceptually* tied to the aim of surviving? By being a happy human, would you be making a *mistake by your own commitments* if you simply did not feel CFN for survival? I take it that such an inference from (PF2) would be implausible: if we could so easily draw conceptual commitments out of phenomenological regularities, we would be well on our way to a substantive version of

---

[19] Note that here and in (PF1) I talk about the concepts behind *our* experience of CFN. Obviously CFN *simpliciter* does not include these conceptual commitments: we can imagine both shmagents (contra (PF1)) and agents who feel CFN for destruction (contra (PF2)).

constitutivism like Korsgaard's (which Street expressly wants to avoid: 2008, 244). And surely this cannot be right: robust psychological regularities are insufficient to ground conceptual claims. Even if every (functioning) human has an aversion to the smell of feces, this does not mean that the concept 'feces' constitutively entails that humans are averse to its smell. As Joyce puts it, "if you tell me that Jane saw a hungry-looking tiger coming towards her, then I'd expect her to run away – but that doesn't make it a conceptual claim that agents run away from tigers" (2001, 63).

So we are not intuitively comfortable with using (PF2) to infer any conceptual commitments of our CFN. But since such a move would seemingly be parallel to Street's inference from (PF1) to conceptual commitments, her justification of claim (B) is on shaky ground.[20] If we can (1) explain why the inference from (PF2) to constitutive norms is unacceptable, and (2) marshal parallel worries for Street's IBE from (PF1), her account will be in trouble. To this end, I submit that the problem with using (PF2) to infer the existence of conceptual commitments is that we can offer a fully adequate explanation of (PF2) without any reference to such cognitively loaded concepts. Instead of illustrating our conceptual commitments, (PF2) is simply a psychological regularity.

As it turns out, Street is the one who gives us the kind of evolutionary explanation I have in mind.[21] She uses this as an argument against an inference to *normative* facts from our experience of CFN, but it can likewise be used against an inference to *constitutive* facts. To begin: suppose two creatures, *A* and *B*, come into existence. *A* feels CFN for its own survival, whereas *B* feels CFN for its own destruction. Obviously (*ceteris paribus*) *A* is the one that survives and propagates, whereas *B* dies off rather quickly. And we can offer an entirely satisfactory explanation of this fact without making any sort of appeal to normative truth. As Street puts it, "The first creature *survived*, of course, but this isn't because its judgment was *true*, but rather simply because that creature tended to do what promoted its survival" (2008, 221). Now she admits that this may be unintuitive for some, but she offers a clever retort:

Those with strong realist intuitions may wish to insist that the second creature was making some kind of mistake… But if you're tempted to think this, ask yourself *why* you're tempted to think this… The answer is not that [creature *A*] was *right*, but rather that you're *his* descendant. (221)

---

[20] Some may contend that there is a disanalogy, as (PF1) is (purportedly) a much more robust phenomenological fact than (PF2). Two points in response: (1) just as Street builds in idealization to full awareness to make (PF1) robust, we could build in idealizations to *happiness* to make (PF2) robust. But though (seemingly) all *actual* happy humans feel CFN for survival, surely we can imagine a perfectly happy human that does not do so, and therefore we still would not want (PF2) to get us to a constitutive norm. (2) Though I have referred to (PF1) as a 'phenomenological fact,' we could certainly question how robust it is by looking to examples of people who are depressed, cognitively impaired, or simply irrational. At the very least, it is far from obvious where the disanalogy between the two is supposed to come.

[21] 2008, 221-3. I translate her talk 'valuing' into 'having CFN.'

Thus Street contends that our tendency to feel CFN for survival (as instanced by (PF2)) does not tell us anything about normative truth; the objects we feel CFN for are ultimately normatively arbitrary. And if (PF2) does not tell us anything about what we are *normatively* required to feel CFN for, then *a fortiori* it does not tell us what we are *conceptually* required to feel CFN for. Just as an evolutionary explanation makes claims to normative truth unnecessary, it likewise makes claims to conceptual commitments unnecessary: (PF2) is merely an evolutionarily advantageous psychological regularity.

Street's explanation gives a simple, naturalistically appealing explanation of (PF2) that undercuts the (unnecessary) inference to conceptual commitments. And, as it turns out, we can offer a directly analogous explanation of (PF1). So suppose that two creatures, C and D, come into existence. They both feel some CFN, but it manifests itself along different patterns in each of them. C feels CFN *only* for the prospect of its own survival: when it thinks about survival it feels CFN, and yet this motivation simply does not translate into motivation to take the necessary means to its survival. It is perfectly possible for C to feel CFN for survival, recognize the boulder, and yet feel absolutely no CFN to jump out of the way. In contrast, D feels CFN for survival *and* whatever it recognizes to be the necessary means to survival. This is *not* because it makes use of the concept of a 'reason' and recognizes its conceptual commitments; these patterns of CFN are simply a brute psychological regularity (just as feeling CFN for survival is a brute psychological regularity). Of course, *ceteris paribus*, D survives and C does not. This much is fairly straightforward, but the point of drawing it out is that it allows us to raise points that are precisely parallel to those raised in Street's explanation (compare the following to the quotes above):

D is the one that *survived*, of course, but this is not because its judgments about what it should do were *true*, but rather simply because its motivational structure tended to promote its survival.

Those with strong *constitutivist* intuitions may wish to insist that the first creature was making some sort of *conceptual mistake*. But if you are tempted to think this, ask yourself *why* you are tempted to think this. The answer is not that D was *right*, but rather that you are *his* descendant.

Thus I submit that we can explain (PF1) (and all other instances of the phenomenological fact) as merely an evolutionarily advantageous psychological regularity: there is no need to posit Street's conceptual commitments. And since Street (rightfully) holds that (PF2) fails to establish the existence of conceptual commitments, she seemingly must admit the same of (PF1). So at the very least the burden has been shifted: we are owed some further justification for the jump from the phenomenological fact to Street's constitutive norms. As it stands, this fact appears to represent a mere psychological regularity, and so it fails to "set

up standards" in the way needed to vindicate Street's account of normative error. [22]

To see why this is so, return to our analogy with assertions in which John asserts '*P&Q*.' Since asserting *'P&Q' entails* asserting *'Q'* (that's just what it *means*), if John goes on to assert *'¬Q,'* then he has clearly made a mistake on his own terms. But suppose the scenario is different. John* is simply the type of creature who has a brute psychological regularity: he tends to make the sounds, "*pee-and-cue,*" and when made aware of certain further facts he always goes on to make the sound, "*cue.*" But he is *not asserting* '*P&Q*' (suppose he does not speak the relevant language); he is just making noises. If John* makes the noise "*pee-and-cue,*" and goes on to make the noise "*not-cue,*" has he made an *error* on his own terms? Of course not. This kind of raw psychological regularity makes it *likely* that he will make such noises, but he does not count as making a *mistake* if he does not do so. Translating back to the question at issue, I contend that the situation with our experience of CFN is parallel to that of John*. We have a *robust psychological regularity* to feel CFN along certain patterns when made aware of certain further facts. But that is all. And just as John*'s mere psychological regularity fails to set up standards for him, *our* mere psychological regularities fail to set up standards for us. He is perfectly free to make such noises as they come; we are perfectly free to feel such motivations as they come. *This* is the best explanation of the phenomenological fact, for it puts forward a purely naturalistic account and does not add any unnecessary (normative or constitutivist) claims on top of that account. [23]

IV. Constitutivism and Naturalist Explanations

Street proposes an interesting account of metaethical constructivism in which our actual normative reasons are constructed out of the constitutive entailments of our own judgments about reasons. Though I find much of her approach appealing, I have argued that her use of constitutive norms to generate normative error is deeply problematic. If any such norms are to be efficacious, we need an explanation of why they inevitably apply to us. And given the conceptual possibility of shmagents, this explanation must come from facts about us *as humans*. However, the obvious (phenomenological) facts that can be appealed to do not proffer such an explanation, as they both (1) underdetermine the correct account of constitutive norms, and (2) fail to show that *any* account of constitutive norms is necessary: these facts are perfectly

---

[22] Why is the failure to justify an appeal to conceptual commitments more intuitively obvious with (PF2) than (PF1)? It seems to me that this is due to the fact that the *objects* of our motivations are so clearly selected for (as Street brings out), whereas it is not as obvious that the *structure* of our motivations (i.e. the patterns in which we feel CFN) is likewise contingent and selected for.

[23] Could Street respond by taking a minimalist understanding of "setting up standards" such that what it *means* to "make a mistake by our own lights" is simply that we would be motivated differently under certain idealized counterfactual conditions? I contend that such a move would undercut the significance of her account, as it would rest on a mere semantic shift in the meaning of "normative error" and "mistake."

compatible with an evolutionary account of mere psychological regularities. Thus I contend that, at the very least, Street needs to offer further justification for her claims about constitutive norms.

Having focused directly on Street's version of constitutivism, one may wonder whether my arguments have any implications for constitutivism more broadly. I believe that they do. First, the mere fact that I have revealed problems for Street's *minimalist* account of constitutivism is *prima facie* reason to think that more robust accounts (such as those offered by Korsgaard or Velleman) will run into problems that are at least as serious. Second, I believe I have shown that constitutivists cannot fully avoid the worries raised by Enoch by simply making their accounts of constitutive norms 'thin' enough. All such claims about the constitutive entailments of certain concepts must be grounded in a claim that we as humans *inevitably* make use of these concepts in the precise sense the constitutivist uses them. And finally, my evolutionary argument suggests that such claims to inevitability can *never* be justified by invoking them as the best explanation for our patterns of behavior or phenomenology. As long as we accept a naturalistic worldview, there will always be a complete explanation of these patterns that makes no reference to the concepts and classifications of the constitutivists. Put in slogan form: we are merely pieces of the physical world, and the physical world does what it does irrespective of the "folk-theoretical names"[24] we paint onto it. Thus the constitutivist will have to invoke other justifications for their claims to inevitability, and it is far from clear what such justifications would look like.

Works Cited:

1. Enoch, David. 2006. "Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Action." *Philosophical Review* 115, (2): 169-198. DOI: 10.1215/00318108-2005-014.

2. ——. 2009. "Can there be a global, interesting, coherent constructivism about practical reason?" *Philosophical Explorations* 12, (3): 319-339. DOI: 10.1080/13869790903067683

3. ——. 2011. "Shmagency Revisited." In *New Waves in Metaethics*, edited by Michael Brady, 208-233. United Kingdom: Palgrave Macmillan.

4. Ferrero, Luca. 2009. "Constitutivism and the Inescapability of Agency." In *Oxford Studies in Metaethics: Volume 4*, edited by Russ Shaffer-Landau, 303-331. New York: Oxford University Press.

5. Joyce, Richard. 2001. *The Myth of Morality*. New York: Cambridge University Press.

---

[24] Lewis 1996, cited in Enoch 2006, 169.

6. Korsgaard, Christine M. 1996. *The Sources of Normativity*. New York: Cambridge University Press.

7. Lewis, David. 1996. "Desire as Belief II." *Mind* 105: 303-13. Reprinted in *Papers in Ethics and Social Philosophy*, 55-67. New York: Cambridge University Press.

8. Schopenhauer, Arthur. [1840] 1995. *On the Basis of Morality*, edited by David E. Cartwright. Indianapolis: Hackett.

9. Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127, (1): 109-166. http://www.jstor.org/stable/4321684

10. ——. 2008. "Constructivism about Reason." In *Oxford Studies in Metaethics,* vol. 3, edited by Russ Shaffer-Landau, 207-245. New York: Oxford University Press.

11. ——. 2009. "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues* 19: 273-298. DOI: 10.1111/j.1533-6077.2009.00170.x.

12. ——. 2010. "What is Constructivism in Ethics and Metaethics?" *Philosophy Compass* 5, (5): 363-384. DOI: 10.1111/j.1747-9991.2009.00280.x.

13. ——. 2011. "Evolution and the Normativity of Epistemic Reasons." In *Belief and Agency*, edited by David Hunter, 213-245. Calgary: University of Calgary Press.

14. ——. n.d. "Objectivity and Truth: You'd Better Rethink it." Unpublished manuscript, https://files.nyu.edu/jrs477/public/Sharon%20Street%20-%20Objectivity%20and%20Truth.pdf

15. Velleman, David J. 2000. "On the Aim of Belief." In *The Possibility of Practical Reason*, 244-281. New York: Oxford University Press.

16. ——. 2009. *How We Get Along*. New York: Cambridge University Press.

On the Indexicality of 'know'

Gabriel Larivière
McGill University

Abstract:
This paper argues that epistemic contextualism (EC) is true, that is, that the English predicate 'know' is context-sensitive. It is divided into three sections: I first present what epistemic contextualism is and give a *prima facie* reason why it might be true. Second, I respond to the gradability objection to (EC) and I give a second reason why (EC) might be true *and* an argument based on conceptual relations between 'know' and 'justification' in favor of it. Third, I respond to the error-theory objection and conclude that (EC) is true. [1]

1. Epistemic Contextualism (EC)

(EC) is the view that the English verb 'know' is indexical, that it changes its content in different contexts of use. For contextualists, 'know' has what Kaplan calls an unstable character, that is, a set of rules that determines a different content in different contexts of utterance.[2] Granted compositionality, a sentence of the form 'x knows p' will express different propositions in different contexts of assertion. Hence, the content of 'know' is a function of the ascriber's context and not of the subject's context: in some contexts, 'x knows p' is more difficult

---

[2] (Kaplan 1989).

to satisfy because epistemic standards are higher. In short, 'know' behaves similarly to the gradable adjectives 'is flat' or 'is tall'.

Part of the attractiveness of (EC) is that it provides a solution to the following skeptical argument (SA):

(i) If x knows she has a hand, then x knows she is not a brain in a vat.
(ii) It is not the case that x knows she is not a brain in a vat.
(iii) Therefore, x does not know she has a hand.

But what is more obvious than (iv) x knows she has a hand? {SA}[3] and (iv) taken altogether form an inconsistent set of equally plausible propositions, i.e. a paradox. A solution to any paradox of that kind consists in saying which of these propositions is false and why that proposition appeared to be true.[4] (EC)'s solution to this paradox consists in saying that 'know' is more difficult to satisfy in high standards contexts; hence, in those contexts, (SA) is a sound argument and (iv) is false. In a low standards context, 'know' is easier to satisfy which makes (SA) unsound and (iv) true. Moreover, the reason why each of these sentences appeared plausible is that competent speakers are sometimes unaware of the context-sensitivity of 'know'. Hence, (EC) provides a solution that indicates which proposition (in a given context) is false and that explains why {SA} and (iv) appeared to be inconsistent. Thus, (EC) could be said to give a "happy-face solution" to the above paradox.

(EC) is supported by many every day 'knowledge'-ascriptions. It is compatible with our intuitions about the truth-values of everyday 'knowledge'-claims in cases such as the following.

ZOO

Suppose Tom is at the zoo with his classroom and, pointing at the zebras, asks his teacher "What are these animals?" The teacher, after looking at the animals in the cage, answers that they are zebras. Tom then asserts

(1) "She knows they are zebras."

Suppose, further, that two post-modernist artists are looking at the scene and have the bizarre idea that these animals could in fact be cleverly painted mules. Pointing at Tom's teacher, one post-modernist artist says

(2) "She does not know they are zebras."

---

[3] I shall use '{SA}' to refer to the set containing the sentences (i), (ii) and (iii), as opposed to '(SA)' to refer to this skeptical argument. Note that if (EC) is true, these sentences have no content if they are not used in a context.

[4] The idea of a paradox and its resolution can be found in (Schiffer 1996, 317-318) and in (DeRose 1995, 3).

In ZOO, intuition says that both (1) and (2) are true. (EC) can easily explain that intuition by saying that the epistemic standards are higher in the artists' context than in Tom's, which makes '*x* knows *p*' more difficult to satisfy in their context. Thus, because 'know' is context-sensitive, the propositions expressed by Artist 2 and Tom are both true and they do not form a contradiction. The fact that (EC) is compatible with our intuition in ZOO (and in many other similar every-day cases) provides a *prima facie* reason to think that it is true.

In the proceeding sections, I shall consider two objections to (EC): the gradability objection and the error-theory objection. The gradability objection states that 'know' has no unarticulated semantic link to a scale because it is not syntactically gradable like gradable adjectives. Thus, the analogy between 'know' and 'is flat' or 'is tall' is misguided and it is very unlikely that 'know' is indexical. In the next three sections, I will argue that the gradability objection fails because 'know' is limitedly gradable and because it allows different syntactic constructions to express its unarticulated semantic link to a scale. I will also give an argument in favor of (EC) based on conceptual relations between 'know' and 'justification'. Next, I will discuss the error-theory objection, i.e. that competent speakers do not fail to recognize indexicals. If (EC) were true, competent speakers would not fail to recognize 'know' as being context-sensitive. But unfortunately for (EC), not only do competent speakers fail to realize that 'know' is context-sensitive but also well-trained epistemologists and linguists argue that 'know' is not context-sensitive, so 'know' is not an indexical. I will respond to that objection by showing that competent speakers can fail to realize that gradable adjectives are context-sensitive, which greatly diminishes the threat posed by semantic blindness.

2. The Gradability Objection

Contextualists typically use an analogy between 'know' and gradable adjectives, e.g. 'is flat', 'is tall', 'is rich', 'is bald', etc, to explain its context-sensitivity.[5] About gradable adjectives, many linguists hold that

(a) they map their argument onto abstract representation of measurements, or degrees;
(b) the set of ordered degrees according to a certain dimension to which they map their arguments is a scale.[6]

For example, 'is tall' is a function from the subset of individuals that have a height to degrees of height. In lambda notation:
(3) $[|\text{is tall}|] = \lfloor x.^{\text{TM}}_{\text{height}}(x)$[7]

---

[5] E.g. (Cohen, 1999, 60).

[6] (Kennedy, 2007, 4).

[7] '$[|\text{is tall}|]$' is to be read as 'the predicate 'is tall'', and '$\lfloor x.^{\text{TM}}_{\text{height}}(x)$' is to be read as 'the function that takes an individual with a certain height as argument and that maps it onto a degree of height'. This notation is used by Kennedy (2007).

A typical compositional implementation of (a) and (b) is that a sentence containing at least one gradable adjective can express different contents in different contexts of use because the standards of comparison vary in different contexts. By generalization, in any context C, 'Michael is tall' is a function whose extension (a truth-value) is determined by an equivalence between the appropriate standard of comparison of height in C ($d_{height-C}$) and Michael's height, i.e. between the minimal degree in C for satisfying 'is tall' and by Michael's tallness.

(4) $[|\text{Michael is tall}|] = \lfloor x. ^{™}\text{height}(m) \geq d_{height-C}$

Thus, the sentence (5) asserted in C can be properly analyzed as (6).

(5) Michael is tall.
(6) $\geq(^{™}\text{height}(m), d_{height-C})$

(EC) claims that 'know' has an unarticulated semantic link to a scale of epistemic strength[8], similarly to gradable adjectives, which makes it indexical.

(7) $[|x \text{ knows } p|] = \lfloor x \lfloor p.\text{Believes}(x, p) \text{ and Istrue}(p) \text{ and } ^{™}\text{epistemic-strength}(x, p) \geq d_{epistemic-strength-C}$

Stanley objected that if a word has an unarticulated semantic scale, then syntactic constructions should make possible the expression of that link to a scale.[9] If the analogy between gradable adjectives and 'know' holds, one would expect 'know''s syntax to be the same as the syntax of gradable adjectives. But because 'know' is not *gradable* – it does not seem to allow for modifiers and for natural comparative constructions – it is not semantically linked to a scale and thus not indexical.

MODIFIERS – (8)[10]

---

[8] I use the term 'epistemic strength' as a broad notion encompassing most of the different forms of (EC). As an example, some contextualists argue that 'know' is linked to a scale of justification (Cohen, 1999). Also, I set aside the question if 'know' is linked to a scalar scale or a partial ordering, like 'is interesting'.

[9] (Stanley, 2004).

[10] 'Really', 'hardly' and 'perfectly' are disputed cases of modifiers: some linguists see them more as "slack-regulators" – that they do not affect a statement's truth conditions but indicates that they are to be taken *literally* – or "hedges" – not to be taken literally (like 'so to speak' and 'more or less') (Lasersohn 1999).

Acknowledging that 'really' is not a modifier solves some problems faced by (EC). Conee (2005) argues that 'know' does not seem to accept truth-conditions modification from "really and truly". According to Conee, 'really' and 'truly' function as intensifiers, as in "a really hot day" and "a truly long drive", but when combined with

> x is very flat / tall / rich / bald
> x is quite flat / tall / rich / bald
> x is extremely flat / tall / rich / bald

MODIFIERS – (9)
> * x very knows that p
> * x quite knows that p
> * x extremely knows that p

NATURAL COMPARATIVE CONSTRUCTIONS – (10)
> x is flatter / taller / richer / balder than y
> x is the flattest / tallest / richest / baldest

NATURAL COMPARATIVE CONSTRUCTIONS – (11)
> * x knows p more than y
> * x knows p the most

Clearly, 'know' is not gradable. Thus, the analogy with gradable adjective seems to fall apart and doubts are cast on the core contextualist idea that 'know' has an

---

'know', they behave more as ways to get serious about the truth of certain knowledge claims.

> (1_{10}) I know Obama is in Mexico this week-end.
> (2_{10}) Do you *really* know?
> (3_{10}) Well, ok, I think so but I don't *really* know it.
> (4_{10}) So you don't know it.

Conee correctly identifies that the use of 'really' allows the second speaker to "get serious about the truth" of (1_{10}), which is revealed to be false, i.e. the truth of (4_{10}) entails the negation of (1_{10}). Thus, 'really' does not modify the content of 'know' but behaves as a slack-regulator. The problem for Conee is that 'really' also seems to be a slack-regulator with gradable adjectives, so his argument fails. The fact that 'really' is not a modifier is supported by the oddity the following statement.

> (5_{10})   (a) * He is tall but he is not really tall.
> (b) He is tall but he is not very/extremely/(?) quite tall.

(Some interviewed speakers were puzzled by 'quite' in (5_{10} – b) but some others were not; all of them found (5_{10} – a) contradictory).

Cohen (2005) argues that (1_{10}) and (4_{10}) were both true, i.e. that 'really' introduces a change in the epistemic standards. Although 'really' can introduce such change, it clearly does not do so all the times and it does not change the content of the predicate 'know'. Note that Stanley (2004, 125) thinks (5_{10} – a) is perfectly fine.

unarticulated semantic link to a scale. Stanley also uses two gradable verbs to make (EC) look more suspicious.

LIKE – (12)
> x very likes y
> x quite likes y
> x likes y more than z

REGRET – (13)
> x very regrets p
> x quite regrets p
> x regrets p more than y / more than q

Stanley concludes that it is very unlikely (EC) is semantically linked to an unarticulated scale if it is not gradable. His argument could be summarized by the two following principles – the weak *Syntactical Imperative* (SI) and the strong *Gradable Imperative* (GI):

(SI) IF an expression e has an unarticulated semantic link to a scale s, THEN some syntactic constructions can express this semantic link.

(GI) IF an expression e has an unarticulated semantic link to a scale s, THEN e is gradable along s.[11]

Because 'know' fails (GI), Stanley concludes that it has no unarticulated link to a scale and it is not context-sensitive. Note that for Stanley, the equivalence (SI) ⇔ (GI) holds.[12]

2.1. 'Know' is Limitedly Gradable
A response to Stanley's argument can go two ways: either argue that (GI) and (SI) ⇔ (GI) are false, or argue that 'know' is gradable in some weaker ways. In this paper, I argue for both: I shall first start by the latter.
Contrary to what Stanley thinks, 'know' is modifiable by the qualitative modifiers 'very well' and 'well'. Here are interesting cases where 'know' behaves like 'regret'.

---

[11] This formulation of (GI) is taken from (Blome-Tillmann 2009, 43).

[12] Even if 'know' is not semantically linked to a scale, (EC) might still be true. For example, Halliday (2007) suggested the analogy between 'know' and gradable adjectives should be replaced by an analogy with modified comparatives such as 'is sufficiently tall'. Since those modified comparatives are not linked to an unarticulated semantic scale but are still context-sensitive a context's standards, 'know' does not need to be linked to an unarticulated semantic scale for (EC) to be true. I shall leave aside Halliday's response and consider Stanley's objection as posing a serious threat to (EC).

(14)   (a) I regret very much[13] that you still do not understand what it is that has moved us in this matter. (Letter from Kennedy to Khrushchev, during the Cuban missile crisis)
(b) I [know very well] that you still do not understand what it is that has moved us in this matter.

(15)   (a) I know very well that all these public assurances were false and that your military people had set out recently to establish a set of missile bases in Cuba.[14] (Same letter)
(b) I [regret very much] that all these public assurances were false and that your military people had set out recently to establish a set of missile bases in Cuba.

(14) and (15) are cases where the modifiers 'very much' unquestionably modifies 'regret'. (14 – a) expresses *significant* regrets Kennedy felt vis-à-vis Khrushchev's behavior during the crisis, and 'very much' semantically qualifies 'regret'; it is difficult to see why (14 – b) would be a different case, where the modifier 'very well' would not semantically qualify 'know'. The same reasoning applies with (15), but reversely: if 'very well' plays a pragmatic role in (a), then 'very much' plays a pragmatic role in (b), which seems implausible. Here is another convincing case where 'very well' modifies 'know':

(16)   I know very well that the easy thing to do, particularly for a president preparing for reelection, is to avoid any controversy. (Obama, 2011)

If one is not convinced of the effect of 'very well' on 'know', one can compare (16) with (16)*, that is (16) *without* 'very well', and appreciate the modifiability of 'know'.
        Stanley argues 'very well' is a pragmatic indicator when combined with 'know' instead of a modifier, as 'very much' in

(17)   2 is very much a prime number.[15]

Pragmatic indicators like 'very much' in (17) do not play any semantic role, but are merely part of speech acts. They are not sign of any genuine modifiability. As a response to Stanley, first, although 'very much' is a pragmatic indicator in (17), it does not prevent it to be a modifier with 'regret'. In fact, clearly, some other modifiers can behave as pragmatic indicators.

---

[13] I take as unproblematic the fact the 'very much' is used after 'regret' in that sentence. First, because this letter is a real letter written by J. F. Kennedy; second, because "I regret very much" typed in Google gave 929 000 entries ("I very much regret" gave 1 620 000 entries).

[14] I modified the original latter, where, interestingly enough, "know very well" was replaced by "learned beyond doubt".

[15] (Stanley 2004, 126).

(18)   Burundi is in South America, isn't it?
        Not quite!
(19)   Did you have a good sleep?
        Yes, I had a very good sleep!

Cross-linguistic data in French support the view that modifiers can also be pragmatic indicators.

(18)$_{fr}$ Le Burundi est en Amérique du sud, n'est-ce pas?
        Ce n'est pas tout-à-fait cela!
(19)$_{fr}$ As-tu fait une bonne sieste?
        Oui, j'ai fait une très bonne sieste![16]

 So 'very well' could be a pragmatic indicator in certain cases and a modifier in some other cases.
Second, there are genuine cases very similar to (15), (16) and (17) where 'very well' is a modifier. If 'very well' is a modifier in the cases below, it is hard to see why it would not be a modifier with 'know'.

(20)   The new mayor's speech was very well received.
(21)   I remember very well that he gave me his phone number last evening.
(22)   Greek politicians understand very well that they will have to observe international lenders requirement. (Newspaper The Baltic Course)

Third, 'very well' does not seem to be a modifier *prima facie*. It is ironic that Stanley gives an example of 'very much' being a pragmatic indicator to explain how 'very well' can be a pragmatic indicator: are there such cases? Consider the unnaturalness of these:

(23)   (a) 2 is very much an even number.
(b) * 2 is very well an even number.
(24)[17]   (a) I very much believe/suggest that she will come.
(b) * I very well believe/suggest that she will come.[18]

---

[16] 'tout-à-fait' and 'très' can be modifiers.

> (1$_{24}$)   Il est très grand.
> (He is very tall.)
> (2$_{24}$)   Il est tout-à-fait chauve.
> (He is quite bald.)

[17] This example is taken from (Dutant 2007, 6). This paper also argues that 'know' is gradable with the modifiers 'very well / well'.
[18] Again, cross-linguistic data in French are evidence that 'very well' is not a pragmatic indication at all. 'well' is *strongly* equivalent to 'bien' – 'well-being' = 'bien-être'; 'I'm doing well' = 'Je vais bien'; 'Well, let's go' = 'Bien, allons-y'. Thus, 'very well' is equivalent to 'très bien'; it is also very unlikely that 'très bien' is a pragmatic indicator.

That 'very well' is a genuine modifier is also revealed by the fact that following syntactic constructions are permitted.

(25)   He understands/remembers/sees/knows *p* but he does not understand/remember/see/know it very well.

The correctness of (25) suggests that there is a semantic difference (in any context) between "he understands/remembers/sees/knows *p*" and "he understands/remembers/sees/knows *p* very well". Thus, 'very well' is a genuine modifier with most factive verbs (including 'know' as (25) suggests) and I see no reason why it would be a pragmatic indicator with 'know' and not with 'understand/remember/see'.
Provided that *some* modifiers can also be pragmatic indicators; that (20), (21) and (22) are cases where 'very well' seems to be a genuine modifier; and that it is doubtful if 'very well' can ever be a pragmatic indicator, why would 'very well' not be a modifier with 'know' – as strongly suggested in (14), (15) and (16)? There seems to be no grounds to support the argument that 'very well' is a pragmatic indicator with 'know'.
        But Stanley gives other reasons why 'very well' does not modify 'know': both make weird combinations in negative assertions and in questions.[19]

(26)   (a) I don't like Bush very much.
(b) * I don't know very well that Bush is president.
(27)   (a) Do you like Bush very much?
(b) * Do you know very well that Bush is president?

But these do not clearly point at the fact that 'very well' is not modifying 'know'. They seem to indicate a more general phenomenon since similar patterns can be observed with other gradable verbs that can take sentential complement:

        (28)   (a) * I don't understand very well that Bush has to go to war.
                (b) * Do you understand very well that Bush has to go to war?
(29)   (a) * I don't see very well that Bush is sitting over there.
                (b) * Do you see very well that Bush is sitting over there?[20]

Also, as Dutant points out, pragmatic integrator such as 'really' are allowed in those construction; it might be the case that only a pragmatic role is allowed to

---

[19] (Stanley 2004, 126).
[20] It might be argued that 'understand' is a borderline case. Nonetheless, other examples of gradable verbs that do not permit modifiers in negative assertions and questions are 'hear', 'hope' and 'remember'. However, other verbs such as 'regret' do permit modifications in those contexts.

adverbs in the above syntactic constructions with 'know', 'understand' and 'see'.[21] Moreover, 'know' fits naturally in embedded 'how well'-questions.

(30)   How well do you know that she is not your enemy?

Therefore, this objection does not do much wrong to the convincing argument above that 'know' is modifiable by 'very well'. Moreover, that 'know' is modifiable gives very good reasons to think it has an unarticulated semantic link to a scale, and a *second* reason to think that it is context-sensitive. To sum up this section, the fact that 'know' is modifiable by 'very well' casts significant doubts on Stanley's objection that 'know' is not gradable and hence not indexical. Of course, 'know' has what we could call a *limited gradability* compared to gradable adjectives; but still, it is far from clear that 'know' fails (GI) and thus far from clear that 'know' has no semantic link to an unarticulated scale.

2.2. 'Know' in Comparative Constructions
As shown above, 'know' is not naturally comparable; moreover, '*x* knows *p* better than *y*' is idiomatic, so it does not support its gradability. It is very likely that an expression that has a semantic link to an unarticulated scale allows some comparative constructions to express its link to a scale. Thus, rejecting (SI) seems implausible and the contextualist needs to find non-idiomatic syntactic comparative constructions with 'know'. As Blome-Tillman argues, 'know' is adverbially modifiable *and* allows adverbial comparative uses.[22]

ADVERBIAL MODIFICATION AND COMPARATIVE USES – (31)
        x knows *p* with very / quite good evidence / justification
        x knows *p* with greater certainty than *y*
        x knows *p* with the best justification possible

So (SI) is saved but (GI) fails. Blome-Tillmann rejects the too strict (GI) and rather proposes the *Modifiability Constraint* on the grounds that other verbs such as 'snore'[23] have a semantic link to an unarticulated scale but are not gradable.

(MC) If an expression *e* has an unarticulated semantic link to a scale *s*, then *e* is either gradable or adverbially modifiable along *s*.

---

[21] (Dutant, 2007).
[22] (Blome-Tillmann 2009, 46).
[23] Other verbs seem to be context-sensitive and semantically linked to an unarticulated scale but not gradable: 'dream' [the French equivalent 'rêver' is gradable along a qualitative scale of details a subject remembers about his dreams, e.g. "J'ai beaucoup rêvé la nuit dernière" (* "I much dreamed last night") and "J'ai plus rêvé cette semaine que la semaine dernière" ("I dreamed more this weak than last weak")]; 'wait' would be another one.

This replacement of (GI) by (MC) is highly plausible and supported by much linguistic evidence, which makes Stanley's bi-conditional (SI) ⇔ (GI) false.[24] Moreover, (31) is good evidence that 'know' is adverbially modifiable along a scale of epistemic strength. Hence, the contextualist has found a satisfactory answer to Stanley's objection.

2.3. The Conceptual Relation Argument or the COHEN Argument
However, the existence of adverbial comparative uses for 'know' – as compared to natural comparative uses – entails that 'know' can be used comparatively along a scale, but of course *not that 'know' is semantically linked to an unarticulated scale*; from the existence of adverbial comparative uses of 'know', the indexicality of 'know' does not follow. Here is one of many examples of adverbial comparative uses of non-indexical factive verbs.

PROVE – (32)
> *x* proved *p* quite rapidly
> *x* proved *p* with greater artistic skills than *y*
> *x* proved *p* with the best pencil available

Obviously, being adverbially comparable is not sign of having an unarticulated semantic scale, since a verb can be adverbially compared to an incredibly surprising number of scales. It is quite clear that Blome-Tillmann is not arguing such a link, but on the other hand, being naturally comparable *is* sign of possessing such scale.[25] In that sense, (MC) – and (SI) – looks more as a triviality than as evidence that 'know' is context-sensitive, and Stanley would be correct in pointing out that (31) is not sign of 'know' having a link to an unarticulated semantic scale.
On the other hand, because 'know' is adverbially modifiable by 'justification', an argument in favor of 'know' having a semantic link to an unarticulated scale can be made. This argument builds on conceptual relations between both expressions similar to some remarks made by Cohen.[26] First, notice that 'wait', 'snore' and 'know' are all respectively adverbially modifiable by the expressions 'a very long time', 'very loud' and 'with very good justification'.
        Second, the expressions used in (31) are different than the ones used in (32). I here want to make a distinction between the word 'justification', a justification-relation and the concept of *justification*. 'with very good justification' is a linguistic item that qualifies the knowledge-relation in which

---

[24] I am unsure if (SI) ⇔ (MC) holds.

[25] I take 'gradable' to mean "modifiable and naturally comparable". Stanley points out that some expressions such as 'taller than six feet' are modifiable but not context-sensitive (2004, 132). But 'taller than six feet' is not naturally comparable and hence not gradable as I defined 'gradable'. In this paper, I take that if an expression is gradable (as I defined 'gradable'), then it has an unarticulated semantic to a scale (I can see of no cases where it does not).

[26] (Cohen 1999, 60).

an individual stands vis-à-vis a proposition *p*. It qualifies the knowledge-relation by expressing something about the justification-relation in which an individual stands vis-à-vis *p*. So 'with very good justification' in a sentence *s* expresses something about a justification-relation in a context *C* (if *s* is true in *C*). However, speakers can also use the word 'justification' to talk not only about justification-relations but also about the concept of *justification*. Someone could say that *justification* is analyzable in terms of *evidence*. The simple fact that the word 'justification' can be used in these two different ways is evidence for the claim that there is a distinction between a justification-relation and the concept of *justification*. I make the obviously plausible assumption that there are semantic and syntactic relations between a word and its concept, i.e. between 'justification' and *justification*. A speaker who uses the word 'justification' in a sentence *s* to express a proposition about a certain justification-relation, also *uses* the concept of *justification* as a tool in order to be understood by other people, i.e. someone who understands the meaning of *s* knows what the concept of *justification* is. Similarly, someone who uses the word 'knowledge' to express a proposition about a knowledge-relation also *uses* the concept of *knowledge* in order to be understood by other people. Of course, the word 'use' in '*x* uses a concept' is used metaphorically, maybe in a different sense than the word 'use' in '*x* uses a word'. But this metaphorical use is sufficient and unproblematic for the picture I want to get at: a speaker qualifying a knowledge-relation with the words 'with very good justification' *uses* the concept of *justification* to qualify that knowledge-relation. A speaker qualifying a proof with the words 'with great artistic skills' uses the concept of *artistic skills* to qualify that proof. The expression 'with very good justification' in (31) is thus different than the expressions 'with great artistic skills' and 'more rapidly' in (32) because the concept of *justification* is a core concept of the concept of *knowledge*, while *artistic skills* and *rapidity* are not core concepts of the concept of *proof*. Similarly, *loudness* is a core concept of *snoring*, and *time* is a core concept of *waiting*. Hence, (31) is different than (32) in that the verb 'know' is qualified and modified by a core concept of *knowledge*, while the verb 'prove' is not qualified and modified by a core concept of *proving*. More formally, I take '*f* is a core concept of *g*' to mean the following:

DEFINITION OF CORE CONCEPT
f is a core concept of *g* IFF when a speaker *uses* *g* to assert something about an object *a* standing in a *g*-relation, that speaker also implicitly or explicitly asserts something about *a* standing in a *f*-relation.

That *justification*, *noise* and *time* are respectively core concepts of *knowledge*, *snoring* and *waiting* is made manifest by the following.

(33)   (a) *x* knows that *p* with very good justification
(b) ?? *x* knows that *p* without any justification
(34)   (a) *x* snores very loudly
(b) ?? *x* snores without any noise

(35)   (a) *x* waited for a very long time
(b) ?? *x* waited 0 second
(36)   (a) *x* proved *p* with great artistic skills
           (b) *x* proved *p* without any artistic skills

(33 – b), (34 – b) and (35 – b) illustrate conceptual mistakes, not syntactic ones. They are evidence of *justification*, *noise* and *time* being core concepts of *knowledge*, *snoring* and *waiting*. The distinction these highlight is a conceptual relation. I shall define this kind of conceptual relation the following:

DEFINITION OF CONCEPTUAL RELATION
Two expressions f and g are conceptually related IFF *f* (or *g*) is a core concept of *g* (or *f*) AND there is a *natural semantic* implication between f and g like the followings:
x snores $\Rightarrow$ *x* makes some noise (scale of loudness)
x waits $\Rightarrow$ there is a period of time *t* during which *x* expects something (scale of time)
x knows that *p* $\Rightarrow$ *x* is justified in believing *p*.[27]

We could say from the above that the concepts of *knowledge*, *snoring* and *waiting* are all naturally semantically linked to other concepts: *snoring* to *loudness*; *waiting* to *time*; *knowledge* to *justification*.
Third, note that the expressions 'is justified', 'is loud' and 'has been expecting something for a period of time *t*' all have a (unarticulated *or* articulated) semantic link to a scale: 'is justified' has a semantic link to a scale of justification strength[28]; 'is loud' has a semantic link to a scale of loudness; 'has been expecting something for a period of time *t*' has a semantic link to a scale of time. Thus, because '*x* snores', '*x* waits' and '*x* knows *p*' are all conceptually related to 'loudness', 'time' and 'justification', they contain a positive conceptual links: '*x* snores' contains the conceptual link that the subject who snores can be mapped onto a positive degree of loudness; '*x* waits' contains the

---

[27] I understand 'justification' in its everyday common sense meaning, i.e. competent speakers *naturally* make these inferences, which explains why (33), (34) and (35) seem very odd. This is not an analysis of knowledge.

[28] As stated above, the fact that 'is justified' is modifiable is not sign of it having an unarticulated semantic link to a scale, as Stanley points out (2004, 132). Stanley thinks that 'is justified' is similar to 'is taller than six feet' and does not have such a link. Blome-Tillmann also suggests it in a footnote (2009, 46). However, I take that 'is justified' has such a link and here are two reasons why: first, 'is justified' is gradable as I defined 'gradable', i.e. it is modifiable and allows for natural constructions, while 'is taller than six feet' only allows for modifiability. Second, one can say "*x* is justified in believing *p* in everyday contexts but not in that context where he needs to be more justified". The fact that 'is justified' allows for these constructions, where *x* can be justified in believing *p* in some contexts but not in some others is strong evidence for 'is justified' having an unarticulated semantic link to a scale. *x* cannot be taller than six feet in some contexts but not in some other ones, although *x* can be much taller than six feet.

conceptual link that the subject who waits can be mapped onto a positive degree of time during which that subject expects something; '*x* knows *p*' contains the conceptual link that the relation between *x* and *p* can be mapped onto a positive degree on a scale of justification (and hence of epistemic strength).

POSITIVE CONCEPTUAL LINK

$[|snores|] = \lfloor x.Sleeps(x)$ and $^{™}_{loudness}(x)$

$[|waited\ for\ y|] = \lfloor x \lfloor y.Expected(x, y)$ and $^{™}_{time}(x, y)$

$[|knows\ p|] = \lfloor x \lfloor p.Believes(x, p)$ and $Istrue(p)$ and $^{™}_{epistemic\text{-}strength}(x, p)$

This positive conceptual link is trivial; invariantists who argue that '*x* knows *p*' is not context-sensitive hold that it is true iff $^{™}_{epistemic\text{-}strength}(x, p) \geq d_{epistemic\text{-}strength}$, where '$d_{epistemic\text{-}strength}$' is a degree of epistemic strength required in every contexts for a subject to know a proposition. The contested part is the context-sensitivity of 'know' simply expressed by adding the standard degree of comparison in *C*.

INDEXICALITY

$[|x\ snores|] = \lfloor x.Sleeps(x)$ and $^{™}_{loudness}(x) \geq d_{loudness\text{-}C}$

$[|x\ waited\ for\ y|] = \lfloor x \lfloor y.Expected(x, y)$ and $^{™}_{time}(x) \geq d_{time\text{-}C}$

$[|x\ knows\ p|] = \lfloor x \lfloor p.Believes(x, p)$ and $Istrue(p)$ and $^{™}_{epistemic\text{-}strength}(x, p) \geq d_{epistemic\text{-}strength\text{-}C}$

The context-sensitivity of 'snore' and 'wait' seems to follow simply from the positive conceptual link of '*x* snores' to 'loudness' and of '*x* waits' to 'time', who are both dimensions of scales; is 'know' also context-sensitive because of its positive conceptual link to 'justification'? I see no reason why it would *not* be the case. Let's say Joe truthfully asserts "I waited for the bus this morning" on a beautiful Monday morning in May, where the bus arrived 2 minutes late. That proposition might be false in another context, e.g. on a stormy Monday morning in January, where most buses would be 2 hours late because of a snowstorm. The only way to explain why that sentence is false in that context while Joe is mapped onto the same value on a scale of time is to acknowledge that the standard of comparison is different in the snowstorm-context, that '*x* waits' is indexical. Similarly, Joe could truthfully assert "I know that arithmetic is incomplete" because, let's say, he heard his introduction to logic teacher mentioning it in a lecture. That sentence would be false if uttered in an international philosophy of mathematics meeting. Again, the only way to explain why that sentence is false in that context while Joe is mapped onto the same value on a scale of epistemic strength is to acknowledge that the standard of comparison is different in the philosophy of mathematics meeting case, that '*x* knows *p*' is indexical. It is also worth noticing that 'waited' can be used in different sentences where the subject who waits is mapped to a scale of time that can account for significant variations.

(37)   Joe waited for the day he would become a chess master. (Asserted at a press conference about Joe's chess player career)
(38)   Joe waited for his computer to download a file. (Asserted by Joe's mother who was (also) waiting for Joe to come and help her with her next chess move)

Taking both sentences above to be true, in (37), because of the positive conceptual link contained in 'x waits', Joe is mapped onto a scale of time to a value that is obviously greater than in (38). Nonetheless, both sentences are true. If Joe had waited few minutes before he would become a chess master, (37) would be false. Hence, from the fact that 'x waits' is positively conceptually linked to 'time' and from the significant variations on a scale of time 'wait' allows, the context-sensitivity of 'wait' follows. If one accepts that (37) and (38) are genuine illustrations of context-sensitivity, then one has to accept that (39) and (40) are also genuine illustrations of context-sensitivity. Consider Joe again and his knowledge of the incompleteness of arithmetic.

(39)   Joe knows that arithmetic is incomplete. (Asserted at lunch time in the cafeteria)
(40)   Gödel knows that arithmetic is incomplete. (Asserted at a philosophy of mathematics meeting)

Because 'x knows p' is positively conceptually linked to 'justification', in (40), Gödel is mapped onto a scale of justification to a value that is obviously greater than the value to which Joe is mapped in (39). Nonetheless, both sentences are true. Hence, from the fact that 'x knows p' is positively conceptually linked to 'justification' and from the significant variations on a scale of epistemic strength 'know' allows, the context-sensitivity of 'know' follows.
To sum up this argument, here is the *Conceptual Relation principle* (CR):

(CR) If an expression e_1 is adverbially modifiable with an expression e_2,
AND e_1 is conceptually related to e_2,
AND e_2 has a (unarticulated *or* articulated) semantic link to a scale s_2;
THEN e_1 has a (unarticulated *or* articulated) semantic link to a scale s_x.

If (CR) is true, then 'know' has an unarticulated semantic link to a scale s and thus 'know' is context-sensitive. I take (CR) to be plausible, although the notion of 'conceptually related' can surely be challenged. Note that the argument above differs from and does not entail Cohen's claim that 'know' comes in degrees because 'justification' comes in degrees.[29] (CR) does not make it necessary that 'know' is linked to the same scale as 'is justified' is linked to. Also, because (CR) is based on observations between 'knowledge' and 'justification', it could be called the COHEN argument, as Stanley names it. For Stanley, The COHEN

---

[29] (Cohen 1999, 60).

argument for (EC) is based on the *Metalinguistic Knowledge-Justification thesis* (MKJ):

(MKJ) For any context C, the word 'know' expresses a relation that, relative to that context, contains as a component the property expressed by the word 'justified', relative to C.[30]

However, (CR) is different than (MKJ) for two reasons: first, it stresses out the importance of adverbial modification; second, it is generalizable. However, one can easily notice the same general idea between (CR) and (MKJ). I take (CR) to be a strong linguistic formalized principle that entails (MKJ).

In sum, the gradability objection fails for a second reason: (GI) is too strong and must be rejected, while 'know' clearly passes the (SI) test as well as the (MC) test. Moreover, (CR) provides an argument that 'know' has an unarticulated semantic link to a scale of epistemic strength. Finally, the analogy between gradable adjectives and 'know' is still valid because 'know''s syntax allows it to

(a) map his argument onto abstract representation of epistemic strength;
(b) and this (partial) order of epistemic strength is a scale.

3. Error-Theory Objection
As said above, another strength of (EC) is that it provides a solution to the following paradox:

(i) If x knows she has a hand, then x knows she is not a brain in a vat.
(ii) It is not the case that x knows she is not a brain in a vat.
(iii) Therefore, x does not know she has a hand.
Moreover
(iv) x knows she has a hand.

(EC)'s solution to {SA} and (iv) consists in saying that in a high standards context, (SA) is a sound argument and (iv) is false; in a low standards context, (SA) is unsound and (iv) is true. Moreover, the reason why each of these sentences appeared plausible is that competent speakers are sometimes unaware of the context-sensitivity of 'know'. Thus, (EC) provides a solution that indicates which proposition (in a given context) is false and that explains why {SA} and (iv) appeared to be inconsistent.

Some philosophers have found this error-theory implausible. Schiffer, for example, said that it was implausible that "those uttering knowledge sentences are both referring, unbeknown to themselves, to different knowledge

---

[30] (Stanley 2004, 133). Here is the COHEN argument:
      1. Gradable expressions are context-sensitive.
      2. 'Justified' is thus context-sensitive.
      3. 'x knows p' means in part what is meant by 'p is justified for x'
      Conclusion: The truth of 'x knows p' depends upon the context.

relations and confounding the knowledge relations to which they're unknowingly referring."[31] He gives the example of a man in London, on a rainy day, saying

(41)   "It is raining".

It is absurd that that the man would mistakenly assert the proposition that it is raining in Oxford. Moreover, if the man mistakenly expressed the proposition that it is raining in Oxford, then it follows that he did not understand (41) at all, which is even more absurd. Schiffer's argument is based on what we could call the *Competent Use of Indexicals principles* (CUI). Let's *s* be a sentence containing at least one indexical and let *p* be the proposition expressed by *s* in *C*.

(CUI$_i$) IF a competent speaker *x* asserts *s* in *C*, THEN *x* knows in *C* that she is asserting *p*[32]

(CUI$_{ii}$) IF (CUI$_i$) fails, THEN there is no proposition *q* such that *x* believes in *C* that she is asserting *q* by asserting *s*.

Schiffer uses (CUI$_i$) and (CUI$_{ii}$) to reject (EC) by *reductio ad absurdum*. If (EC) is true, then his error theory is needed to explain the above paradox. But by (CUI$_i$) and (CUI$_{ii}$), competent speakers do not fail to recognize the indexicality of 'know' and (EC)'s error theory is false, so (EC) is also false. Therefore, (EC) must be false.

---

[31] (Schiffer 1996, 327).

[32] Michael Blome-Tillmann made the comment that this principle is false in cases where a speaker is under hypnosis. Cases where the competent speaker is drunk, under minimally powerful drugs, extremely tired, in a psychotic crisis or in any other unusual psychological states, count as cases where this principle fails – although they might arguably be cases where (CUI$_{ii}$) does not. Nonetheless, Schiffer could simply add a condition to his principle, something like:

> (CUI$_i$)* IF a competent speaker in *C x* asserts *s* in *C* THEN *x* knows in *C* that she is asserting *p*.

Specifying the notion of "competent speaker in *C*" might more or less avoid these counterexamples. Still, there are some other more subtle counterexamples to (CUI$_i$)*: a competent speaker could confuse the two predicates 'is bold' and 'is bald', the latter being obviously indexical. Thus, in a philosophical debate, he could qualify a friend's argument as being a 'bald statement', which is clearly a violation of both (CUI$_i$)* and (CUI$_{ii}$). Many other counterexamples can be built following similar ideas: at this point, it seems that Schiffer's principle can probably not be *completely* fixed. But clearly, Schiffer's argument remains: how can competent speakers be blind to the indexicality of 'know'?

Schiffer's argument can be interpreted as making two claims. First, 'know' is not context-sensitive because *ordinary speakers* fail to see that they assert different propositions in different contexts with the same sentence. Second, 'know' is not context-sensitive because *well-trained philosophers* fail to see that the set {SA} and (iv) is not inconsistent. I shall first address the former.

First, it seems that not all ordinary speakers fail to see that 'know' is context-sensitive. The ZOO case illustrates common sense intuitions that both Tom and Artist 2 speak truly when they respectively assert "*x* knows *p*" and "*x* does not know *p*". Some of them might even have the intuition that Tom and Artist 2 do not contradict each other, so they must express different propositions.[33]

Moreover, there are counterexamples to (CUI_i): ordinary speakers can fail to realize the indexicality of certain gradable adjectives.[34] Imagine the following scenario.

BOYFRIEND
Suppose Catherine and Megan are discussing how their perfect boyfriend would look like.
(Catherine) "My perfect boyfriend would be tall like Johnny."
(Megan) "But Johnny isn't tall."
(Catherine) "Yes, Johnny is tall!"
(Megan) "No, Sidney is tall, but Johnny is not tall!"
(Catherine) "No no no, Sidney is a giant and Johnny is tall!"
(Megan) "Are you blind? Sidney is not a giant and Johnny is not tall!"
…

Clearly, in BOYFRIEND, Catherine and Megan fail to see that 'is tall' is context-sensitive and that they are not contradicting each other. Thus, (CUI_i) fails. Moreover, BOYFRIEND is clearly a case where both Megan and

---

[33] In everyday language, it is common to say things such as "I know *p*, but Agatha knows it better (or with greater certainty)". Thus, ordinary speakers, because they have the intuition that Tom and Artist 2 do not contradict themselves, might respond that the teacher knows they are zebras, but someone who knew they were zebras and not cleverly painted mules would know better (or with greater certainty). It seems plausible that ordinary speakers would intuitively say Tom and Artist 2 express different propositions. This point remains untouched by Stanley's correct observation that "knows better" is idiomatic. The fact that speakers might use this idiomatic expression suggests that they feel a need to clarify what they mean by 'know', i.e. propositions containing the verb 'know' might have different content. Out of 10 non-epistemologists interviewed, 7 of them had the intuition that Tom and the artist were both speaking truly and not contradicting each other; 2 had the intuition that only Tom was speaking truly; and 1 had the intuition that only the artist was speaking truly.

[34] For some philosophers who do not think that 'is tall' and 'is flat' are context-sensitive, see Unger (1975). For Unger, 'is flat' does not have different truth-conditions in different contexts, but its assertability-conditions vary.

Catherine believe they are asserting a certain proposition, so (CUI$_{ii}$) also fails. Schiffer might want to weaken his two principles by inserting modality.

(CUI$_i$)$_{modal}$ IF a competent speaker x asserts s in C, THEN x could know in C that she is asserting p.

(CUI$_{ii}$)$_{modal}$ IF (CUI$_i$)$_{modal}$ fails, THEN there is no proposition q such that x believes in C she is asserting q by asserting s.

Here is BOYFRIEND$_{modal}$.

BOYFRIEND$_{modal}$
Suppose Catherine and Megan are reading two magazines about how the perfect boyfriend should be. Suppose that Catherine takes what is in the magazine A to be the most certain and indubitable thing in the world and that the same thing applies for Megan and magazine B.
(Catherine) "My magazine says that the perfect boyfriend should be tall, that is, at least 1,80m high. This is perfect, because I like Johnny and Johnny is 1,82m tall."
(Megan) "But Johnny isn't tall, look, my magazine says the perfect boyfriend should be tall, that is, at least 2m high, like my Sidney."
(Catherine) "No, Johnny is tall! Your magazine is wrong."
(Megan) "No, Sidney is tall, but Johnny is not tall!"
(Catherine) "No no no, Sidney is a giant and Johnny is tall!"
(Megan) "Are you blind? Sidney is not a giant and Johnny is not tall!"
…

Both (CUI$_i$)$_{modal}$ and (CUI$_{ii}$)$_{modal}$ clearly fail in BOYFRIEND' and there seems to be no way to weaken (CUI$_i$)$_{modal}$. Thus, ordinary speakers might fail to recognize the indexicality of a gradable adjective used in a discussion and the first interpretation of Schiffer argument fails.

There are at least four complementary answers to the second interpretation, that 'know' is not context-sensitive because *well-trained philosophers* fail to see that the set {SA} and (iv) is not inconsistent. The first one is to repeat what has just been established, that if ordinary speakers might fail to recognize the indexicality of certain gradable adjectives, why not philosophers? BOYFRIEND$_{modal}$ could be turned into an argument between a Moorean furiously pointing at *Proof of an External World* and a Skeptic pointing at (SA), both unable to recognize the indexicality of 'know'. The second one is to recognize that 'know' is a special word not obviously indexical, that is, it is not easy to recognize that it changes its content in different contexts. Its limited gradability, as seen above, makes it more difficult to recognize its indexicality. Thirdly, 'know' is non-obviously context-sensitive because standards usually do not get very high. The lack of variations makes us unaware of this particular feature of 'know' so that, while doing epistemology, philosophers fail to see that they are raising standards to an extreme level that

greatly changes the content of 'know'. Fourthly, 'know' behaves similarly to *absolute* gradable adjectives, i.e. 'is flat', 'is empty', 'is full', etc.[35] These are gradable adjectives which value required in a given context for truthfully asserting that an object is flat (e.g.) is calculated in terms of a contextually fixed difference between a maximum or a minimum. In other words, in a certain context, a glass that can contain 50 mL of water needs to contain at least 45mL for the sentence 'That glass is full' to be true; in another context, that glass needs to contain at least 48 mL for 'That glass is full' to be true. Because absolute gradable adjectives contain an implicit reference to an absolute maximum, it is more difficult to recognize that there are indexical – ("No field is *really* flat, there is always at least one small, maybe microscopic, bump on it!").[36] Similarly, a contextualist could say that 'x knows p' is true iff x's evidence that p eliminates all possible worlds in which $\sim p$ holds – pssst! – except those properly ignored in C.[37] The absolute to which 'know' refers is that of x evidence that p eliminating *all* possible worlds in which $\sim p$ holds, although in most contexts, if not all, some possible worlds can be properly ignored – ("You never *really* know that p because you never eliminate all possibilities in which $\sim p$ holds!").

Therefore, Schiffer's objection is greatly diffused by the above counterexamples and by the explanations provided in the last paragraph. Thus, it fails to pose serious challenge to (EC). Considering the two reasons pointing at the context-sensitivity of 'know' given in the precedent sections; considering the (CR)-argument in favor of (EC); and rejecting the threats posed by the gradability and the error-theory objections, I conclude that the English predicate 'know' is indexical, that (EC) is true.

I want to make a final comment here. Someone could accept (EC) but argue that when philosophers are discussing epistemology, epistemic standards are necessarily higher, which automatically changes the content of 'know' in all philosophical contexts. One could argue that it is with *that* content of 'know' that philosophy is concerned, that the only relevant way in which philosophers should discuss what knowledge is and what can be known is in high standards, when 'x knows p' is more difficult to satisfy. Thus, (EC) might be true, but it is philosophically irrelevant and the paradox formed by {SA} and (iv) is not resolved because (SA) looks true but cannot be philosophically accepted. The answer to such critique based on a normative view of how philosophy should be conducted requires two clarification: first, if philosophy ought to be conducted only with one content of 'know' (when standards are very high), then philosophers should no longer make 'knowledge'-ascriptions and should no

---

[35] On a useful typology of gradable adjectives from which the category of 'absolute gradable adjectives' is taken, see (Kennedy and McNally, 2005).

[36] Interestingly enough, Unger (1975) makes an argument in favor of skepticism on the basis that no surface is *really* flat, which illustrates the idea that well-trained philosophers can fail to realize that absolute gradable adjectives are context-sensitive. In any case, I take that if 'is flat' is not context-sensitive, 'know' is not.

[37] (Lewis, 1996).

longer use the word 'know' as it is used in English. They ought to use a different word – let's call it 'knowledge*' – that has as no unstable Kaplanian character, i.e. that does not change its content in different contexts. 'x knows* p' can be false in some contexts where 'x knows p' is true, and if 'x knows* p' is true in one context, then it is true in every other contexts. Therefore, this critique is equivalent the claim that philosophers should not use the word 'know' but 'know*'. Second, it is false to claim that the paradox formed by {SA} and (iv) is unresolved by (EC): the one formed by {SA}* and (iv)*, where 'know' is replaced by 'know*', would be the one unresolved and the one with which philosophers would be concerned according to that critique – if it forms a paradox at all. Beyond these clarifications, I do not see any other answers to that critique. The claim that epistemology should be concerned with 'knowledge*' rather than with 'knowledge' is very strange – as if all philosophers mistakenly thought that epistemology was about 'knowledge' – and there is no obvious or intuitive good reason why this is or should be the case.

4. Conclusion

In sum, in this paper, I argued that 'know' is context-sensitive. First, I showed that (EC) is compatible with every day 'knowledge'-claims, which makes it *prima facie* plausible. Second, I rejected the gradability objection on the grounds that 'know' is limitedly gradable by 'very well' – which gave a second reason why (EC) might be true – and that it is adverbially comparable. I then built on the fact that 'know' is adverbially comparable with 'justification' to give an argument based on conceptual relations between to the two words in favor of (EC). Finally, I showed that the error-theory objection does not pose serious threat to (EC) because competent speakers can fail to recognize that gradable adjectives are context-sensitive.

Works Cited:

1. Michael Blome-Tillmann, "The Indexicality of 'Knowledge'," *Philosophical Studies* 138 (2009): 29-53.

2. Stewart Cohen, "Contextualism, Skepticism, and the Structure of Reasons," *Philosophical Perspectives* 13 (1999): 57-89.

3. Stewart Cohen, "Contextualism Defended" and "Contextualism Defended Some More," in *Contemporary Debates in Epistemology*, ed. Steup and Sosa (Oxford: Blackwell Publishing, 2005), 56–62, 67–71.

4. Earl Conee, (2005). "Contextualism Contested" and "Contextualism Contested Some More," *Contemporary Debates in Epistemology*, ed. Steup and Sosa (Oxford: Blackwell Publishing, 2005), 47–56, 62–66.

5. Keith DeRose, "Solving the Skeptical Problem." *The Philosophical Review* 104 (1995): 1-52.

6. Julien Dutant, "The Limited Gradability of 'knows'," (forthcoming)

7. Daniel Halliday, "Contextualism, Comparatives and Gradability," *Philosophical Studies* 132-2 (2007): 381 – 393.

8. David Kaplan, "Demonstratives," in *Themes from Kaplan*, ed. J. Almog, J. Perry and H. Wettstein (Oxford/New York: OUP, 1989), 481-563.

9. Christopher Kennedy, "Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives," *Linguistics and Philosophy* 30 (2007):1.

10 Christopher Kennedy and Louise McNally, "Scale Structure, Degree Modification, and the Semantics of Gradable Predicates," *Language* 81-2 (2005): 345-381.

11. Peter Nathan Lasersohn, "Pragmatic halos," *Language* 75-3 (1999): 522–551.

12. David Lewis, "Elusive Knowledge," *Australian Journal of Philosophy* 74-4 (1996): 549-567.

13. Stephen Schiffer, "Contextualist Solutions to Scepticism," *Proceedings of the Aristotelian Society* 96 (1996): 317-333.

14. Jason Stanley, "On the Linguistic Basis for Contextualism," *Philosophical Studies* 119 (2004): 119-146.

15. Peter Unger, *Ignorance: A Case for Scepticism*. (Oxford: Clarendon, 1975).

On Presentism and Time Travel

Thomas Hall
Western Washington University

1. Introduction

Presentism is a thesis about temporal ontology according to which it is necessarily the case that only present objects exist.[1] Many philosophers have said that presentism and time travel are incompatible. For example, William Grey has said that the possibility of time travel would have to "presuppose that the past or future were somehow real," (1999: 56) and Steven Hales has said that "there is no such thing as time travel under presentism." (2010: 360) In addition, several philosophers have endorsed the view that time travel is possible only if eternalism (also called 'four-dimensionalism') is true – the view that past, present, and future objects all exist on an ontological par. William Godfrey-Smith, for example, has said that "the metaphysical picture which underlies time travel talk is that of the block universe," (1980: 72) and Ken Perszyk and Nicholas Smith agree that "For time travel to be possible, we need a different conception of time: four-dimensionalism." (2001: 4) This appears to be the common view.

The aim of this paper is to show that the common view isn't any good. Towards this end, I will evaluate two arguments that intend to show that there is

---

[1] There is some variation in how presentism is formulated. Some authors add that *it is always the case that* everything is present (e.g. see Crisp (2003: 16)), and others add that *the only properties and relations that present entities instantiate are those they currently instantiate* (e.g. see Sider (2005: 329)). Also, some authors do not (explicitly) formulate presentism as a necessary thesis. For defenses of presentism, see Crisp (2003), Markosian (2004), and Bourne (2006).

something about presentism *in and of itself* that rules out the possibility of time travel: The Nowhere Argument (§2) and The Suicide Machine Argument (§3). I will ultimately respond to these arguments on behalf of the presentist.[2]

2. The Nowhere Argument

Simon Keller and Michael Nelson dub the following argument 'The Nowhere Argument':[3]

On the presentist model, the past and the future do not exist, so there is nowhere for the time traveller to go. Travelling to Portland is possible, because Portland is right there waiting for you. But travelling to the Land of Oz is impossible, because there is no such place. Travelling to the past or future is more like travelling to the Land of Oz, if presentism is true. You can't travel to somewhere that doesn't exist, so, if presentism is true, you can't travel to other points in time. (2001: 334-5)

Initially the argument is quite compelling.[4] However, we should take a closer look:

> *The Nowhere Argument:*
>
> (P1) If presentism is true, then necessarily, the past and future do not exist
>
> (P2) It is impossible to travel somewhere that does not exist
>
> (C1) If presentism is true, then it is impossible to travel to the past or future
>
> (P3) If time travel is possible, then it is possible to travel to the past or future
>
> (C2) If presentism is true, time travel is impossible

The first premise goes analytic, so only (P2) and (P3) are up for grabs. The rationale for (P2) can be put as follows: *traveling* is a two-place relation, one that holds between a traveler and a destination. Given the general principle that no relation can hold without coexisting relata, it follows that no one can travel to a nonexistent destination. The rationale for (P3) is presumably intuitive: in order to travel in time, one must travel to the past or future.

Given the rationales above, the argument naturally progresses as follows. In cases of *time* travel, the destination relatum would be a concrete time. But, if presentism is true, there is only one time that is real – the now. As such,

---

[2] One noteworthy restriction: I will only be concerned with *closed-future* presentism – that is, presentism plus the view that future-tensed contingent statements are determinately true or determinately false. Kristie Miller has argued that time travel is incompatible with *open-future* presentism – that is, presentism plus the view that future-tensed contingent statements are neither determinately true nor determinately false. See her (2005) and (2008).

[3] This argument goes back (at least) to Grey (1999: 56-7).

[4] The argument is endorsed by: Grey (1999), Perszyk and Smith (2001), Eldridge-Smith (2007), among others.

presentism entails that traveling to the past or future is impossible because (i) no past or future times exist, and (ii) the *traveling* relation cannot hold with respect to the nonexistent.[5] From here, the second conclusion comes easy: if traveling in time requires traveling to the past or future, and presentists can't allow for such travel, then presentism and time travel are incompatible.

2.1. Against objections to the second premise

Unfortunately, many philosophers have focused their scrutiny on (P2) of The Nowhere Argument. The common way of fleshing out this objection is with a simple reductio: if (P2) is true, then the argument rules out too much – namely, the ordinary passage of time.[6] This is supposedly because the passage of time involves persisting objects traveling to the immediate future, which (according to the presentist) is just as nonexistent as the distant past. This is a bad objection. The presentist should not equate the passage of time with *traveling* to the immediate future. To see why, we need to understand what presentists say the passage of time consists in.

Presentism is a version of the *A-theory* of time – the view that tensed statements are irreducible, typically in the sense that tensed sentence tokens cannot be given tenseless truth conditions.[7] In order to express tensed truths, presentists traditionally utilize primitive tense operators (e.g. WAS, WILL) that attach to present-tensed sentences. For example:

> WAS (There are dinosaurs)

According to the presentist, 'There were dinosaurs' has the logical form of (1), which reads '*it was the case that* there are dinosaurs'. A-statements like (1) change their truth value over time. For example, (1) is true now, as it was a thousand years ago. But 100 million years ago, (1) was false. The passage of time for the presentist (*qua* A-theorist) consists in the change of such tensed truths – i.e. you eventually drop the 'WILL' for an instant, and then add a 'WAS'. Importantly, this phenomenon is quite different from traveling. Unlike the passage of time, traveling essentially consists in a change of relations to external objects.[8] The presentist can therefore reasonably say that the passage of time does not amount to anything like traveling. Furthermore, even if we don't choose to make this rebuttal, we can still insist that *time travel* is very different

---

[5] This kind of worry is all too familiar to presentists – dubbed 'the problem of cross-temporal relations'. There is a considerable amount literature on this topic. For some useful discussion, see Chisholm (1990), Sider (1999), Markosian (2004), Crisp (2005), and Bourne (2006: 95-108).

[6] For endorsements of this objection, see: Dowe (2000: 443), Keller and Nelson (2001: 335), Sider (2005: 329), and Daniels (2012: 472).

[7] For example, those that believe in the reducibility of tense (i.e. B-theorists) can say that an utterance of 'There were dinosaurs' expresses a truth if and only if there exists some time t such that (i) there are dinosaurs at t, and (ii) t is earlier than the time of utterance. For more on this detensing strategy, see Sider (2001).

[8] It is commonly thought that time could pass without such change. On this topic, see Shoemaker (1969).

from the ordinary passage of time. This is because time travel requires *leaving the ordinary flow*. After all, ordinary persistence does not make us into time travelers.[9]

There is another common way of fleshing out the objection to (P2). Some philosophers have said that one can travel to a nonexistent place, *as long as it exists upon arrival*.[10] Here is one such example involving spatial travel to a nonexistent location, due to John Carroll:

Suppose you are a big fan of amusement parks and hear that they are planning to build a new one in Argentina. It doesn't exist yet, but you are so excited that you start now to hitchhike your way there from Raleigh, NC. It seems that you are traveling to the amusement part even though it doesn't exist. What seems important is not that the destination exists when you start to travel, but that it exists when you arrive. (2008)

I don't find the example very plausible. A more accurate description of the relevant case would go like this: you start to travel to some *existing* place – a concrete location – which is such that it *will* be the case that when you arrive, *it* contains an amusement park. Given this response, we can say that your utterance of 'I'm traveling to the amusement park in Raleigh, NC' expresses a "quasi-truth"[11] because the following closely related facts obtain:

I'm traveling to Raleigh, NC, and

WILL (An amusement park exists in Raleigh, NC and I arrive there)

Importantly, the presentist's primitive tense operators are *ontologically noncommittal* – that is, the ontological commitments of all quantification and referring terms within their scope are negated.[12] Presentists can therefore endorse (3) without being committed to the existence of the amusement park. This response makes the case of spatial travel on offer inapplicable as a counterexample to (P2); it indicates that the case *does* involve an existing destination, and moreover, that you are (strictly-speaking) not traveling to the amusement park.[13]

---

[9] This is a bit too quick. If "suspended animation" cases (i.e. cases where an individual's physical processes are slowed or stopped via external means (such as freezing)) count as time travel, then it would be possible to persist normally while traveling in time. These cases are controversial, however, and so I have chosen to ignore them.

[10] For endorsements of this objection, see: Dowe (2000: 443), Miller (2005: 226), Carroll (2008), and Wasserman (manuscript: 38-9).

[11] Roughly, a *quasi-truth* is something appropriate to assent to in everyday circumstances, although not literally true. On the notion of quasi-truth, see Sider (1999) and Markosian (2004: 24-5). Alternatively, we could say that the relevant utterance expresses a truth *in ordinary English*, and seek to provide the metaphysical facts which underlie it in the form of what Sider (2011) calls a "metaphysical semantics".

[12] More familiar operators share this feature as well. For example, *Actualists* (philosophers who accept the view that everything is actual) traditionally take the modal operator 'POSSIBLY' to work in this way: e.g. the truth of 'POSSIBLY (There are talking donkeys)' is consistent with there being no talking donkeys.

[13] It may be possible to repair the spatial-analogy case. For example, imagine a case where X knows that (say) God is going to create a new closed spatial region. God's plan

Given that presentists think time is very different from space, they will likely want to resist the likening of time travel to spatial travel. *Contra* spatial travel, time travel cannot be a genuinely relational affair for the presentist. This is because presentists cannot (or at least should not) believe that it's possible for any genuine cross-temporal relation to hold, for the simple reason that *no* relation can hold without coexisting relata. As such, I think presentists should welcome the truth of (P2) with open arms.[14] However, now aware of the explanation for this concession, we should be skeptical of the third premise:

(P3) If time travel is possible, then it is possible to *travel* to the past or future
In order for the argument be valid, the sense of 'travel' at work here must be the same as that in (P2). This exposes a crucial bit of the rationale that wasn't initially apparent – namely, that *time travel* is genuinely relational. But this position, despite straightforwardly begging the question, has yet to be argued for. As such, if we can provide a coherent account of presentist time travel that makes no use of cross-temporal relations (and thus brings no commitment to nonpresent objects), this will give us a reason to reject (P3). And in fact, I think this can be done.

2.2. A presentist-friendly account of time travel
Keller, Nelson, and Ryan Wasserman all agree that presentists can translate traditional time travel talk into appropriately tensed language.[15] In this section, I follow them and show exactly how presentists can mirror the standard account of time travel articulated by David Lewis:[16]
What is time travel? Inevitably, it involves a discrepancy between time and time. Any traveler departs and then arrives at his destination; the time elapsed from

---

is to connect the new region to the finite open space that X inhabits. Given her knowledge, X begins to travel to the nonexistent space at t_1, arriving when it appears at t_2:
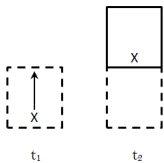
But is it true that X is traveling to the nonexistent closed space at t_1? I think not. Again, we can cite "in the ballpark" facts that suffice to make it quasi-true, like: (i) X is traveling to the "edge" of the open space at t_1, and (ii) WILL (there exists a closed space connected to the "edge" of the open space and X arrives there).

[14] At least insofar as 'travel' is being used in its ordinary relational sense. I fail to see how to motivate the premise otherwise, and so I think this is a fair assumption.
[15] See Keller and Nelson (2001), and Wasserman (manuscript).
[16] This account is endorsed (in some form or other) by Lewis (1976), Dowe (2000), Sider (2001) and (2005), Keller and Nelson (2001), Monton (2003), Richmond (2003), Arntzenius (2006), Hales (2010), Wasserman (manuscript), among many others.

departure to arrival (positive, or perhaps zero) is the duration of the journey. But if he is a time traveler, the separation in time between departure and arrival does not equal the duration of his journey. (1976: 145)

Lewis goes on to make a well-known distinction between *external time* (i.e. objective time)[17] and *personal time*, where an object's personal time is the assignment of coordinates to its stages that maintains the regularities and physical processes common to that kind of object.[18] The motivation for this distinction is easy to see. Before entering her time machine, it seems a pastward time traveler may appropriately utter 'In just a few seconds I will see dinosaurs', but in this case the traveler is really going millions of years into the past – not a few seconds into the future. With the relevant distinction in hand, this utterance becomes less paradoxical: we can say that just a few seconds of the traveler's *personal time* will elapse between her entering the time machine and arriving millions of years in the *external* past.

Recently, Wasserman has shown Lewis's account of personal time to be inadequate:

…consider the case of a single, non-time traveling electron. Suppose that some of its stages are labeled in order (1, 2, 3, etc.) according to external time. And suppose further that every electron remains intrinsically unchanged throughout its entire career. In that case, one can assign coordinates to the stages of our particular electron in many different ways and still preserve the kinds of regularities we ordinarily see in other electrons. For example, one can simply reverse the ordering of all the electron-stages. In that case, there would be discrepancy between the "personal" time of the electron and external time, but that would not make the electron a time traveler. (manuscript: 34)

To solve this problem, Wasserman proposes a revised account according to which "an object's personal time is the assignment of coordinates to its stages that matches the coordinates given by the *relevant causal relation*" – where 'relevant causal relation' picks out whatever causal relation makes for identity over time.[19] The background assumption for this account is the common belief that identity over time (for physical objects) requires the right kind of causal dependence between an object's stages.[20] Equipped with this important revision, the traditional account of time travel can be put as follows:

---

[17] Traditionally, the structure of external time is thought to mirror the structure of a single straight line (perhaps equipped with an end-arrow representing its privileged direction). In this way, time is thought to be a one-dimensional continuum, as opposed to multidimensional or discrete. Moreover, time is thought to be properly represented by *one* line, thereby excluding separate timelines, and *straight*, thereby excluding branching timelines and curvature. For the purposes of this paper, I will assume this traditional conception is correct. As such, I will not address whether presentists can believe in "closed time-like curves". On this topic, see Monton (2003).

[18] Lewis explicitly only gives an account for persons, but it easily generalizes. See his (1946: 146).

[19] Wasserman (manuscript: 35) – my emphasis

[20] See Lewis (1976b).

Lewisian-Wasserman Time Travel (L-W Time Travel):
Necessarily, something travels in time $\leftrightarrow_{IVO}$ there is a discrepancy between its relevant causal ordering and external time.[21]
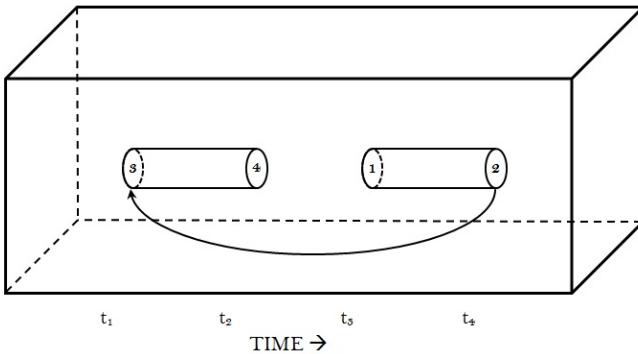
To see the account at work, consider **Figure 1**:

TIME →

Figure 1 depicts a case of discontinuous time travel to the past. The x-axis (t_1-t_4) represents external time, whilst the numbered circles (①-④) represent the ordering of the object's stages[22] according to its personal time. The object persists normally from t_3 to t_4, where its departure-stage (i.e. ②) activates a time machine, causing it to be the case (represented by the curved arrow) that its arrival-stage (i.e. ③) appears at t_1, where it again persists normally until t_2. According to L-W Time Travel, this is a case of genuine time travel because there is a discrepancy between the relevant causal ordering of the object's stages and external time: ③ is both a *causal successor* and an *external predecessor* of ②.

        As made apparent by the block-universe representation above, L-W Time Travel presupposes eternalism. To see why, assume that presentism is true and t_4 is present. In this case, ③ simply doesn't exist, and so it can neither be temporally prior to anything nor a causal successor of anything. The general approach I wish to take towards a presentist-friendly reformulation involves making use of appropriately tensed claims, as suggested by Keller and Nelson:

---

[21] L-W Time Travel is meant to express not only what conditions are necessary and jointly sufficient for time travel, but also what it is *in virtue of* that something travels in time (thus the '$_{IVO}$' subscript).

[22] I will often use the vocabulary of a *perdurantist* – basically, the view that objects persist through time by virtue of having different temporal parts, or "stages", at each moment they exist. *Endurantists* accept the denial of perdurantism, and will therefore need to understand stage-talk neutrally where needed. Nothing turns on this.

But the presentist can have just the same patterns of events happening at just the same times. Or at least, it can be the case on the presentist model that the right sorts of events *will* happen, or *did* happen, or *are* happening, at the right sorts of times. (2001: 338)

With respect to Figure 1, presentists can say the following (where t_4 is the present time):

> ② exists and ③ does not,
> WAS (③ exists and ② does not), and
> WILL-BE-FOR-O (③ exists and ② does not)

Importantly, (4) and (5) allow us to capture the external difference between the relevant stages without being committed to the existence of both stages. Likewise, the conjunction of (4) and (6) allows us to capture the relevant causal ordering of the object's stages without being committed to the existence of both stages. The *personal tense operator* in (6) reads '*it will be the case for O that*', where 'O' picks out the particular object in question. Following Wasserman, personal tense operators "are reducible to ordinary tense operators and causal facts." (manuscript: 44) More specifically: For any object, O, WILL-BE-FOR-O (Φ) $\leftrightarrow$ either (i) WILL (Φ and O has an appropriate causal successor), or (ii) WAS (Φ and O has an appropriate causal successor). (manuscript: 44)

On the revised Lewisian account of personal time, to say that for any object, O, *it will be the case for O that Φ* is to say that there exists some time t such that (i) Φ at t, and (ii) O's current stage has an appropriate causal successor at t. Once again, the presentist cannot accept this account, for she cannot quantify over nonexistent times. Fortunately, she can accept (7), because it employs appropriately tensed statements. Importantly, (7) allows objects to have appropriate causal successors in the external future *and* the external past, thereby allowing for pastward time travel. But in order to complete this account of personal time, we need to flesh out what it means – *in presentist-friendly terms* – for an object to have an appropriate causal successor.

Unfortunately, the metaphysics of causation has been especially troublesome for presentists. We can appreciate the standard worry given a few popular assumptions: (i) the causal relata are physical objects (e.g. events), (ii) causation is not always simultaneous, and (iii) if c causes e, then both c and e exist. In other words, if causation at least sometimes occurs at a temporal distance, and if the relevant physical relata need exist in order for said relation to hold, then presentism is straightforwardly false.[23] One immediately evident response on behalf of the presentist is to drop the *physical* requirement – perhaps the causal relata are best taken to be abstract objects that exist in the present. The presentist already thinks there are a variety of tensed *facts*, for example, and can happily say that some such facts provide a causal explanation

---

[23] An argument along these lines is considered by Bigelow (1996), Zimmerman (1997), Tooley (1997), Sider (1999), Markosian (2004), Bourne (2006), and McDaniel (2009).

for others.[24] Unfortunately, there are many hard questions here which would distract from my purposes.[25] Although I can't argue for it here, the option I prefer the presentist take is to leave behind the idea that causation is a metaphysically distinguished cross-temporal relation.[26] As such, I think presentists ought to deny that 'c causes e' (where 'c' and 'e' purport to name existing relata) is the correct fundamental locution for causation. Following Sider, presentists will benefit from utilizing a two-place *causal operator* that attaches to sentences (Φ, Ψ):[27]

Presentist Forward Causation:
BECAUSE (Φ), WILL$_{\text{N-UNITS-OF-TIME-HENCE}}$ (Ψ).
Presentist Backward Causation:
BECAUSE (Φ), WAS$_{\text{N-UNITS-OF-TIME-AGO}}$ (Ψ).

This approach allows the presentist to engage in causal-talk without commitment to nonpresent objects.[28] To see the causal operator at work, take:
The current appearance of a time traveler is caused by someone activating a time machine two years ago.

By utilizing Presentist Forward Causation, (8) can be translated into:
> WAS$_{\text{TWO-YEARS-AGO}}$ [BECAUSE (Someone activates a time machine), WILL$_{\text{TWO-YEARS-HENCE}}$ (A time traveler appears)].

Likewise for (10) and (11) via Presentist Backward Causation:
Jen's activating the time machine causes it to be the case that she viewed her younger self 30 years ago.
BECAUSE (Jen activates the time machine), WAS$_{\text{THIRTY-YEARS-AGO}}$ (Jen is viewing her younger self).

Certain tangential issues aside, I will adopt this approach to causation.[29] But before we can complete an account of presentist personal time, we also need to

---

[24] On fact causation, see Bennett (1988: 21-49).

[25] One such question is: if there are facts, what kind of entities are they (ontologically speaking)? Are they true propositions, or instantiations of universals? And if facts are abstract, how would they cause anything – it seems the causal relata need be *pushers* and *shovers*, and you need to be physical for that!

[26] This is where my view differs substantially from that of Keller and Nelson, who agree that "if causation can be a cross-time relation on the four-dimensionalist view, then it can be one on the presentist view too." (2001: 341)

[27] See Sider (1999: 338) and (2005: 5). It is worth noting that if (i) sentences express propositions, and (ii) facts are true propositions, then this approach supports moving to fact causation.

[28] It is also worth noting that presentists don't have to take the causal operator as primitive. One plausible reductive view of causation that the presentist could opt for, for example, says that the causal operator reduces to laws plus the instantiation of qualitative properties and single-time relations.

be able to pick out the right kind of causal explanation – namely, the one that makes for identity over time. To explicitly represent the explanation required, I will subscript the causal operator with '$_{ID}$'. The presentist can then say 'BECAUSE$_{ID}$ (O exists and has such-and-such features), WILL/WAS (there exists a y with such-and-such features)' to express that some object, O, has an *appropriate* causal successor in the external future or past. More carefully, the analysis would go as follows (for simplicity, I've left out the metric variants of the tense operators):

For any object, O, 'WILL/WAS (O has an appropriate causal successor)' $\leftrightarrow_{DF}$ O exists and has features F_1-F_n,
WILL/WAS (There exists some y such that y has features G_1-G_n), and BECAUSE$_{ID}$ (O exists and has features F_1-F_n), WILL/WAS (There exists some y such that y has features G_1-G_n).[30]
If we employ this analysis in (7), we can produce a complete account of presentist personal time (for simplicity, I've only provided non-metric truth conditions for the future-tensed personal operator):

Presentist Personal Time:
WILL-BE-FOR-O (Φ) $\leftrightarrow$ O exists and has features F_1-F_n, and either:
WILL (There exists some y such that y has features G_1-G_n, and Φ) and BECAUSE$_{ID}$ (O exists and has features F_1-F_n), WILL (There exists some y such that y has features G_1-G_n), or
WAS (There exists some y such that y has features G_1-G_n, and Φ) and BECAUSE$_{ID}$ (O exists and has features F_1-F_n), WAS (There exists some y such that y has features G_1-G_n).[31]
As utilized in L-W Time Travel, eternalist personal time appeals to genuine cross-temporal causal relations and qualitative differences between timelessly existing stages. The presentist is able to substantively mirror this account by making use of appropriately tensed statements that describe qualitative facts and the relevant causal explanation.[32] Putting this all together, we are now in a

---

[29] For some worries involving this approach, see Sider (1999: 339-40).

[30] The type of biconditional here is meant to express what it *means* (as a matter of definition) to have an appropriate causal successor. Note that the same tense operator is intended to be used throughout. Credit here is largely due to Wasserman (manuscript: 43). Formulation differences are apparent in tenet (iii): Wasserman takes a fact causation approach, rather than using Sider's two-place causal operator.

[31] This account is largely due to Wasserman as well (manuscript: 44). Again, formulation differences are due to different approaches to causation.

[32] It is worth noting that Presentist Personal Time can be used to (accurately) describe the personal futures of ordinary persisting objects. For example, imagine that I plan to order pizza tonight so that it is now true that *it will be for me that I order pizza*. This is true because *it will be the case that* someone appropriately causally related to me orders pizza. More carefully:

position to understand Wasserman's account of *non-relational* presentist time travel:

Presentist Time Travel:

Necessarily, something travels in time $\leftrightarrow_{IVO}$ there is a discrepancy between its personal tense operators and ordinary tense operators. (manuscript: 45)

For a visual representation of this account, we can reconstruct (part of) Figure 1 into a presentist-friendly model, depicted in **Figure 2**:

Figure 2

WILL-BE-FOR-O
(③ exists and ② does not)

t_1 t_4
WAS   TIME→   NOW
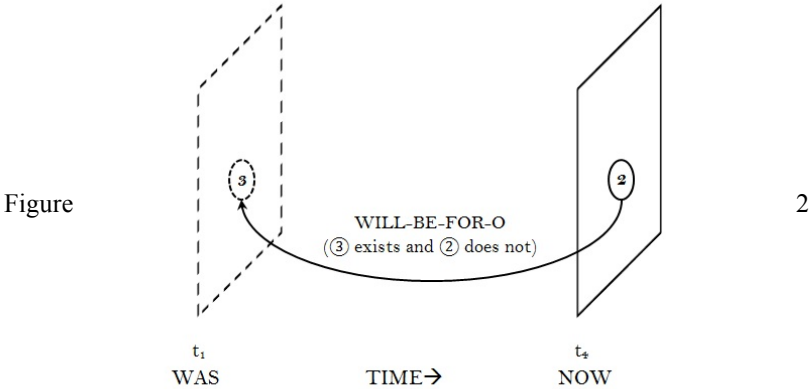
depicts two instantaneous cross-sections of the eternalist's block in Figure 1.[33] The present time is t_4, and thus according to presentism, ② exists and ③ does

---

WILL-BE-FOR-O (O orders pizza) iff (i) O exists and has features F_1-F_n, (ii) WILL (There exists some y such that y has features G_1-G_n, and O orders pizza) and $BECAUSE_{ID}$ (O exists and has features F_1-F_n), WILL (There exists some y such that y has features G_1-G_n).

As such, Presentist Personal Time seems to successfully undermine a worry articulated by Sider (my emphasis):

That I will view a dinosaur in my personal future amounts merely to the fact that I once viewed a dinosaur, and moreover that this is caused by my entry into a time machine. Since this fact *bears little resemblance to the facts that constitute a normal person's genuine future*, I could not enter the time machine with anticipation and excitement at the thought of seeing a dinosaur, for it is not true that I am about to see a dinosaur, nor is the truth much like being about to see a dinosaur. (2005: 333)

[33] Some might object to the possibility of *discontinuous* presentist time travel. One way to do this is to endorse the view that temporal continuity is necessary for identity over time. However, as Keller and Nelson (2001: 339) argue, this is a problem equally suited

not. However, the presentist is happy to say that it *was* the case (externally speaking) that ③ exists and ② does not, thereby capturing the external difference between the stages. Similarly to Figure 1, the direction of the relevant causal explanation is represented by the curved arrow. Given Presentist Personal Time, we can describe the object's personal future as follows: (i) ② exists *now* with various features, (ii) *it was the case that* ③ exists with various features, and (iii) ② existing *now* with various features *causally explains* (in the relevant sense) why ③ once exist*ed* with various features. Moreover, this gives us the needed discrepancy: it both *will be for O that* ③ exists and ② dos not, and *was the case that* ③ exists and ② dos not. Thus, according to Presentist Time Travel, we have depicted a case of genuine time travel.

   To summarize and further clarify how Presentist Time Travel mirrors L-W Time Travel, the following chart shows how the relevant discrepancy (for Figures 1 and 2) is characterized by each account:

L-W Time Travel	Presentist Time Travel
③ is temporally prior to ②	② exists and ③ does not, and WAS (③ exists and ② does not)
③ is an appropriate causal successor of ②	BECAUSE$_{ID}$ (② exists and has features F_1-F_n), WAS (③ exists and has features G_1-G_n)

2.3. *Against the third premise*
Given the account of time travel outlined above, a strong objection can forged against (P3). As I suggested above, I think it should be of no great surprise presentist-friendly time travel will not involve literally traveling to the past or future. As Kristie Miller has said, "Of course, in some sense the presentist can never travel to the past. Rather, what is the case is that it is now true that some current individual did exist in the past, and that individual's existence in the past is caused by her existence in the present." (2005: 226) Similarly, Wasserman suggests that presentists may say time travel "involves making certain past- or future-tensed statements true by acting on presently existing objects (programming flux capacitors, turning on time machines, etc.)." (manuscript: 38) As illuminated by Presentist Time Travel, this approach allows the presentist to say that time travel is *not* genuinely relational, and yet maintain all the substantive features of the traditional Lewisian-inspired account. Of course, providing a coherent account falls short of proving it to be genuinely possible –

---

for eternalists and presentists alike. Another way is to endorse the view that discontinuous *causation* is impossible. But again, it's not clear how this view could be motivated on purely presentist grounds. Furthermore, even if this was a problem particularly for presentists, we could simply have the diagrams depict continuous pastward time travel instead.

the point here is simply that The Nowhere Argument does not have the resources to rule it out.

3. The Suicide Machine Argument

Hales endorses a similar argument which he dubs 'The Suicide Machine Argument'. The basic idea is quite simple: If all of reality is confined to the present, then leaving the present to travel in time ultimately amounts to killing yourself. Hales also says that "the moves that presentists make to get around the Nowhere Argument are not successful to fend off the Suicide Machine Argument." (2010: 353) In this section, I express my disagreement. I will argue that Hales' argument fails for the same reason The Nowhere Argument fails. First, some thoughts about suicide. Hales says that "[f]or presentists, getting into a time machine is suicide – the occupant goes out of existence." (2010: 357)[34] To evaluate this claim, we first need to distinguish two possible readings of Hales' suggestion:

If activating a time machine causes the occupant to go out of existence *in the temporal sense*, then such activation is sufficient for their committing suicide.

If activating a time machine causes the occupant to go out of existence *in the atemporal sense*, then such activation is sufficient for their committing suicide. To go out of existence in the temporal sense is to *never exist again*. Yet even from the eternalist's perspective, an object which travels back in time will never exist again with respect to external time. For example, take Figure 1. While it's true that the time traveler does not exist at any moment after t_4, it's certainly *false* that the time traveler commits suicide at t_4. After all, he or she has a personal future to look forward to. So, (13) is off the table – what about the second reading? Well, consider again the eternalist's model in which an object's stages exist timelessly within spacetime. On this picture, there is a definite sense (i.e. the atemporal sense!) in which it is impossible for an object to go out of existence - but surely, suicide would not be an impossible feat if eternalism were true. Of course, the impossibility of going out of existence in the atemporal sense would not entail the impossibility of suicide on this view (even if p strictly implies q, we cannot infer that it's not possible that q from its not possible that p). But what alternative view would the eternalist opt for? That is, what would it take for a time machine occupant to commit suicide (by activating the machine) *if eternalism were true*? Plausibly:

If activating a time machine causes the occupant to *have no personal future*, then such activation is sufficient for their committing suicide.

      This account is good news for presentists. If Presentist Personal Time is on the right track, then presentism allows time travelers to go out of existence (in the atemporal sense), and yet have personal futures. Analogously to the eternalist's account, Presentist Personal Time says that the direction of one's

---

[34] For a response, see Licon (2011), a rebuttal from Hales (2011), and a second reply from Licon (2012).

personal future is determined by the direction of the relevant causal explanation. As long as the relevant causal explanation can be directed (say) backwards in time, presentism allows time travelers to have personal futures in the (nonexistent) external past.[35] Therefore, by rejecting (14) and instead following eternalists in accepting (15), presentists can sensibly say that time travel is *not* suicide.

Given the worries above, I think the best chance this argument has of succeeding is to formulate it without talk of suicide:

The Suicide Machine Argument:

(P1) If presentism is true, then necessarily, if an object leaves the present, it goes out of existence (in the atemporal sense)

(P2) Necessarily, if something travels in time, it leaves the present

(C1) If presentism is true, then necessarily, if something travels in time, it goes out of existence (in the atemporal sense)

(P3) If time travel is possible, then it's not the case that necessarily, if something travels in time, it goes out of existence (in the atemporal sense)

(C2) If presentism is true, time travel is not possible.

The first premise is uncontroversial: according to presentism *everything* is present, and so, if an object is no longer present, it no longer exists. The rationale for (P2) is straightforward as well – there is no such thing as a time traveler who never leaves the present.[36]

That leaves (P3). Unfortunately, the rationale here is unclear. At one point, Hales says that "[eternalists] insist that any sort of successful travel, spatial or temporal, involves the traveler existing at departure and safely arriving, intact and still in reality." (2010: 358) So perhaps the rationale for (P3) is supposed to be this: in order to successfully travel in time, the traveler must exist upon departure *and* arrival. But this poses no threat to presentism. With respect to Figure 2 for example, the following is true (leaving the relevant causal connection implicit):

② exists, t_4 exists, ② departs t_4, and
WAS (③ exists, t_1 exists, and ③ arrives at t_1)

---

[35] Some presentists deny the possibility of backwards causation (e.g. see Bourne (2006)), and on such grounds could argue that backwards time travel is impossible. But in order for such an argument to show that there is something about presentism *in and of itself* that rules out the possibility of time travel, the motivation for ruling out backwards causation would need to be presentist-inspired as well. One such motivation is noted by Sider (2005: fn.5) who suggests that presentists could rule out backwards causation via their acceptance of anti-reductionism about the direction of time, plus the view that causation reduces (in part) to the direction of time. However, eternalists can be anti-reductionists about the direction of time as well. For example, see Maudlin (2007).

[36] Again, this may be too quick. See fn.10.

As shown in (17), the traveler's arrival-stage existed at its respective arrival time. A more plausible take on Hales' rationale for (P3) requires us to focus on the 'still' in the quote above. The idea here, I presume, is not only that time travelers must exist upon departure and arrival, but that they must exist (timelessly) upon arrival *after* their departure. This *does* pose a threat to the presentist; however, it poses a very similar threat that we have already dealt with. Just like the third premise of The Nowhere Argument, this amounts to the (arguably question-begging) claim that traveling in time requires the existence of a nonpresent time.[37] As such, (P3) of The Suicide Machine Argument can be rejected on the grounds that Presentist Time Travel substantively mirrors the traditional account of eternalist travel, but does not commit the presentist to nonpresent objects. Moreover, Presentist Time Travel illuminates why this rejection is plausible. In cases of pastward time travel, for example, the traveler going out of existence upon departure may serve as a necessary part of the causal explanation for their once appearing in the past. If this is right, (P3) is not only false but its contrary is *true*: if time travel is possible, then it *must* be the case that time travelers go out of existence upon departure. This is what is required of the traveler to make the relevant past- or future-tensed statements true.[38]

Works Cited:

1.   Arntzenius, Frank. 2006. "Time Travel: Double Your Fun." *Philosophy Compass* 1, 6: 599–616.

2.   Bennett, Jonathan. 1988. *Events and their Names*. Indianapolis: Hackett Publishers.

3.   Bigelow, John. 1996. "Presentism and Properties." *Philosophical Perspectives* 10: 35–52.

4.   Bourne, Craig. 2006. *A Future for Presentism*. Oxford: Oxford University Press.

5.   Carroll, John. 2008. "The Nowhere Argument." *A Time Travel Website*, ed. John W. Carroll, URL = <http://www4.ncsu.edu/~carroll/time_travel/nowhere.html>.

---

[37] This claim is arguably question-begging because, as Daniels (2012: 481) says, "[t]o conceive of backwards time travel in such a way that it requires a time traveller exist at her arrival time after her departure time would, by definition, preclude the compatibility of presentism and time travel."

[38] I owe a great deal of gratitude to Ryan Wasserman for providing helpful comments on an earlier draft of this paper.

6.   Chisholm, Roderick. 1990. "Referring to Things That No Longer Exist" *Philosophical Perspectives* 4: 546–556.

7.   Crisp, Thomas. 2005. "Presentism and "Cross-Time" Relations." *American Philosophical Quarterly* 42, 1: 5 – 17.

     — 2003. "Presentism." *The Oxford Handbook of Metaphysics*, eds. Michael J. Loux and Dean W. Zimmerman, 211- 245. Oxford: Oxford University Press.

8.   Daniels, Paul. 2012 "Back to the Present: Defending Presentist Time Travel." *Disputatio* 4, 33: 469-484.

9.   Dowe, Phil. 2000. "The Case for Time Travel." *Philosophy* 75, 3: 441-451.

10. Eldridge-Smith, Peter. 2007. "Paradoxes and Hypodoxes of Time Travel." *Art and Time*, eds. Jan Lloyd Jones, Paul Campbell & Peter Wylie, 172-89. Melbourne: Australian Scholarly Publishing.

11. Godfrey-Smith, William. 1980. "Travelling in Time." *Analysis* 40, 2: 72–73.

12. Grey, William. 1999. "Troubles with Time Travel." *Philosophy* 74, 1: 55-70.

13. Hales, Steven. 2011. "Reply to Licon on Time Travel." *Logos & Episteme* 2, 4: 633-636.

     — 2010. "No Time Travel for Presentists." *Logos & Episteme* 1, 2: 353-360.

14. Hudson, Hud, and Wasserman, Ryan. 2010. "Van Inwagen on Time Travel and Changing the Past." *Oxford Studies in Metaphysics*, ed. Dean W. Zimmerman, 41-59. Oxford: Oxford University Press.

15. Keller, Simon, and Nelson, Michael. 2001. "Presentists Should Believe in Time Travel." *Australasian Journal of Philosophy* 79, 3: 333-345.

16. Lewis, David. 1976. "The Paradoxes of Time Travel." *American Philosophical Quarterly* 13, 2: 145-152.

     — 1976b. "Survival and Identity." *The Identities of Persons*, ed. A. Rorty. Berkeley: California. Reprinted in his *Philosophical Papers* 1, 55-80. Oxford: Oxford University Press, 1983.

17. Licon, Jimmy. 2012. "Still No Suicide for Presentists: Why Hales' Response Fails." *Logos & Episteme* 3, 1: 145-151.

— 2011. "No Suicide for Presentists: A Response to Hales." *Logos & Episteme* 2, 3: 455-464.

18. Markosian, Ned. 2004. "A Defense of Presentism." *Oxford Studies in Metaphysics* 1, ed. Dean W. Zimmerman, 47-82. Oxford: Oxford University Press.

19. Maudlin, Tim. 2007. *The Metaphysics within Physics*. Oxford: Oxford University Press.

20. McDaniel, Brannon. 2009. "Presentism and Absence Causation: An Exercise in Mimicry." *Australasian Journal of Philosophy* 88, 2: 323–332.

21. Miller, Kristie. 2005. "Time Travel and the Open Future." *Disputatio*, 19: 223-232.

— 2008. "Backwards Causation, Time, and the Open Future." *Int Ontology Metaphysics* 9, 2: 173-191.

22. Monton, Bradley. 2003. "Presentists Can Believe in Closed Timelike Curves." *Analysis* 63, 3: 199-202.

23. Perszyk, Ken, and Smith, Nicholas. 2001. "The Paradoxes of Time Travel." *Maui and the White Rabbit: Maori and Pakeha concepts of time*, ed. Hamish Campbell. National Museum of New Zealand: Te Papa Press.

24. Richmond, Alasdair. 2003. "Recent Work on Time Travel." *Philosophical Books* 44, 4: 297–309.

25. Shoemaker, Sidney. 1969. "Time Without Change." *Journal of Philosophy* 66, 12**:** 363-381.

26. Sider, Theodore. 2011. *Writing the Book of the World.* Oxford: Oxford University Press.

— 2005. "Traveling in A- and B- Time." *The Monist* 88: 329–335.

— 2001. *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford: Oxford University Press.

— 1999. "Presentism and Ontological Commitment." *Journal of Philosophy* 96, 7: 325-347.

27. Tooley, Michael. 1997. *Time, Tense, and Causation*. Oxford: Clarendon Press.

28. Van Inwagen, Peter. 2010. "Changing the Past." *Oxford Studies in Metaphysics* 5: 3-28.

29. Wasserman, Ryan. manuscript. *The Paradoxes of Time Travel*.

   — 1997. "Chisholm and the Essences of Events." *The Philosophy of Roderick M. Chisholm*, ed. Hahn, Lewis E. Chicago: Open-Court.

Resources for Philosophy Undergraduates

This section includes listings of journals, contests, and conferences — all of which are available to undergraduates in philosophy. If you have comments, suggestions, or questions, or if you would like to be listed here in the next issue, please contact us and we will gladly accommodate your request.

JOURNALS:

There are numerous journals, published both in print and online. The information is as recent as possible, but contact the specific journal to ensure accurate information.

> **Aporia:** Brigham young University. Submissions are due early fall. Papers not to exceed 5,000 words. Send submissions to: Aporia, Department of Philosophy, JKHB 3196, Brigham young University, Provo, UT 84602. Visit: http://aporia.byu.edu/.

> **The Bertrand Russell Society Quarterly:** Edinboro University of Pennsylvania. Visit: http://www.lehman.edu/deanhum/ philosophy/BrSQ/

> **The Dialectic:** University of New Hampshire. Submissions are due in April. Essays (15-20 pages), short critical articles, book reviews, artwork. Send submissions to: The Dialectic, c/o Department of Philosophy, University of new Hampshire, Hamilton Smith 23, Durham, NH 03824. Visit: http://www.unh.edu/ philosophy/Programs/dialectic.htm

> **Dialogue:** Phi Sigma Tau (international society for philosophy). Published twice yearly. Accepts undergraduate and graduate submissions. Contact a local chapter of Phi Sigma Tau for details or write to Thomas L. Predergast, Editor, Dialogue, Department of Philosophy, Marquette University, Milwaukee WI 53233-2289. Visit: http://www.achsnatl.org/society.asp?society=pst

> **The Dualist:** Stanford University. Submissions are due early 2013. 10-30 page submissions. For more information, see http://www.stanford.edu/group/dualist/ or contact

the.dualist@ gmail.com. Check website for information on submitting a pa- per and updates on the submission deadline.

Ephemeris: Union College. For more information, write: The Editors, Ephemeris, Department of Philosophy, Union College, Schenectady, NY 12308. Visit: http://www.vu.union.edu/~ephemeris/.

Episteme: Denison University. Due November 14. Submissions must be at most 4,000 words. Contact: The Editor, Episteme, Department of Philosophy, Denison University, Granville, Ohio 43023. Visit: http://www.denison.edu/philosophy/episteme.html

Interlocutor: University of the South, Sewanee. Questions can be addressed to Professor James Peterman at jpeterma@sewanee.edu. Send submissions to Professor James Peterman, Philosophy Department, 735 University Avenue, Sewanee, TN 37383-1000. Visit: http://www.sewanee.edu/Philosophy/Journal/2006/current.html

Janua Sophia: Edinboro University of Pennsylvania. Submissions and inquiries sent to Janua Sophia, c/o Dr. Corbin Fowler, Philosophy Department, Edinboro University of Pennsylvania, Edinboro, PA 16444. Visit: http://www.edinboro.edu/cwis/philos/januasophia.html

Princeton Journal of Bioethics: Princeton University. Visit http://www.princeton.edu/~bioethic/journal/.

Prolegomena: University of British Columbia. Visit http:// www.philosophy.ubc.ca/prolegom/ or write prolegom@hotmail.com or Prolegomena, Department of Philosophy, 1866 Main Mall, Buchanan E370, University of British Columbia, Vancouver, B.C., Canada. V6T 1Z1.

Prometheus: Johns Hopkins University. Prometheus strives to promote both undergraduate education and research, and looks for submissions that originate from any scholarly field, as long as those submissions clearly demonstrate their applicability to philosophy. Visit http://www.jhu.edu/prometheus/. Write prometheusjhu@hotmail.com or Prometheus, c/o Philosophy

Dept., 347 Gilman Hall, Johns Hopkins University, Baltimore, MD 21218.

Stoa: Santa Barbara City College. For more information, write The Center for Philosophical Education, Santa Barbara City College, Department of Philosophy, 721 Cliff Drive, Santa Barbara, CA 93109-2394. Visit: http://www.sbcc.edu/philosophy/web- site/CPE.html

The Yale Philosophy Review: Submissions due February 14. Visit: http://www.yale.edu/ypr/submission_guidelines.htm

CONFERENCES:

There are many undergraduate conferences, so contacting the philosophy departments of a few major schools in a particular area or researching on the web can be quite effective. The conferences below are by no means an exhaustive list.

American Philosophical Association: The APA website, http:// www.apa.udel.edu/apa/opportunities/conferences/, contains an extensive list of conferences.

Butler Undergraduate Research Conference: Butler University. The conference is held in mid-April. See http://www.butler.edu/urc/in- dex.html for details.

National Undergraduate Bioethics Conference: Notre Dame. Visit http://ethicscenter.nd.edu/events/nubec.shtml or write bioethic@nd.edu.

Pacific University Undergraduate Philosophy Conference: Pacific University. The conference is held in early April. Visit http://www. pacificu.edu/as/philosophy/conference/index.cfm for details.

Rocky Mountain Philosophy Conference: University of Colorado at Boulder. Visit: http://www.colorado.edu/philosophy/ rmpc/rmpc.html

ESSAY CONTESTS:

The essay contest listed below aims at a broad range of undergraduates, but there are many other contests open to students enrolled at specific universities or interested in particular organizations.

> **Elie Wiesel Essay Contest:** open to undergraduate juniors/seniors with faculty sponsor. Questions focus on current ethical issues. Submissions are due in late January. The top prize is $5,000. For more information, visit: http://www.eliewieselfoundation.org/EthicsPrize/ index.html

The Dualist would like to thank the following contributors from Stanford University:

The Philosophy Department
The ASSU Publications Board

Special Thanks:
Lenia Constantinou
Vera Haugh
Nadeem Hussain
Krista Lawlor
Alberto Martin
Teresa Mooney
Eve Scott
Sunny Toy

THE DUALIST is a publication dedicated to recognizing valuable undergraduate contributions in philosophy and to providing a medium for undergraduate discourse on topics of philosophical interest. It was created by students at Stanford University in 1992 and has since featured submissions from undergraduates across North America.

If you would like to receive an issue of THE DUALIST or to submit a paper, please contact us at the address below. We prefer that submissions be formatted according to the Chicago Manual of Style guidelines.
See http:// www.chicagomanualofstyle.org/

Papers should be submitted in electronic form only.

Visit our website for submission information:
http://www.stanford.edu/group/dualist

Please email us with any inquiries:
the.dualist@gmail.com

Or write to:
The Dualist
Philosophy Department
Stanford University
Stanford, CA 94305