Summer 8-9-2022

A Practical Conception of the Kantian Bifurcation

Nikolaus Kennelly

Follow this and additional works at: https://scholarworks.gsu.edu/philosophy_theses

A Practical Conception of the Kantian Bifurcation

by

Nikolaus Kennelly

Under the Direction of Juan S. Piñeros Glasscock, PhD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2022

ABSTRACT

There is an ambition to conceive of the human being as a composite of perceptual and desiderative faculties belonging to a causal order and a rational faculty belonging to a normative order. The problem is that this conception is unstable: If we locate the perceptual/desiderative faculties in a causal order, no room is left for the rational faculty. Consequently, to conceive the human being in full, one must alternate between two different points of view. In this paper, I argue that the solution is to reevaluate how we think about causes and norms: To say something is determined by causes is not just to locate it within a causal order but is more fundamentally to exclude it from our evaluative practices. Further, to say something is constrained by norms is not just to identify a set of evaluative practices but is more fundamentally to include it in our evaluative practices.

INDEX WORDS: Immanuel Kant, Christine Korsgaard, Constitutivism, Normativity, Necessity, Meaning

A Practical Conception of the Kantian Bifurcation

by

Nikolaus Kennelly

Committee Chair: Juan S. Piñeros Glasscock

Committee: Christie Hartley

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

August 2022

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

1   INTRODUCTION

The Kantian bifurcation of the human being into two parts, one determined by causes and the other constrained by norms, has often been given a psychological characterization. On this characterization, the cause-determined part is the locus of perceptual and desiderative capacities shared with the other animals, while the norm-constrained part is the locus of uniquely human reflective capacities. Given the presence of these parts in the human being, two states are possible: Either the human being's reflective capacities are disengaged, in which case her motions are caused by perceptions and desires which themselves have a causal origin, or her reflective capacities are engaged, in which case her motions are the causal outputs of norm-constrained discursive activity taking perceptions and desires as causal inputs.[1]

In this paper I aim to call into doubt the tenability of a psychological characterization of the Kantian bifurcation. My worry is that from a psychological standpoint, what I have called 'norm-constrained discursive activity' are causal proceedings and are therefore unintelligible as reasonable behavior. After developing this worry, I will put forward a practice-based characterization of the bifurcation. On this view, the human being is divided into a reflective part that is included in our practices and a perceptual/desiderative part that is excluded from our practices. My hope is that by providing a practice-based account, the ground will be set for developing a positive Kantian account of human agency that captures in more realistic terms the way in which desires and perceptions relate to reason.[2]

---

[1] This is of course a simplification, but I take it to be acceptable in broad outline. I have intentionally avoided engaging too heavily with the extensive literature on Kant's empirical psychology (see, for example, Frierson (2014)), particularly as developed in the *Anthropology,* because much of this literature takes for granted what I have set out to critique here. Instead, my primary interlocutors are those philosophers who attempt to ground normativity in our psychological natures (e.g., Korsgaard (1996) and Frankfurt (1971)).

[2] What do I mean by a 'practice'? As a first pass attempt to answer this question, we might turn to Rawls, who identifies a practice as "any form of activity specified by a system of rules which defines offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure. As examples one may think of games and rituals, trials, and parliaments." (Rawls 1955: 3).

Although I will spend much of this work defending a practice-based view against a psychological view of the bifurcation, my broader aim is to argue against a common way of thinking about causes and norms.[3] Call this way of thinking 'objectivist' and take as its basic commitments the views 1) to say something is determined by causes is to say there do not exist practices of an appropriate type, and 2) to say something is constrained by norms is to say there do exist practices of an appropriate type. Rather than taking an 'objectivist' approach, I argue that we should think of the distinction between causes and norms in terms of individual and communal appraisal. The basic commitments of this appraisal-based view are 1) to say something is determined by causes is to *exclude* it from our practices and thus to not hold it to standards of correctness, and 2) to say something is constrained by norms is to *include* it in our practices and thus to hold it to standards of correctness.

To see the difference between these two approaches, consider the distinction between a norm-constrained meaningful utterance and a cause-constrained belch. An objectivist determines whether a vocalization is a meaningful utterance or a belch by appealing to whether there exist linguistic practices of an appropriate type. If there do exist linguistic practices of an appropriate type, then the vocalization is a meaningful utterance, while if there do not, then the vocalization is a belch. Were an objectivist to reassess a meaningful utterance as a belch or vice-versa, they would hold that they were formerly mistaken about the existence or non-existence of the appropriate linguistic practices. On my appraisal-based view, in contrast, to assess a vocalization as a meaningful utterance is to include it in our linguistic practices, while to assess a vocalization as a belch is to exclude it from our linguistic practices. When we reassess a meaningful utterance

---

[3] The 'Objectivist' view that I criticize is extremely pervasive in analytic philosophy and the social sciences. Exemplars include Bicchieri (2005), Hawkins, Goodman, & Goldstone (2019), and Boyd & Richardson (2001). Such views typically identify specifically 'social norms' with regularities in group behavior and some cluster of psychological states.

as a belch, we exclude an utterance that once found a place in our linguistic practices from our linguistic practices and thereby 'make it into' a mere belch.[4]

Given this admittedly rough characterization of the distinction between cause and norm, I will proceed with my arguments. In §2.1, I will examine two different accounts of reflection through an examination of two different views of what 'thinking about' amounts to. The first view understands 'thinking about' in terms of causal relations, while the second view understands 'thinking about' in terms of proper inferences. Following this, I will turn in §2.2 to a third 'Korsgaardian' account of reflection that attempts to combine both causal and normative elements by means of what I call 'pre-normative first-personal necessity.' Here, I will argue that this account of reflection involves a viewpoint shift, and that once we become aware of this viewpoint shift, we encounter two very different sorts of reflection: Reflections identified as 'one's own' and reflections identified as 'merely caused'. Following this, I will examine in §2.3 a few ways that these two kinds of reflection might differ from one another. I will reject a 'deep metaphysical' distinction where the former originate causal chains and the latter belong to causal chains, and instead argue for a practice-based distinction where the former is included in our evaluative practices and the latter is excluded from them. I will then in §2.4 refine this practice-based distinction and apply it to the 'Korsgaardian' bifurcation.

2   TOWARDS A PRACTICAL CONCEPTION

2.1   The Interface

"A lower animal's attention is fixed on the world. Its perceptions are its beliefs, and its desires are its will. It is engaged in conscious activities, but it is not conscious *of* them. That is, they are not the objects of its attention. But we human animals turn our attention on to our perceptions and desires themselves, and we

---

[4] As we shall see in §3, such a reassessment may still be practically *improper* even if it cannot be referentially *mistaken*.

are conscious *of* them. That is why we can think *about* them." (Korsgaard 1996: 79)

The italicized prepositions hint at a mystery in this otherwise commonsensical account of the 'human animal'. If we excluded the prepositions, the claim that human beings can 'think' in addition to desiring and perceiving would be no more mysterious than the claim that a bat has a faculty of echolocation in addition to faculties of sight and touch. However, once we've included the prepositions, a general theory of 'aboutness' is needed.

Given the distinction between cause and norm discussed in 1, there are two directions in which we might try to develop such a theory. First, we might try to account for 'aboutness' in terms of causal relations between thought contents and perception/desire contents. In this case, the properties that we ascribe to 'aboutness' would be of the sort that we ascribe to entities that are ordinarily excluded from our evaluative practices. Consider for example the properties we typically ascribe to entities like atomic nuclei, clouds, and solar systems. In most cases, these properties are such that if we were to find atomic nuclei, clouds, and solar systems lacking them, we would not consider them defective or incomplete. Instead, we would introduce a new genus or reclassify the property as accidental. To account for 'thinking about' in these terms would similarly involve identifying properties that are such that absent them, a case of 'thinking about' would not be defective or incomplete. On the other hand, we might try to account for 'thinking about' in normative terms by ascribing it properties of the sort that we ascribe to entities that are included in our evaluative practices. In this case, 'thinking about' would involve a kind of *accountability*: One could 'think about' more or less well.

There are clear benefits and drawbacks to both directions. If we began by giving a causal explanation, then the interface between thought contents and perception/desire contents would be explained in the same way that the interface between perception/desire contents and objects-in-

the-world is (that is, in causal terms).[5] Such an approach would thus lend itself to a high degree of explanatory elegance. However, such explanatory elegance would come at the expense of feasibility, requiring a large-scale recharacterization of terms. If, on the other hand, we began by giving a normative explanation, we would already have a reservoir of terms drawn from ordinary practical life at our disposal. But while explanatory elegance came at the expense of feasibility before, here feasibility may come at the expense of explanatory elegance.

To see how, consider a normative 'inferentialist' explanation of 'thinking about.' On this sort of explanation, someone is only 'thinking about' insofar as they are disposed to draw inferences of an appropriate type. Candidate normative properties of 'thinking about' would include a disjunction of principles of formal logic, a disjunction of grammatical and syntactical rules, and the requirements of temporally extended thought.[6] Such an explanation would also specify normative properties of 'thinking about x.' If 'x' were cause-determined, it would involve principles governing explanations that always avoid the ascription of normative properties. Furthermore, if 'x' were norm-constrained, it would involve principles governing explanations that involve the ascription of normative properties. An example of a case where x is cause-determined would be if it is an atomic nucleus. Typically, when we 'think about' atomic nuclei we don't ascribe properties to them such that absent these properties the atomic nuclei would be defective or incomplete. An example where x is norm-constrained, on the other hand, would be if x is a dog. In many contexts (being a veterinarian, for example), 'thinking about' a

---

[5] I refer to 'thought contents' and 'perception/desire contents' rather than 'thoughts' and 'perceptions/desires' to highlight how, if we were to provide a strictly causal explanation, we'd have to (re-)characterize the latter in a way that doesn't involve the ascription of normative properties.

[6] The first disjunction would contain such principles as *modus ponens*, *modus tollens*, the law of non-contradiction, the law of the excluded middle, and so on. The second disjunction would contain rules governing word order, pronoun use, and so on (I assume—perhaps without much warrant—that the medium of 'thinking about' is one's spoken or written language). Finally, the requirements of temporally extended thought would include retentional and recollective requirements. This account is schematic and thus may prove inadequate upon further study.

dog would involve ascribing that dog a property of having four legs such that if it lacked that property, it would be defective as a dog.

If we understood 'thinking about' in this manner, thinking about perceptions and desires would amount to being disposed to draw inferences of an appropriate type. Thus, to think about these perceptions and desires as cause-constrained would be to pursue explanations that always avoid the ascription of normative properties. For example, to think about my desire for coffee as cause-constrained would be to think that odor particles have traveled through the air from the cup to my nostrils, that olfactory neurons in my nostrils have responded to those, and so on. Furthermore, because coffee, cups, nostrils, and so on are ordinarily characterized in a way that involves the ascription of normative properties, it would be to pursue recharacterizations of these things, and recharacterizations of these recharacterizations, and so on.

2.2   A Bridge Between the Normative and the Non-Normative?

Now, it is obvious that we do not ordinarily 'think about' our perceptions and desires in this way. Rather, our perceptions and desires typically play *roles* in practical and theoretical reflections, viz., reflections about how to act and what to believe. Consider, for example, the role that the desire plays in this account of practical reflection:

> I desire and I find myself with a powerful impulse to act. But I back up and bring
> that impulse into view and then I have a certain distance. Now the impulse
> doesn't dominate me and now I have a problem. Shall I act? Is this desire really a
> reason to act? (Korsgaard 1996: 79)

As I read Korsgaard, what is being described here is a diachronic process in which a 'desire' goes from being a cause-constrained impulse to a normative reason, where the impulse occurs pre-reflectively and the reason occurs in reflection. Because this is supposed to be an account of the *source* of a normative reason (and of normativity more generally), it is thought not to be itself

constrained by norms in the manner of the previous accounts of 'thinking about.'[7] Instead, it is supposed to have a causal origin and to be partly cause-constrained.

Central to this conception of practical reflection is the idea of a kind of pre-normative first-personal necessity. When one is caused to enter the activity of reflection, the only ways that this activity can end are if a) one finds a normative reason, or b) an outside cause interferes with the activity of reflection. In principle, (b) could occur any time one is caused to reflect, in which case the activity would not result in a normative reason. However, Korsgaard thinks that (a) does as a matter of fact occur, and that this explains the existence of rational actions, which are act-end pairs like 'go to the store to get some ice cream.'

For Korsgaard, then, the 'bridge' between the non-normative and the normative is a kind of pre-normative necessity. She takes pains to point out that this necessity isn't causal, logical, rational, or normative.[8] Instead, it is a 'first-personal' necessity arising out of the 'natural' fact of self-consciousness: Upon being caused to reflect, *I* am condemned to seek a reason. *I* cannot not seek a reason upon being so caused, even if some external cause may interrupt the activity of reflection. Consider, e.g., a case where the odor of freshly baked cookies passes by 'my' nostrils, stimulating olfactory neurons and causing a desire, which in conjunction with some other causal events that have primed me to reflect in such conditions, causes the activity of reflection. At this point, *I* emerge out of this activity as someone who *must* find a reason to act, but who at the same time 'has' a desire suggesting an act. So, I reflect on whether the act suggested by the desire is an act that I have a reason to do. I may consider that I am on a diet, and that eating cookies would

---

[7] It was pointed out to me that on this account, the thoughts expressed by 'shall I act?' and 'Is this desire really a reason to act?' are not yet norm constrained. We thus could not rely on normative principles to account for the content of such thoughts.

[8] Korsgaard writes at the very start of *Self-Constitution* that "the necessity of choosing and acting is not causal, logical, or rational necessity. It is our plight: the simply inexorable fact of the human condition" (Korsgaard 2009: 2). Here, I have identified the necessity of choosing and acting with the necessity of reflection.

wreck that diet, in which case I may have a reason to abstain. On the other hand, I may consider

that I haven't eaten anything all day, and that I am hungry, in which case I may have a reason to

eat the cookies. Both my being on a diet and my being hungry may be reflectively scrutinized,

but at some point, I would arrive at a reason to act, which would in turn normatively necessitate a

rational action. This process is illustrated in simplified form in the following diagram:
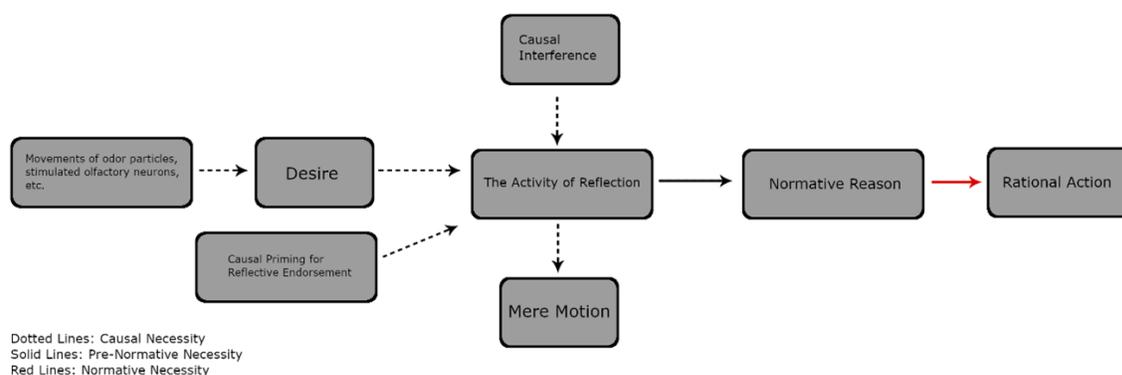
Figure 1

My first critique of this picture concerns the pre-normative necessity that is thought to

generate a reason out of the activity of reflection. The worry is that this pre-normative necessity

is passively experienced causal necessity, and that all on its own this is not enough to generate a

normative reason.[9] Consider that when 'condemned' to seek a reason, we cannot choose to do

otherwise. Ordinarily, when we cannot choose to do otherwise, we ascribe what 'we' do to an

---

[9] We might also consider the possibility of pre-normative necessity being a kind of rational necessity, an example of which is described here: "If Diotima believes that all women are mortal, and Diotima reflects that she is a woman, and the relation between these two propositions is evident to her, then Diotima must believe that she is mortal. She is confronted with the necessity of believing in her own mortality" (Korsgaard, *Response*: 7). This might seem to be an especially attractive candidate for pre-normative necessity because like normative necessity and unlike causal necessity, one can be confronted with rational necessity and not give into it. However, we should be cautious, for this similarity might be evidence that rational necessity is a species of normative necessity, and thus not a viable candidate for pre-normative necessity. Further reason to think that rational necessity is a species of normative necessity comes from the suggestion that the principles of logic may play constitutive roles in rational thought, such that failure to accord with them indicates a defect in thinking (See my account of 'thinking about' in §2.1 and my account of 'incompatibility' in §3.3).

external cause; someone who cannot choose to not fall is caused to fall by the force of gravity, someone who cannot choose to not twitch is caused to twitch by a disturbance to the central nervous system, and so on. Similarly, someone who cannot choose to not reflect is caused to reflect by an external cause.[10] This much is already explicit in the above diagram, but the worry arises if we accept with Korsgaard that in order for our reflective activity to generate a reason, we must value that reflective activity.

As a point of clarification, Korsgaard thinks that in order for our reflective activity to generate a reason, we must value ourselves as human beings, where a 'human being' is "a reflective animal who needs reasons to act and to live" (Korsgaard 1996: 121). The problem is that awareness of being condemned to reflection, understood in causal terms, does not obviously entail valuing oneself as a thing that needs reasons to act and to live. It is conceivable for a being to exist who is aware that she will be caused to reflect in certain circumstances, but who does not value her reflective nature. We may imagine that when she is caused to reflect, she disassociates from her reflective activity, which continues aimlessly until interrupted by an outside cause. If a being like this could exist, then the attempt to derive normative necessity from pre-normative necessity would be in serious trouble.

It may be worried that we are mistakenly assuming that self-awareness does not itself have its source in reflective activity. If we were to accept that self-awareness itself has its source in reflective activity, then we may argue that our imagined being is *reflectively* 'disassociating' from her 'reflective activity.'[11] In this case, it is reasonable to think that she values the reflective

---

[10] It is important to note that on this picture, even though the reflector 'has no say' on whether they reflect, they do 'have a say' about how the reflections proceed. They can, in other words, 'choose' which considerations are relevant and which considerations are irrelevant. What 'having a say' and 'choosing' amounts to, though, is fraught with controversy. In this paper, I argue that these concepts are only intelligible as normative concepts, but there have been attempts to argue that they consist in dispositions and abilities (see e.g., Clark (2009).

[11] This isn't a departure from Korsgaard's position. Consider, e.g., Korsgaard's claim: "[…] the human mind is self-conscious in the sense that it is essentially reflective" (Korsgaard 1996: 93).

activity through which she disassociates from her 'merely caused reflective activity.' If she thinks 'I was caused to reflect, but *I* do not value that part of myself,' she is valuing the part of herself that is reflecting *now*. A few moments later, she may go on to think 'A moment ago I was caused to reflect, but *I* do not value that part of myself,' again valuing the part of herself that is reflecting *now*, and so on. What we find here is that this imagined being is stuck between two conceptions of her own reflective nature. Her occurrent reflections are *her own*, and valued accordingly, while her past reflections that she is reflecting about are *merely caused*.

It follows that pre-normative necessity, understood as awareness that one is caused to reflect, only *indirectly* entails self-valuing. It entails self-valuing in the way that self-awareness *in general* entails self-valuing: To be self-aware I must be reflecting, but for *me* to be reflecting I must value my reflections as my own.[12] But, as suggested by our imagined being, this conception of reflection is very different from the 'merely caused' conception of reflection apparently on display in Figure 1. In the proceeding pages, I will develop a positive account of these two types of reflection, drawing from the distinction between cause and norm discussed in section 1.

2.3   Two Types of Reflection

Consider as prelude an arguably mistaken reading of the Kantian story according to which a being engaged in practical reflection must "be the first cause of all [her] actions" (Lectures on Metaphysics: 28:269). According to this basically objectivist view, there is a deep metaphysical distinction between the two types of reflection: The first type (the one that is identified as 'merely caused') is a part of a causal chain originating in the world, while the other type (the one that is identified as 'one's own') originates a novel causal chain. It is important to

---

[12] As we will see in the next section, this notion of valuing is one in which a kind of authority is conferred upon the valued thing, such that the thing is brought into and changes a broader network of proprieties and improprieties.

notice that on an objectivist account like this, there is no real place for a substantive notion of *valuing* in our characterization of 'her own' reflections: Either valuing on such a view can be given a causal reduction (perhaps it is understood as a psychological state) or it is not essential to 'her own' reflections. Consequently, if we were to hold that valuing is a non-reducible normative property and that it is essential to 'her own' reflections, then we would have to reject the view that the distinction can be apprehended from within the objectivist standpoint.[13]

Is there reason to think that valuing is a non-reducible normative property that is essential to 'her own' reflections? I think so, but to see why, consider the following intuitive sketch: When our imagined being values a thought, it is thereby granted a kind of authority. In the case that we have imagined, the thought reduces another thought to the status of the 'merely caused' and thereby changes what further thoughts would yield an inconsistency. We can think of the thought as analogous to the disavowal of a chess move by a chess player, for just as the disavowal would change what further moves would violate the rules, the thought would change what further thoughts would yield an inconsistency (I will explain this more thoroughly in §3). We should exercise caution, for although the disavowal of the chess move is not itself governed by the rules of chess, it is governed by the rules of further practices, e.g., those which set the boundaries of the game. Consequently, our imagined player could always be in violation of these rules in disavowing the move. From this, we could infer that a valued thought is analogous to a move in a rule-governed practice and a contradiction is analogous to a rule-violation in a rule-governed practice. Just as a move in a rule-governed practice changes which further moves are rule-following and rule-violating, the thought changes which further thoughts are rule-following and rule-violating.

---

[13] Another worry is that the view that reflection R occurring at T1 is a 'first cause' at T1 but then that same R occurring at T1 is 'caused' at T2 doesn't make sense on any standard view of causality.

Were we to accept that reflections identified as 'one's own' are essentially norm-laden in the manner discussed in the previous paragraph, I suggest that there would be reason to conceive of the distinction from an appraisal-based view: Reflections identified as 'one's own' would be included in our evaluative practices, while reflections identified as 'merely caused' would be excluded from them. Consider this by analogy with chess: In recognizing someone as playing chess, we must hold them to be bound by the rules of the game. It is not enough that we merely notice that there are pieces on a board in front of them, that they are picking up those pieces and moving them around, and so on. In a similar manner, to recognize her reflections as 'her own' rather than 'merely caused' would be to hold her to be bound by certain rules.

If we accept this much, we find that 1) because going from recognizing reflections that are 'merely caused' to recognizing reflections that are 'her own' requires us to make a shift such that we are now holding her to be bound by certain rules, an argument that doesn't mention this shift in its premises (like Korsgaard's) wouldn't have derived normative conclusions from non-normative premises, and 2) normativity would not have its source in reflective activity, for the relevant reflections would themselves already be constrained by norms.

Even if we were to accept (1) and (2), it may be objected that the chess analogy suggests a misleading picture of 'one's own' reflections. That is, it suggests that we should think of these reflections as equivalent to a game consisting of a set of inferences that are equivalent to moves. On this picture, a reflector might infer 'go to the library' from 'read Swann's Way', where a part of what would make this inference good would be that it accords with a rule holding that means should be inferred from ends. Even if this were a viable model of reflection, it is not the only one that can be thought of, and it is not the one that Korsgaard has in mind. That is, from a Kantian-Korsgaardian perspective, we have conceived of reflection as maxim *construction*, but reflection

properly conceived is maxim *evaluation*. On a maxim construction model, reflection would involve drawing inferences between the parts of maxims that are good according to whether they accord with rules originating from *outside* the activity of reflection. On a maxim evaluation model like Korsgaard's, in contrast, such rules are supposed to have a source that is *internal* to the activity of reflection.

2.4   A Clarification of Korsgaard's Position

Here, I should clarify an ambiguity in my earlier description of Korsgaardian practical reflection. As I have described it, an agent is prompted to reflect when confronted by a desire suggesting an *act*, e.g., the act 'eat the cookies', and reflection consists in seeking what I have called a 'reason' to do or not do that act. Because reflection is prompted by a suggested *act* and not a *maxim* (a suggested act-end pair), and because the reason seems very much like an end, we may think that Korsgaardian practical reflection is a maxim construction model. However, such a thought is mistaken: For Korsgaard, practical reflection is prompted by maxims, which are then tested to determine whether they embody reasons.[14] To put this another way, rather than being confronted with an act like 'eat the cookies' and then reflectively seeking an end like 'to satisfy my sweet tooth', an agent is confronted with a maxim like 'eat the cookies to satisfy my sweet tooth' and then reflectively tests that maxim to determine whether satisfying her sweet tooth is really a reason to eat the cookies.[15]

---

[14] Korsgaard writes: "Kant's question [and, it turns out, her own] is whether the act and the end are so arranged, or related to one another, that the maxim can serve as a law" (Korsgaard 2009: 15). Further: "the reason for an action is not something outside of or behind or separate from the action at all, for explicating the action and explicating the reason, are the same thing. Rather, an action is an essentially intelligible object that embodies a reason, the way a sentence is an essentially intelligible object that embodies a thought" (Korsgaard 2009: 16).

[15] We should note a tension between the earlier view that desires prompt reflection and the current view that maxims prompt reflection. On a conception of desires as brute facts, it is reasonable to think that they could at most serve as something like 'raw material' on a maxim construction model. That is, because a desire that did not prompt reflection would translate into 'mere motion' lacking any goal or aim, we might think that if it did prompt reflection, it would *at most* specify a goalless act. Consider the case where the odor of freshly baked cookies passes by 'my' nostrils, stimulating olfactory neurons and causing a desire. If we understand this desire as a brute fact, the most

On this picture, reflection generates rational action when a maxim successfully 'passes' evaluative reflection, which involves its being willed as universal law.[16] Willing the maxim as universal law makes the action 'one's own', for "if all my decisions were particular and anomalous, there would be no identifiable difference between my acting and an assortment of first-order impulses being causally effective in or through my body" (Korsgaard 1996: 228). For *me* (emphasis on the 'me') to eat the cookies to satisfy my sweet tooth, I must be committed to making the world such that anyone in it would act similarly in similar circumstances. If I am not so committed, then *I* am reduced to an assortment of causally effective first-order impulses, where these are basically equivalent to desires.

---

reasonable rendering of 'being confronted' by the desire would be: Becoming aware of the kinds of motions that the desire would have caused if left unimpeded by reflection. But if being 'confronted' by a desire is understood in this way, desires are clearly inadequate for serving as suggestions for action. Because actions contain ends, the act specified by the desire would need to be reflectively supplemented with an end. In this case, we may suspect that we would need to posit two phases of reflection: An initial maxim construction phase prompted by a desire and a later maxim testing phase prompted by a maxim. In the first phase, I would initially be aware of bodily motions that would have been caused by the desire had I not been prompted to reflect and then I would reflectively append an end to an act identified with these motions. For example, I would start by being aware that my arm would have reached for the cookies had I not been caused to reflect and then I would append in reflection an end like 'to eat a cookie' to the motions-cum-act of 'reach for the cookie.' The strangeness of inferring ends from acts rather than the other way around should prompt concern here, but this is the position that we are led to if we accept that a) desires are brute facts, and b) only maxims can serve as suggestions for action.

[16] As a point of clarification, the standard here is *not* whether one possesses a psychological state of 'universal willing,' but is rather whether one is committed to 1) acting similarly in similar circumstances, and 2) making the world such that anyone in it would act similarly in similar circumstances. Meeting (1) would constitute one as an egoist agent, while meeting (2) would constitute one as a moral agent. Although demonstrating the constitutive role of (2) is a challenge, the constitutive role of (1) can be demonstrated through consideration of an apparent counterexample: Someone decides to act just this once and never again. Although it is not immediately obvious, this can be interpreted as *both* a case of universal willing and a case of particularistic willing. It is a case of universal willing if they are willing a very precise maxim and what is meant by 'just this once and never again' is just that they would never find themselves in circumstances where they could act on that maxim ever again. It is a case of particularistic willing, on the other hand, if they are not willing a maxim at all. To will a maxim, whether a general one like 'reach for the cookies to take a bite' or a specific one like 'reach for the cookies to take a bite on this day at this time in this place …', is just to make the maxim into a law. To will a maxim that doesn't specify time or location would be to make a law that holds regardless of when and where you are. Thus, to will 'reach for the cookies to take a bite' would have the absurd result of making it law that you reach for the cookies to take a bite all the time and everywhere. From this, we find that if willing maxims is not going to lead to absurdity, one would have to construct very precise maxims. We thus find that (1) captures the *active* dimension of agency, for the prospect of commitment motivates agents to include as much as possible in the scope of what they will.

The claim that first-personal universal willing marks one off as an agent rather than an assortment of causally effective first-order impulses should strike us as odd, especially when we consider that the reflections and actions of a universal willer can always be reclassified as caused motions. We can infer that the recognition that one is a universal willer, and thus an agent, must involve adopting a perspective that is distinct from the perspective we take with respect to them when we tell a causal story. This is the perspective Korsgaard has called the "plural first-person," which is one where "we are joined with others in the processes of deliberation and justification" (Korsgaard, 2002: 55). It is from *within* this perspective that one is intelligible as an agent only if they universalize their maxims and it is from *within* this perspective that not universalizing their maxims reduces them to an assortment of causally effective first-order impulses.

If we are attentive to the perspectival difference, we can avoid a potential equivocation between the 'merely caused' that is ascribed in the plural first-person and the 'caused' that is ascribed in the third-person. To say that something is 'merely caused' is to exclude it from our practices, while to say something is 'caused' is to undertake to tell a deterministic story.[17] The distinction between the greeting 'hello' and a belch that sounds like 'hello' is a plural first-person distinction between an utterance occurring inside our specifically linguistic practices and a 'merely caused' vocalization excluded from our linguistic practices. From a third-personal perspective there is no place for this kind of distinction. Rather, from the third-person, the distinction between the greeting and the belch is a distinction between two different causal streams.

---

[17] Whether we can maintain the perspectival difference and the two kinds of causes depends on whether undertaking to tell a deterministic story is coextensive with undertaking to exclude as much as possible from our practices. I have assumed that it is not, although more work needs to be done in defense of this view.

With this perspectival difference in place, we can reevaluate the bifurcated conception of the human being in terms of the plural first-person. Rather than classifying perceptions and desires as 'caused' in the third-personal sense, we can instead classify them as 'merely caused' in the plural first-personal sense. In that case, the human being would be divided into a reflective part that is included in our practices and a perceptual/desiderative part that is excluded from our practices. Here, to include the reflective part in our practices is to hold it to standards of correctness. It is to treat it as the sort of thing that can succeed or fail, be corrected or damaged, and in general be evaluated. To exclude the perceptual/desiderative part from our practices is to not hold it to standards of correctness. It is to treat it very much as we treat atomic nuclei and clouds, as the sort of thing for which evaluation is inapt.

An advantage of adopting such a practice-based account of the bifurcation is the rational and perceptual/desiderative parts are apprehended from within the same perspective. Rather than alternating between a third-personal perspective that leaves no room for rationality and a plural first-personal perspective that leaves no room for perceptions and desires (or worse, sticking to a third-personal perspective and reducing rationality to a causal stream), the bifurcation can be apprehended from within a perspective accommodating both reason and the relevant forms of perception/desire.[18]

Now, although Korsgaard's plural-first person accommodates something very much like my practice-based view, it does involve a few dubious assumptions. Consider, for example, that it gives priority to the first-personal perspective, which builds in a cause-constrained conception of perceptions and desires. That is, when first-personally conceived, perceptions and desires are

---

[18] By 'relevant forms' I mean the forms that can serve as *suggestions* to reason.

presented as entities that one is 'confronted with' from the 'outside'—as, in other words, entities lying outside of our evaluative practices.

Although I do not have space to argue that a practice-based characterization does not need to be tethered to the first-personal perspective, it is worth noticing that there is at least a conceptual distinction to be drawn here. We may consider, for example, that there is room for both first-personal and second-personal perspectives in the very same practice of baseball: A player stuck on second base may first-personally perceive that the pitcher has thrown the ball to the catcher and then attempt to steal base while an umpire participating in that same practice may second-personally call out that player for attempting to steal base on a foul ball.

3   THE PRACTICAL AND THE REAL

3.1   Some Anti-Realism Worries

The position described here might be thought to be untenably anti-realistic because of the appraisal-dependency of norm and mere cause. That is, by explaining norms in terms of practical inclusion and mere causes in terms of practical exclusion, something that is neither assessed as a norm nor a mere cause might be thought to be left 'indeterminate' between the two. This worry can be offset, although not entirely done away with, by noticing that when we conceive of the kinds of cases that provoke these indeterminacy worries, we are often shifting from a third-personal to a practical perspective and treating whatever is third-personally conceived as 'indeterminate' because we haven't done a pragmatic reduction. In other words, the thought that the unappraised thing is 'indeterminant' results from a failure to acknowledge what *we* are *doing* when we are third-personally conceiving, where the explanation offered here is that we are practically excluding. To make this clearer, consider the following case:

> Someone makes a sound at T1 which is then assessed at T2 as either a meaningful utterance or as a mere sound.

The worry for the appraisal-based view is that the sound at T1 is 'indeterminate': It is neither a meaningful utterance nor a mere sound. To offset this worry, notice that *we* are conceiving the 'person' making a sound at T1 from the third person, in which case *we* are practically excluding and the sound at T1 is merely caused. We are then either going on to give an account of the assessment at T2 from the third person, in which case our account involves a reduction of the distinction between norm and mere cause to two causal streams, or we are giving an account of the assessment at T2 from the practical point of view, in which case we are shifting perspectives to one where the distinction between norm and mere cause can be apprehended as such. If we go with the former then the distinction corresponds to two ways of practically excluding, while if we go with the latter then the distinction corresponds to practical inclusion and exclusion. This explanation offsets rather than defuses the worry because we can always ask in virtue of what we are practically including or excluding. We can then divert this further worry through an analysis of what we are doing when we ask that question, and so on in a regress. Consider next this case:

> Someone assesses the thought 'I want to read Proust' as a reason to go to the library at T1 and reassesses the thought as a mere desire at T2.

The worry for the appraisal-based view is that if something's being a reason or a mere desire consists in nothing other than its being assessed as such, the thought trivially goes from being a reason at T1 to being a mere desire at T2. This is to accuse the view of something like the second horn of the Euthyphro Dilemma vis-à-vis norms and causes. Although there is something to this worry (which I will come back to in a moment) the way that it is framed here makes it susceptible to the same kind of pragmatic reduction as in the last case.

To see how, consider that one and the same 'thought' is supposed to be assessed as a reason at T1 and as a mere desire at T2. What is the status of this 'thought' that persists from T1

to T2 and is assessed differently each time? One possibility is that this 'thought' is a third-personally conceived causal entity, in which case we might suspect there of being either a perspectival shift or a reduction. It might be held that a perspectival shift is off the table because unlike the last case there is no time at which the 'thought' is unassessed, however this would be to confuse what is being conceived with the conception: It isn't as if in order to conceive of a case where something happens at T1 followed by something happening at T2, we must first conceive of the thing that happens at T1 at $T_1 1$ from one perspective and then the thing that happens at T2 at $T_1 2$ from one perspective. We could instead conceive of the thing happening at T1 from P1 at $T_1 1$ and then from P2 at $T_1 2$ followed by the thing happening at T2 from P1 at $T_1 3$ and then from P2 at $T_1 4$. Regardless of how these conceptions proceed, we could then give an account of what we are doing when we are conceiving much as we did before.

Perhaps rather than thinking of 'the thought' as a third-personally conceived causal entity, we should think of it by appeal to the nature of conception. More precisely, it may be that what is designated by 'the thought' is indeterminate conception *itself*: Much as it is indeterminate whether an object seen from a distance is a rock or a house, it is indeterminate whether a thought is a reason or mere desire. We can notice that to think in these terms involves the introduction of a representing (a conception indeterminate between a rock or house) and a represented (a rock or house), thereby making it a basically objectivist view. However—and this will be clearer in the next sections—there may be a way to account for such representings and representeds in terms of proprieties that are fixed by social practices that *we have endorsed*.

In the next section, I will start to develop a more sophisticated version of the appraisal-based view that leaves room for indeterminate conception and even 'mistake', where 'mistake' occurs in cases where e.g., a thought is reassessed as normative even though it is 'really' caused.

The account I will develop will not depend on an 'external' standard (one that does not in the end depend on appraisal) for cause and norm, and neither will it require us to reconceive the assessment that something is a cause or a norm along referential lines.

3.2   Assessments Within Communal Practices We Have Endorsed

One way to start to leave room for something like 'mistake' on the appraisal view would be to notice (what is already hinted at in §2.3) that the assessment that an utterance is meaningful or mere noise is *itself* a move made inside a practice. Not unlike ordinary games, the practices in which such assessments are made vary a great deal, but they are alike in involving proprieties and improprieties. While in one practice, the assessment that an utterance is meaningful rather than mere noise would be proper, in another that very same assessment would be improper. Consider a speaker who assesses her own past utterance as meaningful. In a communal practice embodying the Enlightenment ideal of personal autonomy, it would likely be inappropriate—a violation of communal norms—for another agent to fail to defer to that agent's assessment of the meaningfulness of her own utterance. But now consider a communal practice embodying the serf-king dynamic of feudal Europe. In such a case it would likely be appropriate—where this means in accord with communal norms—for an agent to defer to a king's assessments of a serf's utterances as meaningless even in the teeth of the serf's protestations.

Given these observations, we might try to revise our account of meaningfulness from 'x is a meaningful utterance iff and because it is assessed to be' to 'x is a meaningful utterance iff and because it is *properly* assessed to be, where the propriety of this assessment is fixed by social practices.' Notice that although no room would be left on this view for mistake, understood on objectivist lines, room would be left for impropriety. In a social practice embodying the Enlightenment ideal of personal autonomy, Y's assessment that X's utterance is meaningless,

where X herself assessed her utterance as meaningful, would be improper. Y's assessment would thus fail to ground the meaninglessness of X's utterance.

It may be worried that by grounding meaning in social practices rather than assessments, we appear to have endorsed an objectivist account of meaning. To see the problem with such an objectivist account of meaning, we need only consider the difficulty faced by *us*, denizens of the post-Enlightenment, in assessing the serf's utterance (and protests) as meaningless. Communal practices according to which a king's assessments ground meaningfulness are not ones that we can endorse without violating the norms of the social practices we have endorsed. This may suggest this further reformulation (FR): "X is a meaningful utterance iff and because it is properly assessed to be, where the propriety of this assessment is fixed by social practices *we have endorsed*.'

But with this, the realist may object that we have simply pushed the Euthyphro problem back, for now the endorsement of a social practice seems to be a trivial and arbitrary matter. Further, once we recall that to assess a noise as a meaningful utterance is to include it in our evaluative practices, the view would seem to face a circularity objection: An utterance is included in our evaluative practices when it is properly assessed to be, where the propriety of this assessment is fixed by social practices we have endorsed.

However, both objections can be met. The circularity objection can be met by noticing that it equivocates between 'evaluative practices' and 'social practices.' These need not be the same thing and may even be such that one is embedded in the other: On my view, in fact, evaluative practices are embedded within social practices.[19] This response to the circularity

---

[19] This is just the by now familiar claim that there are proprieties governing the ascription of normative properties. For example, it is improper in the Modern Scientific Worldview social practice to ascribe a property to Tungsten such that a sample of Tungsten lacking it would be defective or incomplete (instead, we'd reclassify). On the other

objection goes a long way toward my response to the objection that we've simply pushed the Euthyphro problem back: Even if it were conceded that the endorsement of a given social practice is a trivial and arbitrary matter—which is not a view I accept—we would still have addressed the worry that meaningfulness is a trivial and arbitrary matter, for we can bracket the question of whether we are (or whether it even makes sense of say we are) 'mistaken' in our endorsement of our post-enlightenment social practice, and instead concern ourselves with whether the assessment of an utterance as mere noise is inapt. Recall, for example, the case where we assess an agent's utterance as mere noise even as she protests that it is meaningful: [In the post-Enlightenment social practice we have endorsed] Our assessment is inapt (assuming she's sane, an adult, and so on) while hers is proper.

3.3 Rationality and Incompatibility

It may be worried that the way we have framed our post-Enlightenment social practice pays insufficient attention to the specifically *rational* constitution of the agents involved. That is, it isn't as if *whenever* an agent affirms the meaningfulness of her own utterance, it is thereby meaningful. Consider cases where in affirming the meaningfulness of her own utterance, an agent is undertaking incompatible commitments. Say that a Proust devotee affirms the meaningfulness of an assertion to the effect that she despises and has always despised Proust. If rather than claiming that she had made a shift in her commitments or that she had been lying all along she continues to proclaim her admiration of Proust while defending the meaningfulness of this assertion, [In the post-enlightenment social practice we have endorsed] either the assertion

---

hand, it is proper in the Veterinarian Office social practice to ascribe a property to dogs (say having four legs) such that a dog lacking four legs would be defective as a dog.

must be revised as mere noise, or the speaker's status as a rational speaker is called into question.[20]

Before continuing, I should say something about how I understand 'commitments.' Rather than thinking of commitments along objectivist lines as intentions (which I understand to be private attitudes), I think of commitments as statuses constituted by certain proprieties. If someone is committed to taking their opponent's queen in a game of chess, this is to be made sense of in terms of them being subject to a separate set of proprieties from if they are committed to promoting their pawn. The moves that it would be appropriate for someone with the former commitment to make would be different than the moves that it would be appropriate for someone with the latter commitment to make. Further, because such proprieties are in the end *ascribed by us*, there is not a serious conflict between an account of meaning that is grounded in whether some speaker has the appropriate commitments and an account of meaning that is grounded in whether some such utterance is assessed as meaningful. What it is to assess an utterance as meaningful is to [properly] hold the speaker accountable for what they go on to say, where holding a speaker accountable for what they go on to say constitutes their having certain commitments.

With this account of 'commitment' out of the way, I will next turn to an account of 'incompatibility.' When I say that commitment to A is 'incompatible' with commitment to B I have in mind cases where commitment to A precludes entitlement to B (and vice versa).[21] Given

---

[20] This claim is analogous to the claim that if a chess player makes an illegal move, either the illegal move must be revised as an accident, or the chess player's status as a competent chess player is called into question. There is room for vagueness here, for a single non-accidental illegal move doesn't usually call into question whether someone is a chess player (although this may depend on the move). It is only when non-accidental illegal moves become frequent that a chess player's status is called into question. Such vagueness also holds in the case of a speaker who asserts a commitment that is incompatible with her other commitments: A single incompatible commitment doesn't seriously call into question her status as a rational speaker, but if they become frequent then we begin to question whether she is a rational speaker.

[21] In this respect, my account draws heavily from Brandom (1994).

this, my use of 'incompatible' picks out an extremely broad genus, including material incompatibilities, practical incompatibilities, and logical incompatibilities as species. Examples include commitments to something's being tungsten and its melting point being 500°C in the material case, commitments to being a good teacher and regularly sleeping in while class would be happening in the practical case, and commitments to that being a dog and that not being a dog in the logical case. Someone who is committed to that being tungsten would thereby also be committed to that having a melting point of 3,422°C and thus not entitled to that having a melting point of 500°C; someone who is committed to being a good teacher would thereby also be committed to showing up to class regularly and thus not entitled to regularly sleep in while class would be happening; someone's committed to that being a dog would also be committed to that not-not being a dog and thus not entitled to that not being a dog, and so on.[22]

3.4 Concept and World

As suggested in the last section, the view developed here is not strictly anti-objectivist, but instead aims to give an account of objectivism in terms of more fundamental normative practices. Consider a case of ostension such as '*that* is Tungsten.' On an objectivist reading, we can think of the expression '*that* is Tungsten' as an abbreviation for a complex proposition that stands in or fails to stand in a representing relation to the facts. An appraisal theorist thinks that when we say that this is an 'abbreviation' we are pointing to implicit proprieties: Someone who asserts '*that* is Tungsten' is thereby be entitled to assert '*that* has a melting point of 3,422°C' and not entitled to assert '*that* has a melting point of 500°C.' Further, the appraisal theorist thinks

---

[22] It goes without saying that speaking *about* incompatibilities doesn't threaten a speaker's rationality, for the asserted commitments may be entirely compatible.

that what it is for '*that* is Tungsten' to mean anything at all is for the speaker to be properly subject to such proprieties, where to be properly subject to them is to be rational.

It may be suggested that we could accept that meaning depends on being properly subject to proprieties and still accept a basically objectivist view according to which propositions stand in or fail to stand in representing relations to facts. We could, for example, conceive of the meaning-constituting proprieties as holding in a specifically discursive 'realm' that is separate from, but potentially representative of, a 'realm' of facts. Against this view, it may be suggested that as both discursive activity and non-discursive motions are temporally extended and spatially situated, it is a prejudice (perhaps of talkative philosophers) to think that only the former could be constrained by meaning-constituting proprieties. Why couldn't the meaning of '*that* is Tungsten' be constituted not just by proprieties governing discursive behavior, but also proprieties governing non-discursive behavior? Perhaps for '*that* is Tungsten' to be meaningful, it is not only proper for the speaker to be entitled to assert '*that* has a melting point of 3,422°C' but also entitled to harvest it for use as filaments in light bulbs and not entitled to harvest it for dinner.

The view that I am gesturing toward is one according to which mastery of certain concepts depends upon us fully being *in* the world. To fully be in the world, on this view, requires the contours of these proprieties to *follow* the world. The issue with someone who asserts '*that* is Tungsten' but who resolutely affirms that it can never be melted is not a failure of correspondence between 'their' concept of Tungsten and the world, but rather that they have failed, by their own lights, to master the concept.[23] Such a failure is 'by their own lights' because in judging something to be Tungsten, they are [in social practices we have endorsed] conferring a

---

[23] Of course, if saying that it can never be melted is a rhetorical exaggeration and they are committed to it having a melting point of 3,422°C this would not be the case.

kind of authority on it according to which *it* is to serve as a standard of correctness.[24] Much more remains to be said on this issue, but for now these remarks must suffice.

4   CONCLUSION

Over the course of this thesis, I have suggested a conceptual shift in how we think about causes and norms. Rather than approaching the distinction from an objectivist standpoint, where assessments are made but the world is left in place, I have defended an approach that locates us in the practical world. That is, on my 'appraisal-based' view, to assess something as cause-determined is not a matter of identifying an absence of appropriate practices but is rather a matter of excluding it from our practices. Further, to assess something as norm-constrained is not a matter of identifying the presence of appropriate practices but is rather a matter of including it in our practices. Once we've made this conceptual shift and applied it to the bifurcated conception of the human being, we find that we can circumvent many of the problems that arise from viewing the bifurcation from a third-personal perspective or both a third-personal perspective and a plural first-personal perspective. That is, rather than having to give up reason or alternate between two very different viewpoints, we can apprehend both reason and perception/desire from within the same viewpoint.

The arguments presented in these pages constitute a kind of partial groundwork, more like a nudge than a fully developed theory, for future ethical enquiry. We do not yet know where it will lead, but we can suspect a few likely directions. If we recall our discussion of the imagined being in §2.3, for example, we encountered the idea that excluding something from a practice is itself a move made within a practice. Think of someone who disclaims the meaning of

---

[24] Along these same lines, Brandom writes: "[…] taking or treating something in practice as a representing is taking or treating it as subject to normative assessment as to its correctness, in a way in which what thereby counts as represented serves as a standard" (Brandom, 2020, pp. 119)

another's utterances in the middle of a conversation. Such behavior would likely violate the rules of proper etiquette, and if done repeatedly, would result in their exclusion from the conversation. In this way, the rules of proper etiquette would contribute to the maintenance of meaningful conversation, for only those who recognize the utterances of their co-participants as meaningful would be included. Similar mechanisms may help maintain the bifurcation, but to demonstrate this in a philosophically rigorous way would be an immense project.

Yet another large project appears on the horizon when we begin to suspect that a practice-based characterization of the bifurcation would imply or at least make possible a kind of agency-pluralism, perhaps of the sort recently defended by Lavin (2017). That is, in much the same way that game-practices can come in many different forms, we might suspect that agency-practices can come in many different forms. Perhaps there are some communities for which the inclusion of perceptions and desires in evaluative practices would be permitted or even actively encouraged. Although these speculations may at first strike us as objectionably objectivist, there may be a way to approach them in an appraisal-based manner.[25] One strategy, already hinted at in the last section, would be to adopt a kind of 'bilingualism' about agency in which communities that maintain forms of agency very different from our own are included in our evaluative practices.

---

[25] Such speculations may strike us as 'objectivist', as I use the term, because they take the form of an inquiry into whether such communities exist. There is no indication here that the intelligibility of such communities depends on us taking a normative stance with respect to them.

REFERENCES

Brandom, Robert. "Freedom and Constraint by Norms." *American Philosophical Quarterly* 16 (3) (1979): 187-196.

Brandom, Robert. *Making it Explicit.* Harvard University Press, 1998.

Boyle, Matthew. Additive Theories of Rationality: A Critique. *European Journal of Philosophy* 24 (3) (2016): 527-555.

Boyd, Robert and Peter J. Richerson, 2001, "Norms and Bounded Rationality", in Gigerenzer and Selten 2001: 181–296.

Cristina Bicchieri, The Grammar of Society: the Nature and Dynamics of Social Norms. New York: Cambridge University Press, 2006.

Clark, Randolph. "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism," *Mind*, 118. (2009): 323–51.

Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person." Journal of Philosophy 68 (1971).

Fix, Jeremy. "Two Sorts of Constitutivism." *Analytic Philosophy* 62 (1) (2021): 1-20.

Frierson, Patrick. *Kant's Empirical Psychology.* Cambridge: Cambridge University Press, 2014.

Hawkins RXD, Goodman ND, Goldstone RL. The Emergence of Social Norms and Conventions. Trends Cogn Sci. 2019 Feb;23(2): 158-169.

Kant, Immanuel. Lectures on Metaphysics, Karl Ameriks and Steve Naragon, trans. eds., Cambridge: Cambridge University Press, 1997.

Korsgaard, Christine M., and Onora O'Neill. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.

Korsgaard, Christine M., "Normativity, Necessity, and the synthetic a priori a response to Derek Parfit". Unpublished manuscript, www.people.fas.harvard.edu/~korsgaar/Korsgaard.on.Parfit.pdf

Lavin, Douglas. "Forms of Rational Agency." *Royal Institute of Philosophy Supplement* 80 (2017): 171–93.

Rawls, John. "Two Concepts of Rules." *The Philosophical Review,* 64 (1) (1955): 3-32.