

University of Mississippi

eGrove

---

Electronic Theses and Dissertations

Graduate School


---

2011

## Korsgaard on the Reason and the Normativity of Rationality

Robert Colin English

Follow this and additional works at: <https://egrove.olemiss.edu/etd>

 Part of the [Philosophy Commons](#)

---

### Recommended Citation

English, Robert Colin, "Korsgaard on the Reason and the Normativity of Rationality" (2011). *Electronic Theses and Dissertations*. 99.

<https://egrove.olemiss.edu/etd/99>

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact [egrove@olemiss.edu](mailto:egrove@olemiss.edu).

**Korsgaard on Reason and The Normativity of Rationality**

A Thesis

Presented for the

Master of Arts Degree

The University of Mississippi

Robert Colin English

May 2011

Copyright Robert Colin English 2011  
All Rights Reserved

## ABSTRACT

Korsgaard (2009) argues, against Kolodny (2005) and Broome (2007), that rational requirements are in fact normative. In her view the normativity of rational requirements is a function of their constitutive role in the deliberative activity of reason. After surveying the treatment of this question in the relevant literature, I explain Korsgaard's theory using pure constructivism as a framing device. I then argue that not only is her account of deliberative reason as an activity unsatisfactory (specifically, it fails to defeat the intuition that charges of boot-strapping are deeply problematic, and makes the adoption of reasons for belief from the deliberative perspective a function of an agent's commitment to principles and not of her seeing the belief as true), but that she is unable to account for the normativity of rationality (because her theory is unable to provide an answer that avoids regress or is not trivial to her own "normative question" when it takes a rational requirement as its object).

## THANKS

I would like to thank the faculty of the Philosophy Department at the University of Mississippi. Whatever philosophical insight found in what follows is owed in large part to their influence. Any mistakes are my own.

## TABLE OF CONTENTS

1 Introduction.....	1
2 Rationality in What Sense?.....	6
3 Korsgaard's View.....	22
4 Objections to Korsgaard's View.....	56
5 Conclusion.....	85
Bibliography.....	86

## INTRODUCTION

“Why should I not violate *modus ponens*?” This question can be asked more generally as well: “why must I obey standards of rationality?” In everyday discourse one often hears rebukes that express a similar idea: “You need to think straight...” or “you should think more carefully...” Philosophers tend to frame such protests differently, but there is a crucial feature common to such utterances. Namely, they purport to be *normative*; they express the idea that we *ought* to think in some ways rather than others, or suggest that some ways of thinking are *better* than others, and that correspondingly, it is *wrong* or *incorrect* to think in ways that stray from sanctioned modes<sup>1</sup>. The practices of praise and blame surrounding *how* we think are a central feature not only of our intellectual endeavors, but inform how we live and interact more generally. This is to say that a concern with the apparent normativity of thought is not solely a philosophical question.

I will be concerned with the nature of rationality. When one thinks as one should, we often say that one is rational, and when one goes awry, we tend to say that one is irrational.<sup>2</sup> There are exceptions. The poet at work on completing a stanza is not usefully instructed to think rationally, we might say, nor is she aptly criticized for failure to adhere to the canons or standards of that phenomenon. At least some artistic creation is

---

<sup>1</sup> One might say: from the sanctioned *patterns*. This, though, would be to frontload a controversial point that should be settled only by substantive debate - that is, it is not obvious that thinking rationally ought to be construed as thinking that instantiates some sanctioned pattern.

<sup>2</sup> The matter is a great deal more nuanced than this crude surmises suggests.

aply characterized as thought, and whether this kind of thought is evaluable in terms of rational or irrational appraisal is a question to be answered elsewhere.<sup>3</sup> But there is reason to think that the character of even artistic thinking is not entirely insulated from rationality. The poet, for example, sees that the line she just composed has a stilted metaphor, one that could be interpreted in a way that runs counter to the general themes of the poem. If after composing this line the poet judges it to have a stilted metaphor that undercuts the theme of the poem yet keeps it, despite believing that the resulting poem will not be as good a poem as it would be without it, then we have the intuition that it is fitting to call her irrational. Laying such thoughts aside, “Why should I be rational?” or alternatively, “why should I think rationally?”<sup>4</sup> will be the general guiding questions for this study.

In “The Activity of Reason” (2009), Korsgaard sketches an account of theoretical reason similar her view of practical reason.<sup>5</sup> Reason is an *activity* undertaken by creatures with a special sort of awareness, and has its origin in the “reflective distance” that each person occupies with respect to the representations that purport to

---

<sup>3</sup> This of course presupposes some conception of thought. But it does make perfect sense to speak of artists as subject to rational appraisal (in the sense of that term whose contrast class is *irrational* rather than *non-rational*) – if the poet aims to write a poem about grief and identity and in order to accomplish this only sketches drawings of ukelele’s (while steadfastly maintaining that she wants and plans to write a poem), then we have reason to think of her as behaving irrationally. Likewise with the akratic poet who never sets pen to paper despite her avowal to write her masterpiece.

<sup>4</sup> My primary reason for framing the issue of the normativity of rationality in precisely *these* terms is strategic (as I discuss later). Korsgaard herself urges that this is the “normative question”, and a domain must answer it if it is to be counted as normative for the person to whom the domain applies. There are different ways to put this question depending on which normative concept one takes as one’s figure: “Why *ought* I to be rational/think rationally”, “Why *must* I think rationally”, “Why is it *good* to be rational and *bad* to be irrational”? These questions have different import, and I do not mean to suggest that they can be substituted for one another with no change in meaning. The point here is that each asks of rationality why it has normative standing for us. This said, the use of evaluatives rather than directives in the question really doesn’t seem to purport to guide one’s actions or attitudes, at least insofar as one thinks of the evaluatives as non-action guiding. Such a view is not universally held.

<sup>5</sup> The view of practical reason she defends in *Self-Constitution: Agency, Identity, and Integrity*. We will see later that for Korsgaard the two are, ultimately, aspects of a single phenomenon.



count in favor of adopting belief (or performing an action). The activity of reason consists in the construction of a unified conception of the world and the constitution of a unified subject that forms beliefs about that world through the application of rational principles.<sup>6</sup> These principles are constitutive principles of the activity of reason, and so rationality is normative because it is defined by the set of principles we must use to engage in the activity of reason.

In this thesis I will criticize both Korsgaard's conception of the nature of reason and rational principles, and her also account of why they are normative. Although her presentation is not completely developed, I think that it is developed enough to see that it is headed in the wrong direction. This project, then, is primarily negative. It is positive in a few respects however. First, it is in the process of trying to *develop* Korsgaard's view of reason and specify the principles of rationality that we will see the implausibility of her approach. It will also hopefully have an additional positive payoff of illustrating the contours of the discussion of reason and rationality, clarifying what is at stake in them, and show at least one way not to conceive of the domain.

In the first chapter I will lay groundwork for the subsequent discussion, and briefly characterize differing ways of articulating what I call *the domain of rationality* in the literature.<sup>7</sup> Next I will establish a rough taxonomy for classifying accounts of rationality as a specific phenomenon (as *structural views*, *process views*, and *mirroring views*).

Finally, I will clarify the sense of normativity in question.

---

<sup>6</sup> It is a consequence of her view that theoretical and practical reason are really dimensions of a single activity, since it is the application of practical rational principles unifies us as agents. For ease, I will sometimes refer to the activity of reason, in Korsgaard's technical sense, as "rational activity".

<sup>7</sup> Specifically focusing on those who articulate it into the concepts of "substantive reason", "reason" (as a faculty), and rationality (as an achievement). I take this way of describing rationality from Jason Bridges.

The second chapter will be devoted to characterizing Korsgaard's account of reason and rational standards. I will contrast her view to that of John Broome (1998, 2009), whose claims that rationality consists in a set of wide-scope normative requirements, a first-order model of reasoning, and connection between the two.

In the third chapter I will present a number of objections to Korsgaard's account, organized along two axes: against her conception of reason and rationality, and also against her conception of the normativity of these rational principles. After drawing attention to a significant tension between Aristotelian and Kantian approaches<sup>8</sup> to normativity latent in her view, along the first axis I will present the *objection from incoherence* and the *objection from diversity*. Along the second axis, I will criticize explanation of the normativity of these principles by presenting the objections *from the incompatibility of constitutivity and normativity* and *from rule-following skepticism*. As suggested earlier, these objections will be brought out in the attempt to plausibly reconstruct Korsgaard's position. The argumentative strategy is to show that even developing it as charitably as possible, Korsgaard's view is unsatisfactory.

The issue of the normativity of rationality is sufficiently deep and expansive to be given thesis-length treatment alone. But given the close relation between rationality and reason in the history of philosophy, an affinity clearly present in Korsgaard's work, to address one independently is certainly to present an incomplete view of the issue at hand. In Korsgaard's view it is true that a person is rational in virtue of her applying rational principles in the deliberative activity of reason. What is of further concern is in virtue of

---

<sup>8</sup> Although, as we will see, these will not be strictly Aristotelian or Kantian approaches, but rather developments of those approaches.

what it is true that one *should* or *ought* to be rational and apply these principles. This question, we will see, poses a difficulty for Korsgaard's constructivism. Korsgaard has said that normativity of a given domain or subject can be expressed in its ability to provide an answer to the “normative question” of “why should I y?” Adopting this as the guiding question enables us to see whether Korsgaard's own view can answer the normative question, and so satisfy itself. It is, therefore, of signal importance that her account be able to answer it when it is applied to the activity of reason. I will urge that it cannot, because the very possibility of asking whether I ought to satisfy rational principles entails that I engage in the deliberative activity of reason and so adhere rational principles.

Korsgaard writes that in her view,

“...it makes perfectly good sense for us to say, both to ourselves and to each other, that because you believe both P and that P implies Q, you ought to believe Q...it is a reminder of the normative commitments that are constitutive of taking a certain kind of action, in this case mental action – believing something, that is, representing the world to yourself in a certain way”. (2009, 34)

In what follows I will illustrate the line of thought that she takes to justify this claim, and urge that it is mistaken.

## 1 RATIONALITY IN WHAT SENSE?

**1.1** The term “rational” and its cognates are predicated of a wide variety of objects, some of which appear to have little in common with each other. In our everyday talk we speak of rational people, rational plans, rational decisions, rational choice – the sense of the term has as great a variety as the number of nouns that can be modified by the adjective “rational.” In philosophy, the term is most often predicated of attitudes (such as beliefs and desires) and patterns of thought, but is also attached to less traditional objects (such as societies).

Some argue that the concept of rationality is concerned with the relations that obtain between the attitudes of some subject.<sup>9</sup> Specifically, one is rational if one's attitudes realize a certain set of relations or, rather, do not realize some impermissible set of relations – coherence, that all of one's attitudes can be true simultaneously, is a prime example of such a requirement.<sup>10</sup> Others think that rationality is concerned with what one has reason to believe. Specifically, one is rational if one's attitudes track the reasons that there are, or if one has the attitudes that there are reasons for one to have.

Other concepts are closely related to the concept of rationality, and must figure into a presentation of that phenomenon. Specifically, these are the concepts *substantive*

---

<sup>9</sup> Scanlon (1998)

<sup>10</sup> It is also, if strictly construed, far too stringent. The amount of intellectual energy needed to root out *all* contradictory attitudes is greater than most people who would intuitively count as rational can muster. A weakened requirement would be preferable, namely one that says that one endeavors to maximize the coherence of one's attitudes when possible.

*reason*, and *reason* as a faculty. For sake of clarity, I will refer to the concepts of substantive reason, reasons, and rationality collectively as *the domain of rationality*. This will help keep matters clear when discussing the relationship between the domain of rationality and normativity, and specifically the differing explanatory approaches adopted by philosophers regarding the two domains. It is important to adopt such terminology insofar as many philosophers analyze normativity using the concepts that belong to domain of rationality.

**1.2** I will begin by introducing the concept of *substantive reason*. At the most general level what I will call a “substantive reason” is typically understood to be a consideration that counts in favor of some action or attitude.<sup>11</sup> This use of the word needs to be distinguished from the term “reason” as it appears in sentences of the following sort: “the reason that you have such back pain is the pressure on your sciatic nerve”. That use of the term is broadly explanatory, and in a great many cases can be substituted for the word “explanation”. As a result it is typically called an “explanatory reason”. What explanatory reasons consist in is a matter of dispute, but most argue that they are either propositions or facts.

The sense of “reason” that is of primary concern here is different from this explanatory sense. Its character can be seen in sentences of the following sort: “but you had no reason to mislead him about the contents of the will”. This is this sense of reason that is typically glossed as “counting in favor”. It purports to *guide* one in action or in the

---

<sup>11</sup> I follow Korsgaard's terminology here. Scanlon (1998) urges that we have reasons for attitudes only; Bridges objects that we can have reasons for actions as well as attitudes. (2007)

acquisition of attitude, and in accordance with this intuition is sometimes called a normative reason<sup>12</sup>. But, as has been noted often, normative reasons also do explanatory work<sup>13</sup>. In saying why my brother bought a gift for our mother I cite his reason for doing so – it was her birthday. That it was her birthday was a reason *for* him to buy her the gift to be sure and so was a normative reason, but citing it as his reason *explains* his action.

It should be noted that one can cite what is perceived by an agent to be a normative reason for action or attitude to explain her performance of that action or adoption of that attitude while not seeing the normative reason that the individual takes herself to have *as a reason*, all things considered. My brother might have thought that he had reason to buy our mother a gift on those Thursdays that he trips over his bedroom slippers and drops his toothbrush in the toilet; as it happens, *this* Thursday both of these conditions obtained and it also happened to be our mother's birthday. So we *explain* his action by citing his perceived normative reason, yet do not grant that his reason is a good one. This is to say that the phrase “for the reason that” which appears in explanations of the sort need not be factive. One may act on what one takes to be a reason, but be mistaken about whether it is a reason.<sup>14</sup>

The concept of a substantive reason is believed by some to resist analysis or eliminative definition.<sup>15</sup> Attempts to define the concept often invoke concepts that are sufficiently similar to substantive reason as to be uninformative, or concepts that can

---

<sup>12</sup> The nature of the relation between explanatory and normative reasons (or, alternatively, between explanatory and normative *uses* of the word reason) is contested and not as neat as this crude sketch suggests: most obviously, what we are inclined to call normative reasons *also* explain, e.g. “the reason he made dinner at home tonight was in order to feed his children.”

<sup>13</sup> Raz (2007)

<sup>14</sup> See Setiya (2003)

<sup>15</sup> Scanlon (1998), Raz (2007)

themselves be understood only in terms of substantive reasons. For example, Scanlon (1998) urges that a reason is a “consideration that counts in favor” of something. (1998, 17) But when one wants to know *how* it counts in favor of that thing, one will be able to say little else other than “by being a reason for it”. Scanlon does think we can say what sorts of things reasons are, though, even if we are unable to provide an illuminating analysis of the concept: they are a four-place relation that figures into normative judgments. This sort of primitivism about substantive reasons is not widely held, though.<sup>16</sup> Joseph Raz, for example, urges that substantive reasons are facts that have a characteristic function, namely to “provide[s] the answer to questions about the reason why things are as they are, become what they become, or to any other reason-why question...”. (2007, 2) On this view, substantive reasons have an essentially explanatory purpose – all substantive reasons are explanatory, but some are explanatory as well as normative. Since he takes substantive reasons to be explanatory, Raz rejects the notion, defended by some, that reasons are propositions. This is because propositions can be false, and false propositions do not explain. So if reasons were propositions then we would be in the uncomfortable position of having to show what false propositions explain or justify, or else provide some ground to count only true propositions as reasons.

For Raz, a normative reason is “a fact which gives a point or a purpose to one's action, and the action is undertaken for the sake or in pursuit of that point or purpose” (2007, 1) Scanlon would perhaps reject this suggestion because it seems possible for something to give point or purpose (either from the agent's perspective or objectively) to

---

<sup>16</sup> Scanlon offers “buck-passing” accounts of other normative concepts such as “good”, meaning they are to be analyzed without remainder in terms of reasons.

an action without being a reason for it.<sup>17</sup> What is to the point, though, is the fact that both Raz and Scanlon are realists about reasons. This means, minimally, that normative facts or truths about what there is reason to believe or do exist independent of our intellectual practices. It is not a condition on realism of this sort, though, that reasons are absolutely independent of reasoners in the sense that if there were no reasoners there would still be reasons. Most plausible forms of realism construe reasons as relations that obtain between people and attitudes or actions that do not persist in the absence of one of the relata.

Intuitively, the concept of a substantive reason is one of the fundamental normative concepts, but it has company: *good*, *bad*, *ought*, *should*, and *must* are all normative terms that pervade philosophical as well as everyday discourse. Judith Thomson has drawn attention to a general distinction between two kinds of normative concepts, the *evaluatives* (of which *good* and *bad* are examples) and the *directives* (of which *ought*, *must*, *should* are examples). It will be useful to keep this distinction in mind because philosophers of normativity often urge that one kind is primary and that the other can be analyzed in terms of it.<sup>18</sup> The normative concepts at work in our guiding question will be 'should' or 'ought'. As will be seen later, Korsgaard disagrees with Scanlon and Raz that substantive reasons are the fundamental normative concept, or at least urge that they are not the sole fundamental normative concept.

---

<sup>17</sup> For Scanlon (2009) reasons are a relation that forms the essence of normative judgment. "R (p,c,a): p is a reason for an agent in circumstances c to do a".

<sup>18</sup> Of those who give priority to the evaluatives: Richard Kraut (2007), G.E. Moore (1903, 1966), J.J.T. Thompson (2008), Talbot Brewer (2010).



**1.3** With the notion of substantive reasons in view, we must now turn to *rationality* and *reason*. On the one hand, to say of something that it is rational or it has reason is to say of some entity that its mind has the property of being active in a way that other entities are not. This is usually thought of as a faculty or capacity that distinguishes some entities, namely humans, from others. Humans, the line of thought goes, are distinctive by virtue of the fact that they have the *faculty* of reason. There are a few activities that intuitively fall under the scope of this faculty. Historically, drawing inferences and computing quantities have been identified as activities conditioned by this faculty or capacity and distinctive of it.

Perhaps as a more fundamental level, the capacity to respond to reasons *as reasons* has been introduced as that ability which distinguishes rational creatures from non-rational creatures, and is the core of this faculty sense of reason.<sup>19</sup> We need not argue that humans always respond to reasons as reasons, but only that at least some humans have the *ability* to do so. And there is reason to doubt that any non-human animals respond to reasons *as reasons*. The contrast class of *reason* and *rationality* in this sense is *non-rational*. I will call this faculty-sense “reason” to distinguish it from “substantive reason”. We will see later that this general characterization of the phenomenon is not congenial to Korsgaard— the faculty of reason is not aptly characterized as a *responsiveness* at all, in her view.

But there is also *rationality* in the sense of that term which has *irrational*, rather

---

<sup>19</sup> McDowell (2009) For example, a Scottish terrier might respond to the charging mastiff by scooting under the porch where mastiff cannot fit. It seems natural to say of this event that the mastiff was a reason for the terrier to scoot under the porch, and that the terrier responded to the reason. But we would not say that the terrier responded to the reason as *what it was*, namely, a reason.

than *non-rational*, as a contrast class. Only rational creatures in the first sense can be rational in this second sense. Rationality in the second sense signifies a positive achievement, and something earns that status just in case it is able to conform to a specific set of conditions. Typically these conditions are taken to be rules or requirements. What is included in this set of rules is a matter of surprising agreement; some usual suspects are the rules of logic (both deductive inference and the probability calculus), evidentiary standards, and canons of belief formation.<sup>20</sup>

Since some draw substantive distinctions between the nature of rules and principles, I will call them the *standards* of rationality to remain as neutral as possible: to call them rules or principles of rationality runs the risk of presupposing a contentious view of their nature that should be settled only by substantive debate.

We can introduce a rough distinction between *structural views* of rationality, *process views* of rationality, and *mirroring views* of rationality. The first suggests that one is rational if and only if one's attitudes and actions instantiate a specific structure, articulated in terms of the relations of one's attitudes to one another. John Broome's position, which will be discussed at length later, is an example of this view. A *process* view suggests that one is rational if and only if one adheres to some process in acquiring or jettisoning attitudes and in determining whether to perform (or not perform) some action. Mirroring views, such as that of Gibbard, suggest that one is rational if and only if one's attitudes and actions are the attitudes and actions that one has most reason to have.

---

<sup>20</sup> In Rational Choice Theory an ideal or standard of rationality is a complete ordinal preference structure (this is a standard of rationality relied on heavily by economists). I think this actually not a standard of rationality – surely one should not be counted as irrational for not having an ordinal preference between two equally horrific choices.

Each of these general views can be expressed in terms of a set of standards or requirements. An example of a structural requirement (with a wide-scope) defended by John Broome is *modus ponens*: rationality requires that (if you believe  $p$  and if  $p$  then  $q$  you have attitude  $q$ ). Another widely held standard of rationality that is strongly structural can be found in so-called “rational choice theory”, where an individual is considered rational only if her preferences have a complete ordinal structure.

An example of a general process view of rationality can be found in Ayer's *Language, Truth, and Logic*: “to be rational is simply to employ a self-consistent accredited procedure in the formation of all of one's beliefs...” (1952, 100) where a procedure counts as accredited just in case it is at that time taken with good reason to be reliable (which is revisable).

There are of course positions within this general scheme that can be plausibly described as belonging to more than one view. If adopted in pure form and taken in isolation, each of these three approaches would perhaps be implausible (counter-examples to each are ready to hand). It is more useful, then, to think of each as *dimensions* in accordance with which a given view of rationality can be understood. Bayesians, for example, believe that one is rational just in case one applies Bayes' Rule to update the prior probabilities assigned to one's beliefs in light of new evidence. For the Bayesian, rationality *just is* cleaving, through reiterated application, to Bayes' Rule. This conception of rationality integrates both structural and process dimensions: the assigning of prior probabilities to some beliefs will require that the probabilities assigned to other beliefs be adjusted. The Bayesian position regarding rationality is therefore geared towards the relations that obtain between our belief (namely the relationships between

their prior probabilities) and so has a significant *structural* component, but insofar as it requires application of Bayes' Rule in the forming of beliefs, also has some degree of *process* involvement as well.<sup>21</sup>

**1.4** An additional point must be noted. The guiding questions about the normativity of rationality include “why is it wrong to violate *modus ponens*?” or alternatively, “why ought one to adhere to *modus ponens*?”, where *modus ponens* is a specific instance of a rational standard. It must be borne in mind that the question of the normativity of rationality at issue here is distinct from the idea that rationality has requirements that one can *fail* to satisfy. That one can intelligibly fail to be rational (not just fall in a contrast class) shows that rationality is normative in some *minimal* sense. But that this is true does not vindicate the normativity of rationality in the more robust sense in view when we ask “why should one not violate *modus ponens*” and “why should I satisfy *modus ponens*?” Consider an analogy that brings this out (and that is also indicative of Korsgaard's general approach).

Chess has rules. But the rules of chess are special: they are also constitutive standards or requirements. For example, it is a rule of chess that a player cannot legally move her knight diagonally across the chessboard; neither can she attack by forward movement with a pawn. But this rule of chess is also a standard of what chess *is*. Chess is *defined* by the set of its rules. It is not the case that one is playing chess *badly* in moving the knight diagonally, but rather that that move is not a chess-move at all – one is not

---

<sup>21</sup> Weberian *Zweckrationalitat* is a rational requirement which says that one should always maximize the results one is after while minimizing the cost to oneself. This is an instrumental rational requirement.

playing chess in that case. This of course has nothing to do with the medium in which the particular game is realized. One can play chess with shoes or stones or ravens or writing desks: the game itself is matter-invariant.<sup>22</sup>

The same is true of some other games. A batter cannot legally bypass first base by running past the pitcher straight to second. Games have requirements that are standards of correctness that are constitutive of the game itself. We can think of the moves in chess as normative in the following, minimal sense: one can *misplace* a chess piece. This can be contrasted with, say, simply throwing stones as a way to pass the time. I might hit a walnut tree, or I might hit the rhododendron on the other side of the road, and so on. But there is no way to *fail* at throwing stones in the same way that one can fail in moving a chess piece.

Rules in this sense are *a priori* regulative of a practice, and if some move fails to conform to the rule it is not the rule that is discredited but rather the move. The rule remains in force even in particular cases in which it is violated. So too with positive laws. Their direction of fit is “mind to world” - if one violates a traffic law one is not entitled to say “well I suppose it was not really a law after all”. This is different from an empirical generalization, of which the so-called laws of nature are paradigmatic examples. In these cases, if the behavior of the entities described by the law violate that law, then we are licensed to say that it was not really a law at all - it does not remain “in force”.

But it doesn't seem to be true that the requirements of chess are normative in a

---

<sup>22</sup> Searle (1996) makes this distinction in terms of convention and rule. It is a convention that chess is played using pieces of a particular shape and material constitution. It is a constitutive rule, however, that, for example, one must advance one's knight along an L-shaped path (and likewise with other rules).

sense that would underwrite the following command “you *ought* to move your knight in a L-shaped path” even if you are not playing chess, and certainly not “one ought to play chess”. One is not under any requirement expressible in the language of “ought” or “should” to satisfy the rules of chess if one is not playing chess. The same is true of baseball: it is not the case that I ought to run to first base if I am not playing baseball.

The point appears to generalize. What is it to say that I ought to satisfy *modus ponens*? If I am not interested in thinking rationally, the line of thought goes, then I am under no requirement to satisfy *modus ponens*. It is precisely this line of thought that has hamstrung so many theorists about the normativity of rationality.<sup>23</sup> We have seen a short version of Korsgaard's answer to this problem: if there is an analogy to be drawn between games and rationality, it is that the standards of rationality are necessary to engage in deliberation at all. How she comes to this answer, and whether she gives it adequate support, is the task ahead.

As an example of how rationality as a set of standards can be conceived such that they are normative in the minimal sense just mentioned, consider Rawls' distinction between two sorts of rules.<sup>24</sup> His concern in that famous paper is to show that the utilitarian principle has application only to the justification of practices as a whole, not to the particular actions that fall within a given practice. These actions are of a different sort. Insofar as they fall under a specific sort of *rule* they have a different structure from the sort of action that is the concern of utilitarian theory.

A practice rule is a rule that is “logically prior” to the cases to which it applies.

---

<sup>23</sup> Some, Kolodny (2005) for example, simply reject the idea that rationality is normative.

<sup>24</sup> Searle (1969, 1995) offers a view of constitutive rules as well, and his view will be discussed below.

Rawls' example is baseball: baseball features a set of rules that define specific offices and the actions that are appropriate to them, stations, movements, and penalties for breach of rules. He says that the logical priority of the practice rules is this: "given any rule which specifies a form of action (a move), a particular action which would be taken as falling under this rule given that there is the practice would not be *described as* that sort of action unless there was the practice." (1958, 25) Contrast this with the *summary rule* that I have to always put a small amount of the salted water in which I have boiled my pasta into the sauce which I have prepared to accompany it. I may of course fail to do this and remonstrate myself because the finished product is inferior to what I would have produced had I followed my rule. But I can boil pasta and make a sauce to accompany it even if I fail to follow this rule. But the same is not so with baseball or chess. There is no such thing as an 'out' or a 'triple play' if the rules of baseball are not in force. Switching the place of my king and rook is not castling unless the practice rules of chess are in force. Indeed, the pieces go by those names in an empty sense if the rules of chess are not in force. Innumerable many objects can be a token of the type "king in chess", and likewise with the other pieces.

These differs from what Rawls calls summary rules. Summary rules are logically posterior to the cases to which they apply. My adding salted pasta water to the sauce, or inserting garlic cloves between the skin and meat of a chicken I plan to roast, are summary rules. Rawls says that the point of having rules at all is that similar cases recur and one can get along better by keeping past decisions in mind such that one can more efficiently decide in a case that has some precedent. But some cases may exist even if there is no rule to apply to it (because it has no precedent). Put generally, Rawls says that

what A's and B's refer to in rules of this sort: “Whenever A do B” may be described even if this rule does not exist, or if there is no practice in which the rule has a place.

In the case of practice rules, however, the actions-types that those rules inaugurate are subject to constitutive constraints. If I intend my moving the knight to count as a moving of the knight then I must move it in an L-shaped path only. Tamar Schapiro uses the example that in order for a judge to punish a defendant she must adhere to the “deliberative constraints of our legal system”, which require that punishment can be meted just in case a judgment of guilty has been delivered. (2001, 101) A judge who delivers a punishment on some other basis, the maximization of expected utility for example, is not issuing a *punishment* at all, but rather doing some other thing. The standards or rules of the practice introduce offices and moves that are subject to constitutive constraints.

This view of rules is naturally suited to a constructivist interpretation.<sup>25</sup> When asking what it is for the rules of chess to be “in force”, the constructivist will respond that this means that some set of agents have committed to playing the game or, more strongly, that they are authors of the rules of the game. Some will go on to say that this is why these rules are a priori regulative of the game, practice, or activity in question. The realist about normativity would demur on this point, urging that the rules of the game, practice, or activity are independent of any mind or, to put a finer point on it, the volitional activity of any agent. Of course one will not likely encounter a realist about the rules of chess. But realism about rational standards such as *modus ponens* is a more plausible, and widely held, position.

---

<sup>25</sup> Not surprisingly, since its first systematic presentation was given by Rawls.



Keeping Rawls in view is helpful for framing Korsgaard's conception of reason as a *rule* or *principle-governed activity*. Korsgaard is of course not a realist, but an avowed anti-realist about rational activity. This can be traced to her deep affinity with Kant, whose central point is that reason cannot be construed as a merely *responsive* phenomenon in the way that the realist suggests and still count as free. Rawls' analysis of practice rules is homologous to the extent that these rules are conditions on the existence of the entities to which they apply. Likewise, for Korsgaard rational principles are conditions on the existence of the fundamental normative entities called substantive reasons and rational agents as well.

**1.5** Insofar as the general question at hand is “why is rationality normative”, at least a few words are necessary regarding the nature of normativity. Theories of normativity take many different guises, but they can be given the general characterization of attempting to define a set of normative concepts, illustrate how they are related to each other, account for how they are deployed in judgments, and situate them in our experience more generally. In the course of constructing a theory that encompasses these aspects one will have given a metaphysics and epistemology of the normative as well.

Theories of normativity can be most easily distinguished by the normative concepts they take to be fundamental. Some begin with the concept *good*, others with *substantive reason*, and analyze other normative concepts in terms of their chosen fundamental normative concept or concepts.<sup>26</sup> Others aim to provide reductive semantic

---

<sup>26</sup> Judith Jarvis Thomson (2008) is an example of a good centered normative theory, Joseph Raz is an example of a substantive reason centered theory.

analyses, without remainder, of normative language (presumably taking this to be sufficient for reduction of normative concepts as well).<sup>27</sup>

The question of the normativity of rationality is complicated for some approaches, especially the two of concern below. Korsgaard's and Broome's accounts of rational standards *belong* to their theories of normativity – that is, the explanation of normativity invokes rational standards. In the case of Broome, the domain of normativity itself can be parsed into ought-facts, reasons, and normative requirements (what rationality consists in). Given this, he expresses the question of whether one should adhere to normative requirements in terms of whether there is any substantive reason to satisfy requirements of rationality individually or collectively. For Korsgaard, as we shall see, the application of rational standards ushers normativity into existence, and so telling a story about why one *ought* to adhere to them becomes even more problematic. In both cases, normativity as a general domain is understood to include, at least in part, in rational standards. But Korsgaard's position is more radical because she says not that rational standards are among the things that are normative for us, but rather that what it is for anything to be normative is to be explained in terms of rational standard, specifically the role they play in deliberation. This has the consequence of making normative appraisal of individuals who adhere or stray from rational standards a complicated matter.

Of course not just any proposition in which these concepts appear are of concern here, but rather those which would be given, as mentioned above, as advice, recommendation, remonstrance, or alternatively those which can be given expression as a question: “why should I satisfy modus ponens?”. To answer this question, we will tell a

---

<sup>27</sup> Stephen Finlay (2009)

story that says in virtue of what it is true that “one ought to satisfy modus ponens”.

## 2 KORSGAARD'S VIEW

**2.1** This section will use the idea of constructivism to frame and situate Korsgaard's position, and then use it as a touchstone in characterizing her conception of the activity of reason. Next I will discuss Korsgaard's account of constitutive principles and standards, explaining how she takes herself to build upon Aristotle's metaphysics of normativity. After providing a few doubts about the plausibility of her interpretation and appropriation of Aristotle, I will then briefly characterize the conception of rationality and reasoning defended by John Broome in order to bring the features of Korsgaard's view into relief. I will note some signal ways in which Korsgaard's view is superior to Broome's, and finally say a few words about differing approaches to the normativity of rationality available to Korsgaard in light of her position.

To locate Korsgaard's position in the conceptual terrain it is useful to have a pure constructivist account of rational standards in view. This is useful because her position is best described as a kind of constructivism (although she seldom claims that moniker for herself). A constructivist conception of rational standards would represent these standards as the products of some procedure performed by a person or set of persons. They are understood to apply to, and be normative for, a person just in case she performed the

procedure.<sup>28</sup> A constructivist about normativity, by contrast, would represent the fundamental normative concept/s or entities as the product of some procedure of construction, and it would represent facts and truths about what one has reason to do or think as issuing from the principles that define this procedure. Constructivists and realists about normativity share the conviction that there are normative facts, or normative truths. They differ in how to account for these truths: realists urge that there are truths about what we have reason to think or do because there are independent normative properties that make normative judgments true; constructivists represent the truth of normative judgments as a function of commitment to the principles that determine them. In Korsgaard's view, these principles are constitutive of deliberative rational activity.

Of course, not all constructivists are the same, and the position is not without problems. In a case where the standards of rationality are to be understood to obtain and apply to someone just in case that person can be represented as having selected them in an appropriately structured position (one ideally structured for delivering rational standards, say), such as Rawls' original position, it is impossible to subject the behavior of the persons selecting rational standards to rational appraisal. And whereas it might be acceptable for the behavior of contractors in contractualism or constructivism about principles of social distributive justice or morality to not be evaluable in terms of justice or morality because they can be evaluated in terms of rationality, it is rather more unsettling to be saddled with a view that excludes rational appraisal of someone selecting

---

<sup>28</sup> This can be given a contractualist construal as well.

principles of rationality.<sup>29</sup>

Constructivist and contractualist approaches are often subject to a similar line of critique, one presented by Hume<sup>30</sup> and echoed (although differently inflected) by Ronald Dworkin.<sup>31</sup> That is, such approaches either suggest that a procedure of construction or contract actually took place which is implausible on empirical grounds, or understand it as “metaphorical” or as a “device of representation” in which case it is difficult to say why one is bound by it.

Constructivists are apt to reconstruct this apparatus so that it is not vulnerable to objections of this general sort. Korsgaard does precisely this, I think, by treating the exercise of reason in deliberation as the reconstructed procedure of construction. Deliberation itself is the so-called procedure of construction. Hers is a constructivism about normativity rather than rational standards: *substantive reasons*, and not rational standards, are the product of construction. Rational standards have a crucial role in her view, however. They are constitutive principles of the procedure of construction, which

---

<sup>29</sup> It is of course possible to represent the contractor as adhering to some loosely conceived notion of reasonableness or rationality in order to select *clearly defined* principles of rationality to govern thought. It seems likely, however, that this loosely defined conception of reasonableness or rationality will be quite similar to the supposedly different clearly defined principles of rationality understood to issue from a procedure of construction once cashed out. Constructivism often runs the risk of smuggling the products of construction into the procedure of construction or characterizing those who perform the procedure.

<sup>30</sup> Hume's criticism actually runs deeper than this, and indicts the authority of promise-keeping or fidelity itself: in his view, that one has made a promise does not have some special normative authority over oneself. The Kantian vindication of the inviolability of promise-keeping depends on a conception of human reason and cognition that Hume rejects. See “Of the Original Contract”. Collected in *The Broadview Anthology of Social and Political Thought: From Plato to Nietzsche*. Broadview Press. Buffalo: 2008

<sup>31</sup> In a review of Rawls' *Theory of Justice* collected in *Taking Rights Seriously*. Harvard University Press. Cambridge: 1978

she calls the activity of reason. We will see later that Korsgaard's answer to the guiding question will be in terms of the relation each person has to this procedure of construction. Specifically, Korsgaard reconstructs the procedure of construction as deliberative activity of reason *itself*.

**2.2** Korsgaard's understanding of rational standards and substantive reasons belong to her account of the nature of reason, and so it is fitting to begin with the necessary condition of this activity: reflective awareness. The source of reason, in Korsgaard's view, is the “reflective distance” that arises when one becomes aware *that* one is aware.<sup>32</sup> This is unquestionably inspired by the notion of “apperception” that constitutes the core of Kant's critical turn. Although she does not put it quite this way, the reflective distance is best construed as a property of awareness. This property accrues to awareness when that awareness meets at least a few conditions: it must be apperceptive (for only in being aware of itself can it be aware of its attitudes *as* attitudes, and so as objects in some sense independent from itself); it must have a conceptual repertoire sufficiently rich and expansive to represent itself, its attitudes, and the authorizing power of endorsement. It is natural to give expression to the idea of this property as the awareness one has *of* the deliverances of one's senses *as that*: one becomes aware of one's seeings *as* seeings (and likewise with the other perceptual media). But most important, one becomes aware of the “potential grounds of our beliefs as *potential grounds*” (2008:18). The phenomenon can

---

<sup>32</sup> Discussed at length in Korsgaard (1996)

be formulated in the following way

*Reflective distance*: the property of awareness that enables it to capture the fact that representations typically cause one to form attitudes or perform actions in a way that enables one to evaluate these attitudes and actions in light of the representations that typically cause them.

Korsgaard sometimes construes this as the “capacity to distance ourselves from” these attitudes and actions and “call them into question.” (1996: 93)

Korsgaard highlights the character of this awareness by comparing it to the awareness of the non-human animal.<sup>33</sup> *Non-human* animals are guided almost entirely by their perceptual capacities and instincts. Non-human animal awareness is, in Korsgaard's terms, “teleologically organized”. (2008: 19) In order to see what she means by this, consider a few examples. For the hound, the squirrel is *something to be chased*; for the squirrel, the acorn is *something to be picked up and cracked*.<sup>34</sup> Animals that are intellectually more sophisticated may have an awareness *that* their peers regard certain organisms or events as objects of aversion or desire, but even if this is so there is good reason (Korsgaard urges) to doubt that these animals have the capacity to step back from their teleologically ordered world and ask, to stay with the example, *whether* the acorn

---

<sup>33</sup> Korsgaard pursues the same line in (1996).

<sup>34</sup> It is not clear just what judgments Korsgaard thinks these observations about the way the world is disclosed to non-human animals licenses. The natural-historical judgments, or Aristotelian categoricals, such as “the squirrel gives birth to two young in early spring”, that Michael Thompson discusses are to the point here. Thompson argues that such judgments are logically distinctive, and cannot be captured either as universally quantified propositions or as statistical generalizations. Would Korsgaard endorse a similar judgment regarding the awareness of non-human animals, such as “The squirrel cracks nuts”? But notice that this judgment, insofar as it neglects to mention the character of squirrel awareness and focuses only on what the squirrel does, fails to capture Korsgaard's point. Is “the squirrel encounters the nut as something to be cracked” any better? This seems to be a rather fantastic claim.



something that *should* be cracked. Nevertheless, she urges, these animals represent their environment to themselves by their perceptions, and these representations *cause* the animals to do what they do. Their environment is represented as having a teleological order, she thinks, not just insofar as their perceptions operate on them causally (for surely our perceptions also have a place in a causal nexus) but because they lack the special form of awareness characteristic of humans.

There was probably a time in our own evolutionary history in which our species, or some species that was our immediate ancestor, had awareness of this sort. It is an awareness in which objects are disclosed as having roles, and *that* they have these “roles” is not a feature of the awareness of the animals who respond differentially to the objects or events in question. But at some point along the way we developed the ability to be aware *that* we are seeing, smelling, and tasting. We also became aware *that* these perceptions, to stick with that modality, affect us in characteristic ways. For example, *that* the sky grows dark and the leaves show their underside typically leads one to seek out shelter. Korsgaard urges that it is a feature distinctive of human awareness that we are conscious of the ways in which our experiences typically activate our instincts, desires, and aversions, and that this consciousness is the source of reason. For as a result of it we have no option, Korsgaard maintains, but to ask *whether* the representations that cause our beliefs and action “really are grounds” for our attitudes or actions.

So much for the idea of reflective distance. It is clear that in Korsgaard's view that the reflective distance is very closely tied to the concept of deliberation. Indeed, it will

come out below that the activity of reason itself bears such strong resemblance to the notion of deliberation that it is sensible to equate the two. Next, it is necessary to characterize how substantive reasons are created as a result of deliberative activity.

Korsgaard suggests that the concept of a substantive reason and the concept of cause are two species of a genus – the species have in common that they are what Aristotle called *aitiai*, or “because”.<sup>35</sup> Following Kant, Korsgaard calls this genus “ground”, and says that it is individuated by the fact that a ground is an answer given in response to a “why question.” Earlier she suggests that the perceptual representation of its environment had by an animal that causes the animal to behave is an example of a ground. Korsgaard is ambiguous regarding whether the representation in question is a ground according to *us* or according to the *animal*. (2009, 19) Presumably she means the former.

Most generally, we can say that the genus *ground* is a form of representation that could figure into an answer to a “why question”. When someone asks “why did the uprights shudder when struck by the football” one can answer that the football *caused* them to shudder; when someone asks “why did the kicker kick the football toward the

---

<sup>35</sup> What are the differentiae between the two species of this genus? Korsgaard is less than pellucid on this point. It is helpful to consider the relationship between species and genera in general. On the common view, a species can be defined in terms of the genus to which it belongs together with a differentia. Horse is a species of the genus animal that is a quadruped, mammal, and so on. In like fashion, Korsgaard urges that reason “might be defined as a ground that has been endorsed by the person who believes or acts.” (2009, 9) This might seem to have as a corollary that the definition of cause is a ground that has not been endorsed. But this is too quick: a cause might be a ground that has been endorsed and has some other differentiating property. It needn't be the case that a cause is distinguished from a reason by lacking the latter's differentiating property.

uprights” one can answer that his *reason* was to score a field goal. This understanding of ground has something in common with Raz's suggestion that a reason is “whatever provides the answer to questions about the reason why things are as they are, become what they become, or to any other reason-why question...”. (2007, 2) Of course Raz does not say that reasons are *only* whatever is given to a why-question – they have to be *facts*<sup>36</sup> and, as we saw earlier, Raz thinks that all reasons are explanatory since it is probably the case that all facts can figure into some explanation (but some are both explanatory and normative).

The anti-realism of Korsgaard's position is evident: the concepts of *cause* and *reason* alike are species of a genus of representation deployed by rational entities and do not pick out features of a mind-independent reality, such as “facts” in the realist. This means that it is not the case that there are reasons, rational standards, or normative entities of any kind in the world which, fortunately enough, rational principles enable us to identify. It is rather to say that reasons are *created* in being picked out by rational principles that we apply. Normative entities are products of the volitional activity of agents and this activity is to be explained in terms of deliberative rational activity (which can in turn be parsed into theoretical and practical aspects).<sup>37</sup>

A representation counts as a *potential* ground in virtue of being the object of the reflective distance. What was before only a representation that caused one to believe thus

---

<sup>36</sup> Raz takes facts rather than propositions to be explanatory reasons (since false propositions explain nothing)

<sup>37</sup> See R. Jay Wallace (2010) for a helpful discussion of constructivism about normativity generally.

and such is put into question by our awareness, and becomes a potential ground. It becomes a 'ground' as a result of this awareness, presumably, because this awareness brings with it the possibility that a 'why-question' of the relevant sort could be given. The awareness interrupts the causal relation between the representation and that attitude or action to which it leads. If we endorse its operation on us, then it becomes a reason. Suppose I see the mouse and, reflecting on the many diseases I believe rodents in general to host, I climb onto my couch. Counterfactually, had I lacked the awareness in question but nevertheless formed the belief or performed the action called for by the representation, then that would have been a cause (and a ground too, given that *someone* could provide that representation as an answer to a why-question) of my behavior rather than a reason for it.<sup>38</sup> But suppose that after coming into the room you ask why I am standing on my couch. When I excitedly answer that “there's a mouse!” because I am afraid of mice and I see one, then the mouse is the *reason* why I am on the couch. Here, my representation of the mouse is a *ground* (given the role it plays in answering a why-question) as well as a reason (since I have endorsed the representation's “operation” on me).

It is not, of course, that every cause of one's attitude or action is transformed into a reason as a result of human awareness. In some cases this is precisely what does not occur – instead of seeing the potential grounds of our desires or beliefs as reasons for

---

<sup>38</sup> Whether we can intelligibly speak of having belief in the relevant sense without apperceptive awareness is a matter of debate. In this case we can say that the belief in the absence of this awareness would be a representation that portrays the mouse as something to be evaded.

them, we will in some cases see them precisely not as reasons (by not endorsing them). If we decide that some potential ground is *not* a reason then one has interrupted the causal relation. Representations that would have exercised a merely causal influence are brought into the sphere of *rational candidacy*, to coin an ugly term, by this awareness. They are treated as the sort of thing that *can* figure into rational justification. She writes that “what would have been the *cause* of our belief or action...now becomes something experienced as a consideration in favor of a certain belief or action instead, one we can endorse or reject” adding that only if we are able to “endorse the operation of a ground of belief or action on us *as a ground*” (2009, 21) do we take the consideration to be a reason. To repeat, the consideration that *purports* to count in favor of belief becomes such by being the object of reflective awareness, and it becomes a consideration that *does* count in favor of a belief, a reason, only by our endorsing its operation on us.

So although this is a view of reason in the faculty or capacity sense, it is not a capacity *only*. It is a capacity that is continually actualized in the form of an activity, namely deliberative activity. The activity of reason can be construed as the procedure of construction, and we now have one side of this procedure in view. We see, that is, that reasons are the products of the procedure of construction. When the procedure is completely specified we shall have Korsgaard's conception of the normativity of rationality in view, and will therefore see her answer to the question why the individual standards of rationality are normative - why it is wrong, to stay with the example, to violate *modus ponens*.

**2.3** The reader will likely have a few questions at this point. On what basis we endorse the “operation” of a representation (potential ground) on us? And why does our endorsing a potential ground's operation on us make it a reason? What is it to endorse the operation of a potential ground? With reasons not yet on the scene Korsgaard cannot entitle herself to the claim that we endorse a representation *as* a reason on the basis of some *reason*. This is, of course, the same worry mentioned earlier that confronted pure constructivism about rational standards: put generally, the constructivist cannot coherently entitle herself to the products of the procedure of construction prior to the procedure. This is a difficulty because, often anyway, the products of construction are fundamental and very important, and it is troubling to be unable to apply them to contractors.

The second question also looms. How does one pick out which representations, which potential grounds, to endorse? Since reasons are unavailable one has only competing potential grounds (representations that typically effect one in a characteristic way) available and, moreover, these are all on the same footing. It is not the case that one potential ground has normative authority that another lacks, since concepts of normative evaluation are to be constructed. One proposal would be to say that one endorses a potential ground in case one finds it pleasurable, or one has desire for it. On this view, to “endorse” might simply be to express a pro-attitude about some ground. It would also be a wildly untenable conception of theoretical reason to say that reasons are the products of the expression of pro-attitudes about representations that activate one in

typical ways. This is not Korsgaard's view. She argues that we endorse some potential ground as a reason instead of a competing potential ground through the “application” of rational principles, and that these are constitutive of the activity of reason. Rational standards, then, are the rules that govern the procedure of construction: deliberation.

Described functionally, the activity of reason consists in two tasks. Each individual with the ability to achieve the form of awareness just mentioned must both “construct a new way of conceptualizing the world” and “construct or reconstruct” her own nature as a subject of the conception of the world. (2008, 23) An obvious objection is to ask which task has priority: if one must construct a new conception of the world then there must be some “self” or “subject” in place to do the construction; but Korsgaard herself says that the self must *also* be constituted so that there is a subject for the reconceptualized world. If we take Korsgaard literally, then it seems that the project can never get started. This initial worry will be addressed later in the form of an objection to her view, but it can be set to the side for now.<sup>39</sup>

The phrase “figure into” still fails to specify how these principles belong to the activity opaque. Earlier, Korsgaard says that the principles of rationality *describe* the activity of reason. What is it to say that a principle describes, or is constitutive, of some activity? To begin to see her answer consider an image she provides that not only

---

<sup>39</sup> One of course engages in many doings automatically, without deliberation. Korsgaard might try to accommodate this sort of point by construing such automating doings as the continued or habituated products of some episode of deliberative rational activity. We will consider later whether this response is satisfactory.

colorfully represents the deliberative stance:

When we become aware that we are representing the world to ourselves, when we turn our attention away from what we perceive and onto the fact that what we are doing is perceiving, then there is a way in which the world loses its unity. What was once simply given to us as the environment is now given to us as a heap of perceptions, or rather experiences, and it is now up to us to put them back together into a picture of the world. And in a similar way where once upon a time we always knew what to do in response to a situation, our own possible responses are now given to us as a heap of desires and fears and impulses, and it is up to us to put ourselves back together. The principles of rationality are constitutive of the activity of reason, I suggest, because they are principles of unification. (2009, 24)

Rational principles describe the activity reason by being constitutive principles of it, and they are constitutive by virtue of their unificatory function. The two primary forerunners for Korsgaard's position are Kant and Aristotle. But her discussion of constitutive standards and principles also has similarities to Rawls (1958), and it is helpful to keep the earlier discussion of Rawls' distinction between practice and summary rules in mind in what follows.

**2.4** Korsgaard's analysis of constitutive standards owes most to Aristotle, and she looks to him to illustrate what it is for a principle to be constitutive. As she says in *Self-Constitution*: “According to this account normative principles are in general principles of the unification of manifolds, multiplicities, or, in Aristotle's wonderful phrase, *mere heaps*, into objects of particular kinds”. (2009, 50) Korsgaard urges that each kind is



specified by a characteristic function, or *ergon*.<sup>40</sup> Each object is what it is and not something else in virtue of having an *ergon* peculiar to it. She builds upon Aristotle's understanding of the being of sublunary substances, specifically the two ways we in which we can inquire into the nature of these substances: either by asking *what* it is or of *what it is constituted*.<sup>41</sup> The form of a thing is *what* it is, its matter is what constitutes it. This is often called his account of hylomorphic compounds. Considered in general, the form of something is its unifying principle<sup>42</sup>; the form is the actuality of a thing whereas the matter is its potentiality<sup>43</sup>; the form of the thing is what makes that thing intelligible *as what it is*.<sup>44</sup>

A clear case is artefactual objects. Consider a trivial example like an ashtray. An ashtray can be realized in a many different sort of matter – clay, glass, wood, rubber. The material of which it is composed is not essential to it. The form however, is essential. The form is the arrangement of the matter that enables it to perform its *ergon*. An ashtray cannot be convex, for example. If it is convex, it won't catch and contain ashes but will

---

<sup>40</sup> On *ergon*, see *Nicomachean Ethics* (1097b 25 – 1098a20). There, Aristotle asks whether happiness is the chief good of man, and thinks that we must have some insight into man's *ergon* (which W.D. Ross translates as “function”) to see if it is. Since the nutritive and sensitive and perceptive life are common to other living things, must be “the activity of a soul which follows or implies a rational principle”. But Aristotle is emphatic in this passage that simply engaging in the activity of the soul guided by the rational principle is not thereby to do well at it. This is as it should be – my adherence to a standard of rationality might be by rote and I might further be unable to adhere to it in a more difficult, complex, or unfamiliar context. It should be added that Aristotle's use of this term is in fact quite varied (encompassing the process of production, the product of production, an achievement without production), and it would be tendentious to reduce its sense to just this one. See Terence Irwin's glossary to the *Nicomachean Ethics*. Hackett Publishing. Indianapolis: 1990

<sup>41</sup> See *Physics* I 7. and *Metaphysics* 1029a5-7.

<sup>42</sup> *Metaphysics* Z 17

<sup>43</sup> *De anima*. II, 1 412.a

<sup>44</sup> *Metaphysics*. Z 1

rather spill them. Nor can an ashtray have a large hole in the middle, and so on.<sup>45</sup> But it is also not the case that the form of the ashtray is able to exist without being embodied in some matter.

Understanding kinds in terms of their *ergon*, form, and matter enables one to make normative judgments about them, Korsgaard continues. The unstable ashtray is bad *qua* ashtray, and so is the ashtray that has a small hole in the bottom of it.<sup>46</sup> A good ashtray (admittedly not a hard kind to come across) really needs to be only convex, stable, and non-combustible. One can express the *ergon* of the ashtray in terms of a set of standards, Korsgaard urges. These standards are special, though, because they are *constitutive*. A standard is constitutive for some kind just in case some entity cannot fail to satisfy that standard and still count as an entity of the kind to which the standard applies. Put specifically, an ashtray that is concave is not a *bad* ashtray, but rather it is not an ashtray at all – it does not hold ashes. Korsgaard explains that this is true not only of object-kinds but also of activity kinds.

Activity kinds are individuated by what Korsgaard calls constitutive *principles* rather than constitutive standards. A constitutive principle of an activity, just as with constitutive standards, is a principle such that if one fails to uphold it then one does not

---

<sup>45</sup> Thomson (2008) makes a similar point in distinguishing goodness-fixing from non goodness-fixing kinds. A goodness-fixing kind is a kind (K) such “that what being a K *is* itself sets the standards that a K has to meet if it is to be good *qua* K.” (2008, 21) She goes on to say that “There is such a property as being a good K if and only if K is a goodness-fixing kind”, which she draws as a conclusion from “There is such a property as being good *qua* K if and only if K is a goodness-fixing kind” and “Being a good K is being good *qua* K”.

<sup>46</sup> This has the appearance of a vagueness problem – if one increases the diameter of the hole incrementally, at what point does it cease to be an ashtray? I will avoid this matter entirely.

count as engaging in the activity in question. To stay with the example, the person engaged in the process of making an ashtray is guided by the constitutive principles of that process. Of course, these principles must have some connection with the constitutive standards of the kind under manufacture: the process of making an ashtray is governed by the principle that one is making something that holds ashes. This, though, is an example of a process and not an activity. There is reason to think that activities and processes are different, and Korsgaard aims to draw on Aristotle's distinction between *energeia* and *kinesis* to explain the nature of the activity of reason.

On Aristotle's analysis, activity (*energeia*), on its strictest construal, is the opposite of production (*poesis*).<sup>47</sup> Production is a sort of doing that has an end beyond itself and that ceases once this end is achieved. Activity is a sort of doing that has no end beyond itself. Korsgaard clearly thinks that activity of reason is an instance of this general phenomenon. In the *Metaphysics*, Aristotle says that activity is the fulfillment of the *ergon* of a functionally organized kind: "For the work [*ergon*] is a completion, and the actuality is the work, hence even the name, '*en-ergeia*' is said with respect to the *ergon*, and aims at the completedness." (1050a21-2) Korsgaard self-consciously fashions her approach after Aristotle's at least to the extent that the latter took the *completed* or actualized individuating function of a kind to be an activity with the characteristics just mentioned. Likewise, the activity of reason, in her view, is the actualization of the feature that distinguishes the kind human from other kinds.

---

<sup>47</sup> *Poesis* is a species of *kinesis*.

Her account of reason as an activity does not, however, meet all the conditions that Aristotle sets on that phenomenon. In Aristotle's view an activity has the distinguishing features of not exhibiting change, being complete in each of its moments, and lacking of an external end that extinguishes the activity once it is reached. As examples, notice the difference between building a structure such as a house, a perfect example of *poesis* and seeing, an example of *energeia*. The analogy is not perfect, though. The activity of reason does not seem to be complete in each of its moments in the same way as seeing. No episode of healthy seeing is more *complete qua seeing* than any other moment, whereas having endorsed some potential ground as a reason does seem to be both the completion of rational activity. Also, Korsgaard repeatedly speaks in terms of the payoff of the activity of reason as its enabling us to successfully navigate the world. (2009, 28) This way of putting it seems very close to suggesting that the activity indeed does have a goal external to it. The point should not be belabored though. Korsgaard's use of "activity" is probably designed to capture the fact that the activity of reason cannot be carried out just once, but must be repeatedly undertaken throughout the life of the individual capable of reflective awareness.

It is helpful to mention another similarity. Consider Rawls' analysis of practice rules: if after rounding first base the batter cuts a hard left and heads straight for the short-stop she is not making a move in baseball. It is a constitutive principle of baseball that one runs to second *after* first, and third *after* second (there is a reason why the bases are individuated by cardinal numbers – there is a uniquely correct order in which they

must be reached). More to the point, practice rules inaugurate the entities to which they apply. Something similar can be said with respect to Korsgaard's view of principles of rationality. The rational agent is simply the product of the activity of reason, which itself is conditioned by the *principles* of rationality. The rational agent, then, comes to exist through the application of these principles. The agent is both guided and constituted by the rules of the activity.

But given her constructivism, it is not quite right to say that one is *guided*. Rather, Korsgaard urges that we should say that the person *guides herself* by means of these principles in performing the activity, and that by *guiding herself* it follows that *she is guided*.<sup>48</sup> Although constitutive principles set standards of correctness *qua* the kind for which they are principles, it is up to the agent/subject to adhere to them, because it is up to the agent to participate in the activity for which they are constitutive. I have the capacity, that is, to determine to run in a bee-line from first base to short-stop. But to do so would be to flout the activity in which I have chosen to engage. Whether this line of thought is acceptable with respect to the activity of reason, though, is not obvious. The suggestion that one must have the capacity to not engage in the activity in question is suspicious: although this is surely plausible regarding activities such as games or episodes of reasoning (or even Aristotelian activities such as contemplation), it appears to be a much less happy fit given Korsgaard's encompassing conception of reason as activity.

---

<sup>48</sup> This way of putting it is akin to Kant's presentation of the Formula of Autonomy in the *Groundwork*.

The reader will perhaps have guessed by now how this view is supposed to have application to the activity of reason. Rational activity is the completion of the *ergon* of the rational creature. We are what we are in virtue of having this ability, and we are fully what we are in actualizing it. However, it may be asked how this analogy with the artefactual is supposed to hold in the absence of something that is made by the activity. Korsgaard has an answer here: rational activity produces further rational activity. Just as Aristotle distinguishes the form of the living from the form of the non-living by saying that the life form *self-modifies* or *reproduces itself* (the dog reproduces itself not only through mating, but also in engaging in nutritive activities, sleeping, eating, and so on: it makes itself *into its own form*), the rational agent reproduces itself as a rational agent.<sup>49</sup>

It is necessary to mention a crucial disanalogy here, one that will become a problem for Korsgaard later. In Aristotle's view the *ergon* of a thing does individuate it, and can be expressed by a principle (it is what is discovered through assiduous empirical research), and does support normative judgments. But supporting normative judgments is a fairly weak grounding (much weaker than constitution). It is not true for Aristotle that, for any kind whatsoever, the *ergon* of that kind determines whether something is flourishing, or doing well or badly *qua* kind. In the case of humans our *ergon* as rational creatures (creatures that have it in us to be rational) must be developed through proper

---

<sup>49</sup> There is precedent for this way of speaking about the continuation of an organic entity through time. In the third Critique, Kant understands organic entities both as causes and effects of their form, but urges that this is simply the way in which discursive cognizers *must* conceive living things: we cannot understand them purely in terms of mechanistic explanation. We are not in a position to know the nature of organisms as they are independent from the conditions in accordance with which we know (with sensory intuition simply given and structured by the pure forms of intuition and the categories).

habituation if we are to count as thinking *well* or rationally, and as a corollary, our *ergon* must have been properly developed for it to make sense to say of one that she *should* think rationally (since here as elsewhere ought implies can).

The question for Korsgaard, then, becomes how the kind human being relates to the kind rational activity. It should have become clear by now that rational activity for Korsgaard is not the *ergon* for humans in the same way that rationality is the *ergon* of humans for Aristotle. For there is a sense in which a person would no longer fully be human if he no longer engaged in rational activity, and so the failure to satisfy the constitutive principles of that activity (standards of rationality) actually puts pressure on one's properly belonging to the kind in question. There is overwhelming evidence that this was not Aristotle's view of the relationship between rationality and humanity.

An additional clarificatory point needs to be made. There is an understandable tendency to see Korsgaard's picture of rational activity as analogous to Kant's conception of cognition. In both cases the individual is bombarded with material that is not fully conceptualized (with potential grounds playing the role analogous to the manifold of intuition) and that must be structured by the active dimension of the cognizer in order to amount to intentional experience. And like Kant's view of cognition, apperception occupies a necessary and absolutely central role in this activity.

But the analogy is not an apt one. This is because Korsgaard places emphasis on our ability to “stand back” and consider a ground as a *potential ground* and determine whether to endorse the way it typically affects us. There is no free space analogous to this

on Kant's conception of cognition, nor on any plausible interpretation of it anyway.<sup>50</sup>

Although the “I think” must *be able* to accompany any of my representations if it is to count as *my* representation, it is not the case that the synthetic unity of apperception consciously *applies* the categories to the manifold of intuition.

Later I will attempt to give content to Korsgaard's picture of reason and rationality by illustrating its structure in greater detail, developing her view as charitably as possible. In doing so it will become clear that her account is unsatisfactory.

**2.5** Before criticizing Korsgaard's picture, it is helpful to consider the view of rationality and reasoning endorsed by John Broome to bring the salient features of her view into relief. As briefly discussed above, Broome urges that the domain of normativity can be parsed into not only the concept of substantive reasons and oughts but normative requirements as well.<sup>51</sup> Normative requirements are a wide-scope relation between attitudes, which means that they are structural requirements that govern combinations of propositions, specifying which combinations one should have and not which propositional attitudes one ought to have taken in isolation. Broome argues that rationality consists in the set of normative requirements. As an example, consider the case of the person who believes that hard back books have sharp teeth (*p*). If *modus ponens* is

---

<sup>50</sup> Indeed, the most plausible way to understand Kant's conception of cognition is to see the categories of the understanding as figuring into the constitution of intuitions themselves. See sections (xxx) of the B-edition Transcendental Deduction.

<sup>51</sup> Broome (1998), (2010)



a standard of rationality then this person ought to believe that she should wear sturdy metal gloves when opening hard back books ( $q$ ) because she believes that hardback books have sharp teeth and a surly temperament, and one should wear sturdy protective gloves when handling anything that has sharp teeth and a surly temperament ( $p \rightarrow q$ ). If *modus ponens* is to count as a rational standard, then this particular way of formulating it has unattractive consequences: namely, it requires us to form beliefs that we have no reason to form. The problem is that this approach treats the requirement as having an “ought” that can be detached. This means that, taking the intuitive rational standard that one should believe what follows from one's beliefs, the “ought” can be transferred from the implication itself to the content of the belief (the proposition) that is implied. Just because  $q$  follows from  $p$  it doesn't follow that I ought or should have  $q$ .<sup>52</sup>

As an alternative, Broome urges that rational requirements have a wide scope, which means that they govern which combinations of attitudes, telling one what combinations of attitudes one should have but not which attitudes taken in isolation one should have. The ought of normative requirements, that is, is non-detaching. For *modus ponens*, as an example, it is a rational requirement that one not be in the state of having the attitudes ( $p, p \rightarrow q, \sim q$ ) or that one should, if it matters to one, be in the state of ( $p, p$

---

<sup>52</sup> Another way to illustrate the problem with treating rational requirements as giving detaching oughts is to note that every proposition implies itself ( $p \rightarrow p$ ). Consider the normative rational requirement which requires you to believe what is implied by your belief. Since every proposition implies itself, if I believe  $p$  and I am required to believe what follows from  $p$ , then it follows that I ought to believe  $p$ . So if we treat the requirements in question as giving detaching oughts, then it makes belief self-justifying.

$\rightarrow q, q$ ).<sup>53</sup> This says that one ought to see to it that one is in the state of satisfying this requirement, or of not violating this requirement. As another example, consider the commonly held view of rationality that one should believe what is entailed by one of one's existing beliefs. If we treat the requirement between  $p \rightarrow q$  as detaching, then one ought to believe that one is over 6 feet tall if one believes that one is 6' 3", *even if* one is only 5'8". But surely it is not the case that one ought to believe that one is over 6 feet tall just because one has the mistaken belief that one is 6' 3".

Broome stipulates an operator to stand for the normative requirement, "O", that operates on a combination of attitudes:  $O [\sim(p, p \rightarrow q, \sim q)]$ . Conceiving of rationality in this way enables Broome to answer to the intuition that the core of that phenomenon has to do with the interrelations between our attitudes (the intuition that informs the idea that the paradigmatic example of irrationality is in which a person adopts some attitude or performs some action that she judges herself to have no reason to have or perform), but also avoid the unpalatable consequences of the other view this just mentioned. Broome characterizes his conception of rationality as requiring that "particular relations to hold among your beliefs, quite independently of any inferential relations that may hold among your beliefs contents." (2010, 7) Like many others, he counts *modus ponens* and *no contradictory beliefs* among the set of rational requirements. Broome's view of rationality, the reader will likely have noticed, has a significant structural dimension.

---

<sup>53</sup> "If it matters to one" is important here to avoid require one to clutter one's mind with attitudes that one doesn't care about.

Rationality is exhausted by avoiding certain combinations of attitudes – so long as one does not instantiate a non-sanctioned pattern, one counts as rational.

For Broome's conception of rationality as a set of wide-scope normative requirements to be of the most use in shedding light on Korsgaard's view, it helps to consider it alongside his view of reasoning. In his view, reasoning is a specific sort of succession of mental states called attitudes (which are relations between a person and content) in some mind.<sup>54</sup> Although we typically do not hesitate to count the inscription of propositions standing in inferential relations in proof-construction or in legal and philosophical argumentation as instances of reasoning, Broome thinks of these as derivative from, and parasitic on, what I will call his essentially *psychologistic* conception. The specific character of this succession distinguishes it from other mental activities. Namely, in reasoning one has a succession of attitudes, called *premise attitudes*, that cause one to adopt a previously unheld attitude or confirm an attitude in which one was previously unconfirmed, called a *conclusion* attitude. As mentioned

---

<sup>54</sup> A content is a proposition and a mark. Many have urged that the content of attitudes are propositions only, and so the belief that 'I am in Oxford, MS' and the desire 'I want to be in Oxford, MS' have the same content 'RCE is in Oxford, MS'. Broome thinks this is not quite right, since to say 'RCE is in Oxford, MS' does not express a desire. He suggests, then, that we think of the content of attitudes to be a proposition 'RCE is in Oxford, MS' with some mark – in the case of my desire to be in Oxford, this mark can be represented with the *marker* 'nice'. So the content of this desire is 'I am in Oxford, MS – nice'. For the content of belief, Broome takes the mark to be the absence of any other sort of mark. The mark seems to be simply an expression of a pro-attitude or a positive association to the proposition. Broome might object, since he thinks that the “mark” does a good amount of work: it makes desire available to reason with, and it enables one to keep track of what sort of attitude one has in view when reasoning. I see how a mark does the latter job, but how it does the former is not clear. Ostensibly it enables one to express the desire and its content *together* whereas before the two were sundered – since they were sundered one could not reason with desires (one had only the propositional content to deal with and not the desire). But it isn't clear how adding some subjective pro-attitude makes the desire rife for reasoning.

earlier, the distinction between theoretical and practical reasoning for Broome is based upon the nature of the conclusion attitude. If the conclusion attitude is an intention, then it is an instance practical reasoning; if the conclusion attitude is a belief, then it is an instance of theoretical reasoning. He defends a first-order model of reasoning which looks like the following. First, consider this example:

The leaves are showing their undersides.

If the leaves are showing their undersides, then it is about to rain.

So it is about to rain.

On the picture of theoretical reasoning that Broome advocates, one initially believes the first two propositions but not the third. One comes to believe the third by saying the first two propositions to oneself, and then saying the third. In saying the third, one comes to believe it. Reasoning is a computational process that operates on the contents of one's attitudes directly, and one does this by saying one's attitudes to oneself. In so doing one causes oneself to adopt new attitudes.

Crucially, one does not say propositions to oneself *about* one's attitudes, but one merely says the propositions themselves. Were one to reason *about* one's attitudes then one would fall into a regress, in Broome's view. For reasoning about my attitudes entails that I form normative beliefs about them, for example that I *ought* to believe  $q$  because it follows from  $p$  and I believe  $p$ . The problem here is that the only way to acquire this belief *about* my other beliefs is through reasoning. But by the second-order model I can reason only by forming normative beliefs about my attitudes, so I will have to have third

order belief about my reasoning which leads to the second order belief (which in turn is about my first order reasoning). But the third-order belief will itself have to be reached through reasoning and will require a fourth-order, and a fifth-order, and so on. This, the second order model, generates a regress in his view.

There are problems with Broome's first-order view as well: first, it is not clear why reasoning should have to conclude in the acquisition of a previously unheld or unconfirmed attitude – in constructing a proof of the law of non-contradiction, recalculating the probability of the joint occurrence of two events, or running through Kant's argument for the transcendental ideality of space, I do not acquire a new attitude nor become confirmed in an attitude in which I was previously unconfirmed. Yet surely these episodes should count as reasoning. Second, given that the primary nature of reasoning is a causal relation between attitudes that is supposed to track the inferential relations between the contents of these attitudes he must provide some account of how the inferential relations between the contents of attitudes relate to the causal relations between the attitudes themselves. For as stated, there is a substantial disconnect between the two.

For Broome, then, reasoning is discrete and bounded *process* that occurs prior to the adoption of belief or formation of intention.<sup>55</sup> By contrast, Korsgaard's conception of reason is an *activity* that is not hemmed in by the formation of some attitude – although she typically describes the activity of reason in terms of its deliberative character, it

---

<sup>55</sup> This conception of reasoning is criticized by Talbot Brewer (2010), especially in the third chapter.

would be mistaken to think of Korsgaardian rational activity as identifiable with explicit *episodes* of reasoning as that phenomenon is construed by Broome. Rather, although seen most clearly in cases of theoretical and practical deliberation, the activity of reason can perhaps be embodied in non-discursive deliberation as well (unlike Broome's first-order account of reasoning). Examples are easy to furnish, such as deciding whether to take one's umbrella by simply looking to the umbrella, then to the sky, then back to the umbrella, and grabbing it. This is plausibly described as an episode of deliberation, and perhaps even reasoning, but it is unaccompanied by the "saying" of premises to oneself.

**2.6** My primary objection to Broome's view shows what is preferable about Korsgaard's approach at the *most* general level, even if I will later criticize her account on not only its details but trajectory more generally. That is, Broome's conception of rationality has at best a tenuous relation to the activity of reasoning. This is because in order to avoid the problems that beset a conception of rational requirements as issuing detaching oughts he suggests that the standards of rationality require relations to hold among one's beliefs independent of the inferential relations that obtain between the belief's contents and that these structural requirements apply only to states and not to processes. At best, we can "be brought to satisfy" the requirements of rationality during reasoning, but it is obvious that they are merely orthogonal to one another. Reasoning is a temporally *extended* process, and in determining what to believe I successively say my premise attitudes to myself and only *after* this form a conclusion attitude that is caused by my reasoning with

my premise attitudes. So at  $t$  I say a premise attitude to myself and at  $t + 1$  I say a different premise attitude to myself and at  $t + 2$  I form the conclusion attitude. The requirements of rationality come in sideways-on relative to this process, and I satisfy them only by instantiating the right combination (*modus ponens*, *no contradictory beliefs*, etc.) of attitudes at some point or instant in the course of carrying out the reasoning process.

Broome will perhaps respond by saying that although the requirements of rationality apply to different levels they are connected at least insofar as one cannot reason correctly and fail to satisfy them. On the one hand this is suspect because there is good reason to think that at least some rational requirements are narrow-scope and apply to processes rather than states.<sup>56</sup> But in addition, it isn't even clear what it is to reason correctly in Broome's view. One *prima facie* promising suggestion is to say that reasoning correctly just is adhering to patterns of valid inference, deductive and inductive. But this cannot be right – for example, using valid patterns of inference one can show that one is equal to three. (Harman, 1999) Although some rules of logic are intuitive candidates for rational standards, as noted earlier, the theory of implication should be distinguished from at least the standards of rationality, and for similar reasons from reasoning understood generally.

Broome might suggest that reasoning correctly is a process that fits the description given above and satisfies at least some normative requirements during the

---

<sup>56</sup> See Kolodny (2005), (2007).

process (or rather, does not violate any normative requirement of rationality). But it is implausible to argue that the only condition on reasoning correctly is that one satisfy some wide-scope normative requirement during the episode. Since on Broome's account normative requirements do not apply to the inferential relations that obtain between *contents* of attitudes, he is ill-equipped to distinguish between satisfying normative requirements that have to do with a specific episode of reasoning (and so satisfying a normative requirement because of *that* piece of reasoning) and satisfying a normative requirement that has a purely accidental relation to the episode of reasoning in question.

Additionally, Broome appears to inherit a difficulty from realist conceptions of normativity. He comes up empty handed when asked *why* one should satisfy normative requirements. Given the details of his theory he phrases this in terms of whether rationality, understood either as a set of requirements or individual requirements, gives us *substantive reasons*. And as yet he has been unable to provide an affirmative answer to this question.<sup>57</sup>

Korsgaard, by contrast, treats the standards of rationality as constitutive standards of the activity reason, and construes the activity of reason as coeval with the distinctively human awareness. Before presenting criticisms of her view, it should be noted that Korsgaard's approach makes a signal improvement over Broome's. Namely, she urges that the standards of rationality have an internal relation to the activity of reason. This is a virtue of her account because any approach that gives independent analysis to rationality

---

<sup>57</sup> See "Is Rationality Normative?" (*Disputatio*, 2008)



and reasoning risks making any relationship between the two occult. Put differently, treating rationality and reasoning as independently explicable courts dualism properly so-called (as a result, we will be unable to see how the two come together). But surely there is *some* connection between the two – the nearly universal acceptance of certain central rules of inference as candidate rational standards illustrates how strong this intuition is.<sup>58</sup>

Recall the quote from Korsgaard which opened this paper: there is no difficulty in saying that one *ought* to believe *q* if it follows from one's belief that *p*. That we *ought* to believe it is a “reminder” of the constitutive standards of an activity to which we are committed, and is a reminder of our commitment to that activity. One is liable to object by saying that it is implausible to say that we *ought* to have a belief that there is no reason to believe. But on Korsgaard's view there is indeed reason to have this belief, and the reason is generated by the application of a rational standard (what she calls rational principle) that relates the two propositions. In forming the belief *q* one applies a rational principle that is constitutive of reason.

Before progressing, it is necessary to note a potential problem for Korsgaard. The application of a rational principle might be understood to be consciously undertaken, such that the principle is held in mind when we determine what to believe or do. This is not a tenable view. Were it Korsgaard's, then she will have made herself vulnerable to an obvious regress: if rational principles are the objects of consciousness then it will be necessary to know what they mean, to interpret them. But we can do that only on the

---

<sup>58</sup> To argue that the intuition is misled calls for a rather ambitious error theory.

basis of some other principle or rule, which we use to interpret that principle, which itself requires yet another rule, and so on.<sup>59</sup> So although Korsgaard speaks of “applying” rational principles, her use of “application” is not properly understood in a traditional sense of that term. The most promising way to give content to the idea of “applying a rational principle” on Korsgaard's theory is, for example, simply forming the belief  $q$  when  $q$  follows from  $p$ .

A final point. Korsgaard's conception of rationality is much more of a process-forward account than Broome's. To be rational, on her view, is to adhere to principles that govern the formation of beliefs by determining which potential grounds are reasons. It is perhaps a bit ironic that I should say of Korsgaard's account of rationality (as part of an *activity*) that it fits a *process* account best. But, as urged earlier, it is far from obvious that Korsgaard's understanding of activity should be understood to conform to Aristotle's strict account. The point, rather, is that rational standards are diachronic, governing something that we *do*.

**2.7** If one were to guess Korsgaard's answer to the guiding question (“why should I be rational?”) in light of the discussion of her view up to this point only, the following would be a fine suggestion: you should be rational because the principles of rationality are constitutive features of an activity that is essential to our kind. Such an approach

---

<sup>59</sup> This is Wittgenstein's point in §198 of the *Philosophical Investigations*. See also John McDowell (2009) and Crispin Wright (draft )

might fall within what is sometimes called Aristotelian *natural normativity*, an approach revitalized by Anscombe (1957) and developed by Foot (2001) and Thompson (2008). This sort of view would treat the normative as entering in as the relation between the general and the particular, specifically, between the flourishing or proper function of an instance of some kind *qua that* kind. Consider the judgment “The female groundhog has two young in the spring”. Thompson calls this a natural-historical judgment, and notes that it has some rather surprising characteristics. First, it is not, despite appearances, a universally quantified judgment. Something does not stop being a female groundhog should it not give birth to two young in the Spring. But neither is this a probable statistical generalization – it is not likely that a female groundhog, as a matter of fact, will give birth in a Spring (much less in every Spring).<sup>60</sup> The truth of the natural historical judgment, or what he also calls an Aristotelian Categorical, is not a function of whether all or even most groundhogs actually do reproduce in the Spring. Judgments of this sort, Thompson urges, are not distributive and exhibit what he calls “non-Fregean generality”. This sort of judgment allows for exceptions, but not of a *ceteris paribus* sort that restrict its generality. Hence, we can say of the groundhog that does not reproduce that there is something *wrong* with it. We can make normative judgments on the basis of the special relation between the general and the particular. This is similar to the point made earlier when Korsgaard urged that understanding objects in terms of their matter and form

---

<sup>60</sup> To make the point more vivid, consider the butterfly: the likelihood that a butterfly will even survive long enough to breed is very low. Yet, we say “the butterfly reproduces in the spring”.

supports normative judgments about them. A different way of characterizing this approach can be seen in Hurka (1993) whose ethical perfectionism urges that rationality, understood as the capacity to acquire knowledge with the properties of *extent* (scope) and *dominance* (number of beliefs subordinate to it), is the essential characteristic of human kind to which we belong. Since we are better *qua* human to the extent that we more perfectly instantiate this capacity, we should satisfy rational standard (such as *modus ponens*) that are a part of this capacity. An application of this line of argument to Korsgaard's approach would say that "The human being adheres to *modus ponens* in its thinking". To the extent that instances of the human form do not adhere to this we can say that there is something *wrong* with this person and that he *should* adhere to *modus ponens*.<sup>61</sup>

But Korsgaard's view is not so clear. She also suggests that the constitutive principles of an activity are normative for the participant of that activity if and only if that participant has *committed* to the activity. She frames the categorical in terms of reflective distance. But in this case, as suggested earlier, the normativity will issue from the endorsement or commitment rather than the relation between the general and the particular expressible in the claim that the "the human engages in the activity of reason". On this approach it is not belonging to the kind that makes them normative, but rather the commitment. Which view best expresses Korsgaard's position regarding the normativity

---

<sup>61</sup> Earlier I questioned the Aristotelian credentials of this view, specifically with respect to human rationality.

of rational standards will be considered below.

With Korsgaard's account fully in view, it is now necessary to draw attention to its inadequacies. The objections below criticize Korsgaard along two axes: first there are objections to Korsgaard's conception of reason and the rational principles that are supposedly constitutive of it (and so her conception of rational standards); second, I will object to Korsgaard's view of the *normativity* of rational principles.

Although I organize the criticisms along these two axes, given how tightly Korsgaard weds her explanation of the concepts of reason and rationality they will occasionally sound similar. Additionally, although Korsgaard's general picture is in view crucial details are not. I have delayed close treatment of Korsgaard's account of the application of rational principles to potential grounds because only in attempting to give substance to Korsgaard's schematic account will we see the objections to which it is vulnerable, and recognize that her approach is not even intelligible, much less satisfactory.

### 3 OBJECTIONS TO KORSGAARD'S VIEW

**3.1** The first objection to Korsgaard's view can be called the *objection from incoherence*. The description of Korsgaard's constructivism above will perhaps have made the reader aware of it: if substantive reasons are *created* in being picked out by rational principles, then there must be something to be picked out. It seems incoherent to suggest that something can be picked out if it is *created* only in being picked out? This is akin to objection sometimes presented against the idea of extensional definition of a concept or class: on what basis can one group like instances so that the that class or concept can be defined as the set of those entities? Without a rule (specifically the concept to be defined extensionally) in view one's selection will be random or rely on some inscrutable power of intuition.

The germ of this objection also appears once the idea of reflective distance receives scrutiny. That concept turns on the notion that I can ask of some representation whether I *should* endorse its operation on me. This looks incoherent insofar as asking whether I *should* endorse the operation of some representation on me is intelligible in terms of substantive reasons. As discussed earlier, the core of this sort of objection is the claim that Korsgaard helps herself to concepts in describing the procedure of construction

to which she is not entitled because they are the products of that procedure. Substantive reasons supposedly come into existence only once I endorse its operation on me – but what could it even mean to ask whether I “should” endorse some representations operation on me without the idea of substantive reason in view? Although I call it the objection from incoherence for this reason, I think in the process of developing Korsgaard's position in response to it illustrates her vulnerability to a somewhat different, but just as serious, problem.

One way to respond on Korsgaard's behalf is to say that normative terms such as “should” and “ought” are not to be analyzed in terms of substantive reasons. This, however, will not work given that Korsgaard herself urges that the domain of normativity comes into existence as a result of the volitional activity of agents, to use R. Jay Wallace's phrase (2006). So, strictly speaking, there is no space for even normative terms like “should” or “ought” prior to the creation of reasons. If this is Korsgaard's position, then it surely is incoherent. Perhaps, though, her view can be developed in a manner that renders this incoherence merely apparent.

Let us return to Korsgaard's description of rational activity. Principles of rationality, she says, are applied in determining what to believe, (2008, 18) and through the application of rational principles representations that were formerly only causes of our behavior are instead encountered as reasons for it. Or better, what was once a cause becomes a rational candidate or potential ground through our special form of awareness – it becomes the sort of thing that can be seen *as a reason* to have an attitude or perform an

action, a reason not to have some attitude or perform an action, or something that is in fact not a reason at all. Only once a cause (or some set of causes) becomes a rational candidate or potential ground as result of the distinctively human awareness can it become a reason in being picked out by a rational principle. But how does the application of a rational standard such as *modus ponens* pick out a reason from a set of potential grounds? The natural suggestion is to say that the conclusion drawn from a substitution instance of *modus ponens* just is the reason that is picked out.

This suggestion, though, requires that premises are available for such an inference to take place. That is: one encounters a set of potential grounds for one's beliefs and one has to determine whether these potential grounds are reasons, that they stand in rational relations to the beliefs that one is disposed to form. If *modus ponens* is a principle that one uses to determine whether some potential ground is a reason, then how does one come to have premises from which one can draw the conclusion (and so the reason)? Only from true premises together with a valid (or strong) pattern of inference can one be guaranteed (or make it likely) that one draws a true conclusion. Adhering to *modus ponens* in deliberating does not transform a potential ground into a reason because it is simply a pattern that preserves truth. We cannot give ourselves reasons to believe something just because it is the conclusion derived in a substitution instance of *modus ponens* – for the premises could be false. Something does not become a reason just because it has been “picked out” by *modus ponens* on the most plausible construal of “picking out”.



To say that the premises are themselves a product of such an application of *modus ponens* only kicks the can down the road, and the buck (as it were) has to stop somewhere to avoid a regress. So this way of picking out a reason from a set of potential grounds, or alternatively endorsing a potential ground as a reason, is *prima facie* inadequate – the application of rational principles such that it enables us to pick out a reason from a set of potential grounds is not correctly construed in this way.

This argument can be pressed at a different level as well. Consider the question: “on what basis does one apply a rational principle?”. On a metaphysical interpretation, the very idea of one applying a rational principle implies that something exists (in some sense or other) that applies it, namely a rational agent. But Korsgaard maintains that the rational agent who applies principles is constituted through the application of principles – so it seems that the entity that applies the principle must exist in order to apply it.

On a logical interpretation of this question, we ask for what reason one applies a rational principle. This seems to come up empty handed also: for reasons are the product of the application of rational principle and are therefore not available independent of them. This is reminiscent of the problem Socrates presents to Euthyphro when the latter suggests that for something to be pious *just is* for it to be loved by the gods.<sup>62</sup> If reasons are the result of the application of rational principles, then the application of rational principles has to be arbitrary. This is sometimes called the *boot-strapping* objection to

---

<sup>62</sup> Shafer-Landau, in *Moral Realism: A Defense* (2005) discusses this objection to constructivism in morality and ethics more generally.

constructivism – it seems very counterintuitive to say that reasons can simply be ushered into existence by the endorsement of an individual. Although Korsgaard aims to give content to endorsement by construing it as the application of a rational principle, it is questionable whether this answers the objection.

Korsgaard might try to counter the arbitrariness objection in its logical guise by urging that there is only one set of rational principles to be applied since the source and nature of the activity of reason is the same for every creature that has reason. The reflective awareness that she treats as the source of the activity of reason (and therefore of the application of rational principles) has one and the same structure across the set of all creatures that have it. In this way, she might urge, the application of rational principles is clearly not arbitrary because for something to be arbitrary implies that there are options other than that thing and any option is as good as any other option. But this is precisely not the case with the application of rational principles – they are our only option. So the application of rational principles is not arbitrary.

If this blocks the objection of arbitrariness, it does so only by invoking a consideration of the wrong kind: this response to arbitrariness is at bottom an invocation of brute fact. It simply says that “we are creatures that *y* and all such creatures *y*” - but surely this is the wrong kind of reason with which to counter the arbitrariness charge. A fitting answer to the charge will be to show that on *rational* grounds the application of rational principles is not arbitrary. But to say simply that, as a matter of fact, we all have the same cognitive equipment or we are all necessitated to form beliefs that enable us to

navigate the world successfully does not help here. For we might have been constituted differently, such that our awareness engendered different rational principles. If this is Korsgaard's response to the objection of arbitrariness then it makes the standards of reason contingent on brute fact.<sup>63</sup>

One might come to Korsgaard's aid here by urging that I have simply begged the question against her. To complain that her view is vulnerable to a charge of arbitrariness because she cannot provide a basis, specifically a *reason*, for the application of rational principles is simply to assume a non-constructivist view of reasons and then blame Korsgaard for not satisfying the conditions on such a view. But the nature of reasons, and *whether* they obtain independent of our rational activity, the objection will continue, is precisely what is in question.

There is perhaps *something* to this response, but it has been addressed above. Namely, Korsgaard's interlocutor and Korsgaard alike have the intuition that there has to be *some* basis for the application of rational principles, and this intuition typically finds expression in normative concepts like "substantive reason". But her interlocutor need *not* say that the bare intuition that *substantive reasons* are the way stay the specter of arbitrariness is sufficient to scuttle Korsgaard's theory, or that they are the *only* way to show the inadequacy of her approach. Rather, he can simply show that Korsgaard is

---

<sup>63</sup> This line of criticism is similar to that urged by McDowell against the plausibility of Kant's conception of cognition, which fails because Kant takes the forms of intuition to be brute features of our cognitive architecture, and so contingent (he repeatedly gestures at the plausibility of cognizers with different forms of sensory intuition or no forms of intuition at all (an example of the latter is a divine intelligence).

unable to provide any satisfactory *basis*, substantive reason or otherwise, to answer the charge of arbitrariness. I suggested just this above, urging that she relies on a basis of the wrong kind (that is, brute fact) to avoid the arbitrariness of the application of rational principles. I did not, of course, endeavor to show that it is *impossible* for Korsgaard to furnish a satisfactory basis for the application of rational principles in a constructivist vein: but her options have been narrowed.

Not only is the brute-fact response to the arbitrariness charge unpalatable, it is especially so given Korsgaard's position. Korsgaard aims to vindicate the rational authority and autonomy of humanity, and show that it is of singular importance: specifically it is the source, and the only source, of normativity. But if we might have been differently equipped than we in fact are, then this source is merely a contingent fact of our constitution.

Korsgaard attempts to settle the metaphysical aspect of the objection from incoherence by an argument from analogy. Just as an organic entity maintains itself by ingesting nutritive material in the way that members of its species typically do (the anteater uses its long proboscis to reach its fare, acquiring the materials its body needs to persist through time and so making itself into itself, ultimately reproducing its form in offspring) the rational agent persists by performing its characteristic activity: put crudely, achieving awareness with the property of reflective distance, applying rational principles, determining what to believe, acting on the basis of these beliefs, and presumably doing it

all over again.<sup>64</sup>

But the strength of this analogy is questionable. The core of the analogy is that the way that constitutive principles describe the persistence of organic systems is the same way as the activity of reason is constituted by rational principles. It is somewhat clear how this hylomorphic view sheds light on the perpetuation of the organism *qua* organism, both the individual through nutritive intake and also the species through reproduction, but is rather less helpful for understanding the nature of reason's activity.<sup>65</sup> The most obvious wedge to drive between the two is to stress, perhaps flat-footedly, that organic hylmorphism is just that: organic. It is embodied. It is not clear that one's awareness and seeing something as a reason at moment is related to one's seeing something as a reason at a later time in the same way that one's organic existence at a later time is related to one's organic existence at an earlier time. At the risk of sounding obtuse, this seems radically different from, say standards of rationality. They surely are not dependent on matter in the same way as the principles that describe the perpetuation of some organic system.

Again, one risks begging the question against Korsgaard in responding this way, urging that a consequence of her view is that we cannot very well help ourselves to realism about rational standards, and since realism about rational standards is correct, we

---

<sup>64</sup> *Self-Constitution* (2009, 35)

<sup>65</sup> Although for this analogy to be its strongest, Korsgaard will need to show that this conception of the organic system is itself plausible, which requires an investigation of how it relates to the most developed biological understanding of organic systems.

should not adopt her view. But to accept this position requires Korsgaard to bite a bullet, the same one mentioned above. That is, why are *these* the rational principles – could we not have been constituted such that the application of some other principle (say some invalid rule of inference) enables us to see considerations as reasons? To respond negatively to this question, at the very least, gives rational principles all the independence they need for realism.<sup>66</sup>

But there is a more obvious and powerful way to press the disanalogy. In Aristotle's view, principles are expressive and not constitutive of the *ergon* of a species. The chestnut tree is what it is in that there is something in it that enables it to develop from an chestnut into a chestnut tree and then produce more chestnuts. But obviously a chestnut tree does not stop belonging to the kind *chestnut tree* should it fail to develop and produce chestnuts. It is just an unflourishing chestnut tree. The principle is the key to explaining the kind in question, but it does not completely describe it, nor is it sufficient for its integrity (an instance of the kind might be blighted). By contrast, Korsgaard attempts to make the principles *constitutive* in a problematically strong sense. She seems to think that the application of rational principles after having occupied reflective distance is both necessary and sufficient for engaging in the activity of reason and therefore for being rational. This, though, is surely is an implausible view. And as I will argue below it leaves no room for the idea of being more or less rational.

---

<sup>66</sup> That is, realism about standards of rationality or rules of logic does not require that one provide a third realm for them to inhabit, but simply to say that that are valid independent of our taking them to be so.

It would be more useful for Korsgaard to draw the analogy between the activity of reason and some other non-organic activity. It is tempting to think of the issue in the following way. In a baseball game the total set of attitudes and movements that can be isolated in a given instant of a baseball game together with the observation of the constitutive principles of that game *enable* or *produce* the next time-slice of the baseball game. This analogy is more promising for Korsgaard's purposes: the obvious discrepancies that appear when drawing an analogy between the constitutive standards of a species-form and the constitutive principles of rational activity do not appear here.

But the crucial disanalogy between the baseball game and the activity of reason is that one can decide to not play baseball. One cannot, however, decide not to engage in the activity of reason, on Korsgaard's view. To do so, one has to engage in the activity of reason. This has unfortunate consequences for vindicating the normativity of rational principles.

There is also the outstanding question of what it is to apply a rational principle on Korsgaard's view. This was discussed earlier when we attempted to clarify and express how a rational principle determines which potential grounds are reasons. But we can also directly inquire into what it is to apply a rational principle. Elsewhere, Korsgaard says of her chosen principle of practical reason, the categorical imperative, that it “is a principle of the *logic* of practical deliberation, a principle that is constitutive of deliberation, not a theoretical premise applied in practical thought.” (2008, 302) We saw earlier that the view which suggests that to apply a principle is to take that principle or standard as the

content of a premise, to reason *about it*, leads to a regress of the sort seen in Carroll's "What the Tortoise Said to Achilles", or of the sort that Broome attributes to the second-order model of reasoning. A promising way to give content to the idea of "applying a rational principle" that avoids such regress is, for example, simply forming the belief  $q$  when  $q$  follows from  $p$ . In this way, when I encounter potential ground I infer what follows from it and in drawing this inference (and forming the belief that follows) I endorse the operation of the potential ground  $q$  on me. But this doesn't show that the application of rational principle is a reason to form a belief that  $q$  on the *basis* of a potential ground for  $q$ , but merely shows that I am treating  $q$  as of credible rational authority by drawing an inference from it. But the supposed rational authority is *just* that I have drawn an inference from it, not that  $q$  itself has reasons in support of it.  $Q$  should have independent rational authority if I am to form beliefs on its basis by the application of rational principles.

The application of the rational principle that any belief implies itself does not help here, for that I can infer  $q$  from  $q$  does not thereby lend rational credibility to  $q$  (as seen earlier with the problem of bootstrapping). Were this Korsgaard's view, then any belief that is inferred from any other belief would have rational credibility (and perhaps even count as justified).

Additionally, this conception of the application of rational principles does not sit well with Korsgaard's general approach. For the most clear way of explaining my forming  $q$  because I have  $p$  and  $p \rightarrow q$  is to say that I simply form  $q$  through some



unconscious process. But surely this suggestion is of questionable constructivist credibility – even to respond that one typically forms such attitudes through an unconscious process that could then be given rational reconstruction by the subject still seems to run afoul of the guiding aspiration of Korsgaard's view – it would require the reflective distance be given a counterfactual presentation (I could ask “why should I see this reason?”), which would eviscerate the point of the idea, which is to locate the source of normativity in the *awareness* characteristic of the deliberative position.

Korsgaard needs a conception of the application of rational principles that does not require these principles to be contents of reasoning (and so avoids regress) but is capable of lending rational credibility to a potential ground of a belief and so to *that* belief and not just which beliefs follow from it. Providing an account of what it is to *apply* a rational principle on Korsgaard's view, then, encounters much the same difficulty as providing an account of how the application of rational principles determines *which* potential grounds are reasons. In both cases the application of rational principles must take an arbitrary starting point, namely a belief that is a potential ground, but which as yet has no rational credibility nor basis.

**3.2** I will frame a second complaint about Korsgaard's conception of reason and rational principles in terms of the *objection from diversity*. It is perhaps less serious than the *objection from incoherence*, but still presents a problem for Korsgaard's view. Korsgaard says precious little about *which* are the principles of rationality. As examples she gives

“rules of logical inference, the principles which Kant identified as principles of the understanding, canons for the assessment of evidence, mathematical principles, and the principles of practical reason.” (2009, 3) In the foregoing discussion I have focused on a formation rule of natural deduction as an example (Korsgaard would agree that this standard falls under “rules of logical inference”, and I have simply assumed that whatever the final list of rational standards is, this would be among them). This objection urges that Korsgaard will be unable to specify the set of rational principles such that they fulfill the constitutive role she assigns them without deeply counterintuitive consequences.

There are surely a great many candidates for rational standards that fall under Korsgaard's brief list. It must be addressed at the outset that we have been speaking as if one view of logic enjoyed unequivocal support. Although the first-order predicate calculus and the rules of natural deduction are certainly ascendant, a variety of non-classical logic also enjoy support as well

Even within the domain of deductive logic there are a plurality of rules. Of course, the number of principles of deductive inference can be twiddled down given the interdefinability of some rules and connectives. But which inductive principles are to be included on the list? Additionally, there are the standards of Bayesian epistemology, principles of decision theory (like maximin and minimax), the rules of the probability calculus, and so on. The objection is not, primarily anyway, that it would be unparsimonious for Korsgaard to admit all of the intuitive candidate rational standards

that belong to her list to the status of rational principles.<sup>67</sup> Rather, the worry is that these standards count as rational *principles* only being *constitutive* principles of reason's activity. But for something to be a constitutive principle of some activity, by Korsgaard's own definition, one cannot fail to conform to the principle and still count as performing the activity in question. But given that we are of course only imperfectly rational, and only *sometimes* satisfy *some* of the principles on offer. Imagine, for example, one who is competent with deductive logic but is a systematically bad inductive reasoner – he almost never adheres to any of the principles of inductive inference. If Korsgaard admits inductive standards to the canon of rational principles, then this person does not count as engaging in rational activity at all. But this is surely a wildly implausible consequence.

She can respond to this objection by urging that the other rational standards mentioned are not constitutive principles. This option is unpalatable since it would require abandoning the normativity of these principles, which would have strikingly revisionary consequences. We often judge, for example, that the individual who reasons in accordance with the standards of inductive inference, decision theory, etc., is doing better than the person who is competent in only one of these domains.<sup>68</sup> If we accept Korsgaard's constitutive view on *this* construal we will not be entitled to such judgments.

There is also the question of whether these less obvious candidates for rational principles fit the description for that kind. To make the case that they do, Korsgaard will

---

<sup>67</sup> An objection of this sort seem to amount to little more than accusations of bad philosophical taste.

<sup>68</sup> We need not go in for a model of an ideally rational agent with unlimited computing power whose only end is maximizing payoffs to itself to be entitled to such evaluations.

need to show that they play a robust unificatory role (and so are constitutive) or that they are related in the right sort of way to the rational principles that do. If she argues that they do play a unificatory role then she will have to determine which principles have priority (if any) and why, and will also face the problem mentioned earlier: if the set of rational principles includes all of our intuitive rational standards then it will be much too easy for one to fail to be engaging in rational activity at all (since on any view of rationality we only imperfectly rational). If she argues that these other principles are not rational principles but are related to them in the right sort way then she will have to specify the nature of this relation. Her best option would be to say that the non-central rational standards are deductively derivable from her rational principles. But we have no reason to think that this is so.

Additionally, although there is a considerable *philosophical* agreement about what should be included on the set of rational standards, empirical evidence suggests that people are systematically incompetent in both deductive and especially inductive reasoning.<sup>69</sup> If these are the standards of rationality, and the activity of reason is the basis of distinctive human experience, then Korsgaard shoulders a great explanatory burden to reconcile their indispensability with empirical facts.

Another potential objection, but one that I will not press, is that Korsgaard must square her view with salient experiential aspects of the deliberative position. Although she draws on that position as the ground for normativity, while occupying that position

---

<sup>69</sup> See Stich (1990)

rational candidates (for belief or action) do not appear to be on equal footing, which they should if her account is correct. That is, if the potential grounds of belief or action become normatively charged only after being endorsed, then it should be the case that all considerations show up to the person deliberating as of equal weight (that is, no weight). But this is surely not an accurate picture of the deliberative position. Rather, we feel pulled towards some considerations rather than others independent of endorsing them. I see that I should stop the pain being inflicted on my friend if I am able, and that I should do this does not seem to result from my ability to ask myself *should I stop the pain being inflicted on my friend since it is within my power to do so and I find it objectionable?*<sup>70</sup>

This itself is not a definitively damning objection, but is simply an additional explanatory weight for Korsgaard to shoulder. It would seem to require a reconstruction of the already reconstructed procedure of construction - one begins to wonder whether and how the law of diminishing returns applies to theoretical reconstructions.

**3.3** Korsgaard's is a constructivism of an odd sort. That is, there seems to be a profound tension at work in her account of the normativity of rational principles that can be expressed in the following disjunctive question: are they normative *because* they are constitutive principles of rational activity, or *because* we have committed (or alternatively “are necessitated”) to undertaking rational activity? This ambiguity is evidence of a

---

<sup>70</sup> Additionally, Korsgaard would need to say more about the times that fall between deliberation. For even on her view deliberation is not constantly taking place. Given that she argues that the activity of reason is what we are at the most fundamental level, then what are we when not engaged in deliberation?

general tension in her efforts to fuse Aristotelian and so-called Kantian constructivist accounts of normativity.

In *Self-Constitution* she claims that her view is informed by a conception of normativity drawn from Aristotle. Normative principles are such because they are the unification of manifolds indexed to kinds. The principle that *unifies* a kind is normative for that kind, meaning that it determines whether some instance is good or bad for the kind in question. This, in turn, could provide an explanation of normative judgments of the sort we have in view: *y should be A*, or “why is it wrong to violate *modus ponens*?”. On some sort of an Aristotelian view<sup>71</sup>, it can be answered that “because you are bad *qua* human being in failing to do so, since it is essential to our kind that we have the capacity to be rational”. This might take the form of a perfectionism which claims that insofar as the faculty of reason is the essential property of the kind human and that the adherence to and development of rational principles is the way to perfect and develop this property, it follows that each person is under a normative requirement to adhere to rational principles.

Crucially, this is a metaphysically realist picture, in which normative facts are fixed by the world independent of the volitional activity knower or actor. But Korsgaard is not a realist, so a vindication of the normativity of rationality along this general line of thought would be difficult to square with her commitments elsewhere. In her view an activity has normative force for her just in case the agent in question has *committed* to the activity.

---

<sup>71</sup> As an example, see Thomas Hurka's *Perfectionism*

But she cannot have it both ways. You should not violate *modus ponens* either because you have committed to it, or you should not violate it because it is a feature of the kind to which you belong that you do not violate it. The reason why one should not violate *modus ponens* cannot be both “because *modus ponens* is a necessary condition for being an integrated rational agent with a determinate conception of the world” and the sort of natural normativity approach just mentioned and discussed earlier.

One might try and gerrymander here, urging that it is a feature of our kind that something is normative for us only if we have “legislated” it or committed to it. But this doesn't “distribute”, as it were, the normativity to both factors, but simply makes the normativity of rational principles a product of our commitment to the activity of reason. Korsgaard, then, has two options. She can argue either that rational principles are normative because they are *constitutive* principles of rational activity and so *constitutive* features of the kind to which we belong, or she can argue that rational principles are normative because we have *committed* to rational activity and they are constitutive features of performing that activity.

The difference between the two perhaps seems subtle. To distinguish them it must be borne in mind that on the first proposal it is the *constitutiveness* of the principles for rational nature (which just is the activity of reason) which is essential to human kind that does the normative heavy lifting, on the latter it is the idea that we have *committed* to performing the activity of reason that does normative work.

**3.4** If Korsgaard's answer to our guiding question follows the first approach, then she is vulnerable to the *objection from the incompatibility of constitutivity and normativity*. If we accept her account then one cannot violate rational principles and still count as a rational subject – for acceptance and application of the principles is a necessary condition for rational activity. This makes it difficult to see how these can be normative: there is good reason to think that for some standard to be normative for a domain one has to be able to fail to uphold it and still count as having the standard apply.<sup>72</sup>

We have seen earlier how the violation of constitutive standards, the rules of chess for example, do not count as chess moves at all. So one is not making a bad chess move or a wrong chess move in advancing the queen in an L-shaped path, but rather not a chess move at all. Korsgaard aims to counter this objection by relying on the notion of *defect*: the rational subject who flouts rational principles is defective *qua* the kind to which she belongs. She is a *defective* rational agent. The best way to flesh this out is to say that, staying with the example, although the move itself is not a chess move this doesn't mean that the player is not a chess player. Rather, it is in virtue of the illegal chess moves that he is a bad chess player. But what distinguishes the illegal L-shaped progression of the queen from the something else that is an illegal chess move, such as flicking the pawn like one would a football made from paper at the opponents king or swatting at the chess pieces with a branch? These too are not chess moves, and we do not want to say that the player is just a bad chess player in making them. If a bad chess player were one who

---

<sup>72</sup> See Clinton Tolley (2006)



failed to adhere to chess rules, then I would count as a bad chess player by flicking the king like a paper football or beating my shoe on the chessboard as a tactic of distraction – there would be no principled way to draw the distinction between playing “badly” in this way and playing badly by always advancing one's rook in an L-shaped path. In both cases, this is just not playing chess at all, not being a bad chess player. I am a bad chess player *properly* so-called if I adhere to the rules of chess but very seldom (if ever) win.

Korsgaard might respond by urging that her understanding of defect can withstand this line of criticism. To do this, she might say that failing to uphold a constitutive principle is compatible with still falling within the domain or kind for which the principle is constitutive. The most plausible way to do this is to treat the principle as admitting of degrees of satisfaction, such that one might satisfy a principle only partially.

Superficially, she could assign values to a range, such that I could count as adhering, for example, .60 to some rational principle. If I fall below some specified degree of satisfaction, then I will not count as a belonging to the kind or domain for which the principle is constitutive. This, though, is not a satisfactory response. For the only way (it seems to me) to give content to the notion of only partially satisfying a rational principle is by adhering to it in only some and not all cases. But in each specific case I will either adhere to the rational principle or not – so this response does not genuinely adopt the idea that *each principle* permits of degrees of satisfaction. In the cases in which I satisfy a principle I will count as engaging in rational activity and in those where I do not I will not count as engaging in rational activity (and in such cases it is not clear in what sense I

“should” adhere to rational principles).

Also, she would need to designate some value greater than .0 as the point of failure to satisfy the constitutive principle (otherwise there is no point in introducing a range system). Korsgaard might simply accept this and say that someone who satisfies rational principles in at least, say .5 cases, counts as engaging in rational activity. But in addition to the problem just mentioned, this value will be merely arbitrarily selected, and not only because of a line of demarcation problem. It will also be arbitrary because there is no normative basis to prefer one value to another given that normativity issues from the satisfaction of constitutive standards – to say that there is some reason or basis to prefer one value to another is to beg the very question at hand. The same holds true were she to argue that satisfaction of greater than .5 of the total constitutive rational principles at least some of the time (greater than .0 for each) is sufficient for engaging in rational activity

An account of the normativity of rationality could be refashioned into a more genuinely Aristotelian approach by abandoning the idea that rationality can be defined by a set of rules or principles. The thesis of the non-codifiability of practical reason characteristic to Aristotelian approaches could be extended to reason in general on this approach, urging that the best way to develop and perfect one's capacity for reason is in the development of a rational sensibility attuned to various desiderata of from the rational perspective (such as sensitivity to evidence, consistency, etc.) but not one codifiable in a set of first-order principles.

This view would not be attractive to Korsgaard either, especially given her

fundamentally Kantian orientation. In her view both theoretical and practical reason must be construed as governed by the application and adoption of principles, an aspect of her position that has its basis in what she thinks is the only available answer to the normative question of “why anyone *should*” adhere to rational principles: because we have authored them.

Another objection to the first approach is the *objection from rule-following skepticism*. If Korsgaard explains the normativity of rational principle solely in terms of their constitutive role then will be unable to distinguish “genuine” rational principles from spurious ones of the general sort presented by Saul Kripke, such as *quus*.<sup>73</sup> Or more precisely, the problem for Korsgaard is not that this interpretation of her view is unable to answer skeptical challenges to the possibility of following a rule (for that is problem that effects constructivist and realist views equally), but rather that it would too easily allow what we have good reason to regard as spurious rational principles as normatively authoritative for agents.

Suppose that you encounter someone who has consistently applied *modus ponens* for as long as you have known him (you have seen him reason in accordance with it, that is), but one afternoon inexplicably infers like this ( $p, p \rightarrow q : \sim p$ ), claiming that he has in fact been adhering to *schmodus ponens* ( $p, p \rightarrow q : q$  for the first 1500 applications and afterwards  $p, p \rightarrow q : \sim q$ ) all along. What distinguishes the genuine rational standard from the spurious one if they both fulfill the requisite unificatory role?

---

<sup>73</sup> See Wittgenstein *on Rules and Private Language* (1982)

Korsgaard will reply that they are to be distinguished by the fact that subsequent to the transition to non-truth preserving inference the subject will begin to become “disintegrated” - they do not fulfill the unificatory role. But it is not clear what this amounts to for rational activity. In the case of artifacts the failure to adhere to the constitutive standard in the production of that artifact will yield something that is bad *qua* kind (a house with drafts or leaks) or worse, not even a member of the kind (a house without a roof). To respond like this to the rule-following objection is tantamount to simply saying that if one flouts rational principles then one is bad *qua* rational subject. But Korsgaard is not entitled to this objection in the absence of standards of goodness and badness *qua* activity of reason, and constitutive principles cannot do this job.

But let us grant that there is some sense of disintegration for the activity of reason *qua rational* that will befall the person just mentioned. Even granting this, Korsgaard's response is still unsatisfactory. Instead, we can frame the objection in terms of, say, *spring Thursday schmodus ponens*. Suppose that *spring Thursday schmodus ponens* is a rule which states that every Thursday immediately following the vernal equinox one infers in accordance with  $(p, p \rightarrow q : \sim q)$  and otherwise  $(p, p \rightarrow q : q)$ . Surely this will do a good enough job “constituting” given that the first Thursday after the vernal equinox falls only once a year. Even supposing that a concept of rational disintegration is in view, this formulation of a spurious rational principle will surely not lead to it.

The idea of rational disintegration merits closer consideration. The organic system which fails to adhere to the “constitutive” principles of its kind (the honey bees that fail

to produce honey; the squirrel that doesn't collect nuts) will grow weak and, ultimately, die. It was suggested above that to respond in this general way to the rule-following objection without a conception of “disintegration” for rational agency in view is merely a metaphorical flourish for saying that one is bad *qua* rational agent. And this doesn't answer the question at hand: in virtue of what is it true to say that I should not violate *modus ponens*? It is perfectly intelligible to have standards of goodness and badness for some kind yet it not be the case that one would be wrong not to satisfy the constitutive standards or principles that are conditions on belonging to that kind.

The idea of rational disintegration might seem to rest on the strength of the analogy between organic systems as defined by constitutive principles and the activity of reason as defined by constitutive principles. But perhaps we can look elsewhere. It would be most charitable to respond on Korsgaard's behalf by looking to her inspiration for this idea in the *Republic*, to see if the account of psychic disintegration discussed there can shed light on the idea of a disintegration *qua rationality*. Although I complained above that the idea of rational disintegration is merely a metaphorical flourish for suggesting that one *should* not violate the standards of rationality, the problem is not, as Murdoch notes, not with metaphor *as such*.<sup>74</sup> Our intellectual lives are pervaded with metaphor, and to try and purge metaphor is a wrongheaded task. The problem is rather the aptness of the metaphor for the task at hand: specifically, the idea of disintegration does not

---

<sup>74</sup> In “The Sovereignty of Good over Other Concepts”, collected in *Existentialists and Mystics*. Penguin Publishing, New York: 1997

provide a satisfactory account of why we *should* be rational when that phenomenon is understood to be satisfied only in the course of as an activity that can be performed only by adhering to principles.

The transition from the aristocratic to the timocratic constitution, which would be analogous to the initial deviation from adherence to rational principles, might be of help in developing Korsgaard's point.<sup>75</sup> The disintegration begins when the guardians of Kallipolis miscalculate the proper timing (since human life and reproduction is subject to a proper cycle just as vegetative life and reproduction) for pairing brides and grooms. This in turn leads to the birth of less excellent children to more excellent parents, and more excellent children to less excellent parent. Even the best of the children born in this time will be less than excellent guardian material, and when they become guardians they will let slide practices necessary for the proper maintenance of the aristocratic constitution, such as musical training of the best youth. As a result the guardians raised by these will lack the musical training that leads to harmony of temperament, and will instead be lead by their sense of honor or *thumos* rather than the rational part of the soul.

A first point to criticize Korsgaard here is the idea that the aristocratic constitution is analogous, or meant to stand for, a constitutively *principle-governed* constitution. Some commentators, for example, urge that the so-called aristocratic constitution is not the state of the soul that adheres to some set of principles, but rather one which is ideally

---

<sup>75</sup> In Book 8 (2004 546a, 547a) of *The Republic*.

unified through having a proper outlook.<sup>76</sup> A second point of critique is that Plato urges that this ideal constitution will itself decay: “not even one so constituted will last forever.”(2004, 546a) Plato seems to think that the aristocratic constitution, whether the constitution of a *polis* or an individual soul, cannot be maintained.

Although Korsgaard might claim to accommodate this point by urging that we are obviously imperfectly rational and so our achievement falls along a continuum, the idea of disintegration makes sense only to the extent that the aristocratic soul is not unified in virtue of the satisfaction of constitutive principles. The conception of deliberation and constitution to which Korsgaard helps herself in building upon Plato is not usefully construed as a *constitutive principle*-governed conception because this cannot do the normative work that Korsgaard needs. So looking at the source of the guiding metaphor fails to vindicate its aptness. The analogy is specious.

**3.5** Alternatively, Korsgaard can argue that rational standards are normative because we are *committed* to performing rational activity. This too can be taken in two senses, one expressible by the locution that we “have committed”, the other by we “are committed”. Either we *have at some point* committed to performing rational activity or we *are* committed, in the sense of being necessitated, to performing rational activity. Korsgaard means the second sense. She at no point references or relies upon a single explicit commitment which has normative authority for the agent in question, and to saddle her

---

<sup>76</sup> Talbot Brewer, 2010

with this view would be uncharitable. She does, however continually urge that we have no option but to form a conception of the world, and that the purpose of a conception of the world is to enable us to navigate it successfully. This commitment is reconstructed as rational activity. That we are so necessitated is a way for Korsgaard to give content to the idea that we *are* committed to the rational principles.

But this attempted vindication of the normativity of rational standards collapses back into incoherence. In addition to the typical objections which say that to suggest that reasons can be generated simply through commitment entails an implausible bootstrapping (mentioned above), the conception of *commitment* at work in Korsgaard's view is unhelpful. The very possibility of committing to the activity of reason, granting that it has the structure and nature by which Korsgaard describes it, would necessarily require that one engage in the activity of reason. There is *no* outside to the activity of reason that would enable one to commit to it – it is normatively barren. So on this approach the content of commitment in its normal sense is drained away. Above, I urged that this means just that we are constituted such that we engage in the activity of reason. But, as suggested earlier, this is a reason of the wrong kind for the normative task at hand, at least by Korsgaard's lights. The suggestion that we are necessitated to conform to the standards of rationality, yet again, is incompatible with their normativity: if we *have no option but* to adhere to rational principles, then how does it make sense to say that we *should* adhere to them?

A final possibility for Korsgaard is the more authentically Kantian position which



would say that the application of rational principles is a condition on our having determinate experience at all. On this view, one could not so much as have a potential ground of belief in view that one might transform into a reason through endorsement without presupposing the operation of rational principles. This view would also render talk of the *application* of rational principles misleading at best. The fully Kantian view of reason and rational principles would *not* see those phenomena as equivalent to the deliberative stance in the way that Korsgaard does simply because deliberation must be *about* something, and there is nothing for deliberation to be *about* unless rational principles structure the presentation of sensible content as intentional experience. On this view, then, the normativity of rational standard is likewise put into question – they are simply what it is to have awareness of objects as objects and the capacity to think about them, and so cannot be what makes it true that we “ought” to adhere to them or that it is “good” to adhere to them.

Korsgaard approximates this approach to the extent that one cannot exempt oneself from the activity of reason save by engaging in it. But even so, her account relies on the possibility that we can encounter our representations as potential grounds independent of rational principles. Korsgaard needs to provide a fuller account of cognition in order to illustrate how representations are formed and in virtue of what they are encountered as potential grounds (are there principles of cognition *other* than rational principles that do this work, or is a purely naturalistic account of the construction of representations her view?). But even were she to do so, her account would still be unable

to vindicate the normativity of rational standards without substantial revision.

#### 4 CONCLUSION

We know that we think, and we know that we can think well or think badly: we need not take it as an article of faith that we *should* think well. To the extent that she attempts to give an account on which it is true that one should not violate *modus ponens*, Korsgaard should be applauded. Some philosophers do not trouble themselves even this much. They simply take it as granted that we should adhere to rational principles, admitting of no explanation. Although this view is not to be rejected outright unless it is simply stated in the absence of some substantive account for why we have no option but to fall back to it, it strikes me as defeatist.

I hope to have shown not only that Korsgaard's account of the normativity of rational standards is unsatisfactory, but that her conception of reason as the most fundamental activity as well. The idea of constructivism served as a framing device for her position throughout, and although I think that it is generally correct to construe her position as some sort of constructivism, it is best not to criticize general positions but rather the details of the view in question. This is what I have tried to do.

## BIBLIOGRAPHY

- Aristotle      *The Metaphysics* Hugh Lawson-Tancred, trans. Penguin. London: 2004
- Audi, Robert “Theoretical Reason: Its Sources, Structure, and Scope” *Oxford Handbook of Rationality* Mele and Rawlings, eds. Oxford University Press. London: 2004
- Bridges, Jason “Rule-Following Skepticism, Properly So Called” *Skepticism, Meaning and Justification*. James Conant and Andrea Kern eds. Routledge: 2009
- Broome, John “Normative Requirements” *Ratio*: 12: 1999
- “The Unity of Reasoning?” from *Sphere of Reason* Simon Robertson, ed. Oxford University Press. London: 2009
- “Reasons” from *Reason and Value: Essays on the Moral Philosophy of Joseph Raz*. Wallace, et al. ed. Oxford University Press. London: 2006
- “Does Rationality Give us Reasons?” *Philosophical Issues*, 15 (2005)
- Finlay, Steven “Oughts and Ends” *Philosophical Studies* 143(3) 2009: 315-40
- Frede, Michael “Introduction” *Rationality in Greek Thought*. Frede, Striker, eds. Oxford University Press. London: 1996
- Frege, Gottlob: *Frege Reader*. Michael Beaney, ed. Blackwell Publishing. Malden: 1997
- Ginsborg, Hannah “Primitive Normativity and Skepticism about Rules”
- Hurka, Thomas. *Perfectionism*. Oxford University Press. London: 1993
- Kant, Immanuel *The Critique of Pure Reason* Wood and Guyer, eds. Cambridge University Press Cambridge: 1999
- *Introductory Lectures on Logic*. Barnes and Noble. New York: 2008
- Kolodny, Niko “Why Be Rational?” *Mind*, 114:455: 2005
- “State or Process Requirements?” *Mind*, 116:462: 2007

“Why be Disposed to be Coherent?” *Ethics*, 118: 2008

Korsgaard, Christine *Sources of Normativity*. Cambridge University Press.  
Cambridge:1996

----- *Self-Constitution: Agency, Identity, Integrity*. Oxford University  
Press. London: 2009

----- *The Constitution of Agency: Essays on Practical Reason and  
Moral Psychology*. Oxford University Press. London: 2008

----- “The Activity of Reason” *Proceedings and Addresses of the APA*  
83:2: 2009

----- “Normativity, Necessity, and the synthetic a priori: A Response to  
Derek Parfit”

Larmore, Charles *The Autonomy of Morality* (2008) Cambridge University Press.  
London: 2008

----- “Reflection and Morality” *Social Philosophy and Policy* 27-2: 2010

McDowell, John *The Engaged Intellect: Philosophical Essays*. Harvard University Press.  
Cambridge: 2009

Nozick, Robert *The Nature of Rationality*. Princeton University Press. Princeton: 1993

O'Neill Onora *Kant: Rationality as Practical Reason*. The Oxford Handbook of  
Rationality, Mele and Rawling, eds. Oxford University Press. London:  
2004

Railton, Peter: “Wittgenstein on The Normativity of Logic” *New Essays on the A Priori*:  
Oxford Scholarship Online Monographs: 2000

Raz, Joseph: *Practical Reason and Norms*. Oxford University Press. London: 1975

----- “Reasons: Explanatory and Normative” (draft) 2007

- Rawls, John: "Two Concepts of Rules" *The Philosophical Review* 64: 1955
- Scanlon, T.M.: *What We Owe to Each Other*. Harvard University Press. Cambridge: 1998
- Shah, Nishi. "How Truth Governs Belief". *The Philosophical Review*, 112, no.4: 2003
- Thomson, Judith Jarvis: *Normativity*. Carus Publishing. Peru, Illinois: 2008
- Tolley, Clinton: "Kant on The Nature of Logical Laws" *Philosophical Topics* Vol. 34, Nos 1&2: 2006
- Wallace, R Jay: "Constructivism about Normativity: Some Pitfalls" *Normativity and the Will: Selected Essays on Moral Psychology and Practical Reason*. Clarendon Press. Oxford: 2006
- Wedgewood, Ralph: *The Nature of Normativity*. Oxford University Press. London: 2007





## VITA

Robert Colin English

[Robertcolinenglish@gmail.com](mailto:Robertcolinenglish@gmail.com)

### **Education:**

The University of Mississippi Aug 2009 - May 2011

M.A. Philosophy

Graduate Teaching Assistant: Introduction to Philosophy, Introduction to Logic

Thesis Title: Korsgaard on Reason and the Normativity of Rationality

Adviser: Robert Barnard, Associate Professor

The University of Chicago Sep 2008 – June 2009

M.A. Humanities/Philosophy Concentration

Thesis Title: Agency and Autonomy: Korsgaard, Kant, and the Constitution of the Self

Adviser: Robert B. Pippin, Evelyn Stefansson Nef Distinguished Service Professor of Social Thought, Philosophy, and in the College at the University of Chicago.

Website: <http://home.uchicago.edu/~rbp1/>

Appalachian State University Aug 2002 – May 2007

Six Hours of Masters Level English Coursework

Author of “The Borders of Ekphrasis and the Boundaries of Art: The Challenge of Gendered Ekphrasis in George Eliot, Charlotte Bronte, and Mary Elizabeth Coleridge.”  
Forthcoming in *Art Objects and Women’s Words: Women’s Ekphrastic Writing 1785 to the present*. Ed. Jill Ehnenn

B.A. in Philosophy and Religion, Cum Laude  
English and German double minor

- Honors Student and President of Philosophy Club
- Co-editor of student literary publication
- Co-editor of ASU Writing Center Newsletter
- Co-organizer of the ASU Philosophy Speaker Series
- Teacher's Assistant for various philosophy courses
- Tutor for Philosophy, English, History, Communications and Geology
- Published in student literary publications and ASU Writing Center Newsletter

Appalachian State University Writing Center  
Writing Consultant

Jan 2007 – May 2007

- Work with clients to facilitate all aspects of the writing process, from brainstorming to editing and formatting.
- Participate in ongoing self and peer evaluative discussions with other writing consultants online (to be reviewed by writing center director).
- Develop *The re-Visioner*, the ASU Writing Center Newsletter, with fellow editors.
- Promote the Writing Center through various media.

ASU Learning Assistance  
Tutor

Jan 2007 – May 2007

- Assist students with various subjects, including Philosophy, English Literature, History, and Geology.
- Help develop and refine study skills.
- Aid in reading comprehension.
- Assist in paper writing (thesis formulation, argumentation, rhetorical structure).