

# The Teleology of Rational Capacities

---

## 1. Virtue Epistemology and “Epistemic Capacities”: A Critique

In describing beliefs as acts of a rational capacity for knowledge, we represent them as acts that, when they are knowledge, realize an end that the believing subject is conscious of. In recent decades, the literature on epistemology has come to rediscover this fundamental teleology of the mind. That is, the literature has regained a consciousness of the fact that the theoretical life of the mind is, in a certain sense, a teleological activity. The growing trend in “virtue epistemology” finds its distinctive contribution in this rediscovery. **Any answer to the question of what knowledge is and how it is possible requires that we grasp the acts we are seeking to understand as acts that realize a *telos*, an end.** Recent debates about intellectual virtues, by authors such as Zagzebski and Montmarquet, as well as alternative language about epistemic capacities, which authors such as Sosa and Greco prefer, stem from this common motive—namely, to register that the idea of virtues and rational capacities is the idea of something that makes reference to an end. The conception of capacities and virtues that is relevant to epistemology, Sosa tells us, is a “teleological conception,” according to which rational capacities and virtues have a teleological

structure.<sup>2</sup> Someone who has a virtue, Zagzebski tells us, is so constituted that her act's relation to its respective end ensures that she is reliable "in bringing about that end."<sup>3</sup>

We can characterize the shared insight of so-called virtue epistemology through the following three claims:

- (1) Believing is an activity, whose end is the truth.
- (2) Capacities are general properties of subjects that explain the occurrence of acts that fulfill the end in terms of which the capacity is defined.
- (3) Therefore, capacities that are defined as having the truth as their end make it intelligible how there can be beliefs that are non-accidentally true, i.e., that constitute knowledge.

With this general insight, virtue epistemology takes itself to be in a position to address the problem that, as we have shown in the previous chapters, structures all discussions in epistemology: viz., the problem of non-accidental truth, which is definitive of the concept of knowledge. We concur with virtue epistemologists that this problem has gone unresolved in contemporary epistemology, insofar as it seeks to do without the idea of virtues or capacities. The intelligibility of the concept of knowledge, however, stands or falls with the solution to this problem. What ultimately makes virtue epistemology distinctive is its claim that the key to solving the accidentality-problem of which the so called Gettier-cases have reminded us is to introduce notions of intellectual virtues or epistemic capacities as fundamental epistemological concepts.

Reference to intellectual virtues or epistemic capacities is motivated by the insight that, in introducing such things, we are dealing with a kind of "cause" that explains why acts that are directed at a particular end—"performances" that have an "essential aim" or "an aim inherent in [them]," as Sosa puts it—are such as to "attain" the end at which they

<sup>2</sup>Ibid.

<sup>3</sup>Zagzebski, *Virtues of the Mind*, 137.

are directed.<sup>4</sup> Virtues and capacities, in the sense relevant here, have to be understood, according to virtue epistemology, as the sorts of “causes” of particular acts that serve to explain why those acts accord with the end at which they are directed. Just as the capacity to play baseball describes a kind of cause that explains why someone who has this capacity performs acts that accord with the end of playing baseball, we can likewise explain knowledge as an act whose “cause” is an epistemic capacity that explains why the act is so constituted that it accords with its end—namely, to be an act of knowledge. And just as we can say that playing baseball consists in the exercise of a kind of “agency,” which characterizes a subject that has the capacity to play baseball, so too can we describe knowledge as the exercise of a kind of “agency,” which characterizes a subject who has a certain “epistemic capacity.”<sup>5</sup>

Let us call this the distinctive insight of virtue epistemology: we employ the concept of intellectual virtues or epistemic capacities in our analysis of knowledge because they provide an explanation for the agreement of an act with an end that is intrinsic to it *qua* act, or, as Sosa puts it, for its agreement with “an aim inherent in it”. In order to understand the idea of non-accidentally true belief that our concept of knowledge carries with it, the suggestion goes, we have to understand knowledge as an act of a capacity whose end is truth. This insight enables us to view the normativity of knowledge as a special case of a more general, everyday, and familiar form of normativity, which has application whenever someone succeeds in doing something by exercising the relevant capacity for doing such things. When we say that someone knows something, what we are saying, on this view, is that she has succeeded in forming a true belief precisely in virtue of exercising an epistemic capacity in forming that belief. Knowledge is, in this respect, like playing baseball or skiing. John Greco expresses the view as follows:

[K]nowledge is a kind of success from ability. Put another way, knowledge is a kind of achievement, or a kind of success for which the

<sup>4</sup> Sosa, *Knowing Full Well*, 14–15.

<sup>5</sup> *Ibid.*, 19.

knower deserves credit. And in general, success from ability (i.e., achievement) has special value and deserves a special sort of credit. This is a ubiquitous and perfectly familiar sort of normativity. Thus we credit people for their athletic achievements, for their artistic achievements, and for their moral achievements. We also credit people for their intellectual achievements. Epistemic normativity is an instance of a more general, familiar kind.<sup>6</sup>

In what follows, our aim is to understand the teleological structure that characterizes a rational capacity, as it is exhibited in such ordinary activities as playing baseball or skiing, in order to then develop an account of the teleological structure of a capacity for knowledge. I will argue that virtue epistemology fails to do justice to what I have called its distinctive insight because it labors under a false conception of the kind of causality that characterizes a capacity for knowledge.

Sosa introduces the idea of capacities in the following way: A shot at a target can hit the bull's-eye without such success manifesting any competence on the part of the shooter. When that happens, we say that the outcome is just an accident, the result of pure chance. But it can also happen that a shot hits the bull's-eye because of the competence of the marksman. A competent marksman doesn't hit the bull's-eye just by accident. In such a case, we have an explanation for her success. It is her competence in shooting that explains why she hits the bull's-eye. The idea here is that we can evaluate beliefs in just this fashion. A belief may be true simply by accident—namely, when it isn't the manifestation of any relevant competence. But a belief can also be true on account of the subject's capacity. What accounts for the fact that the belief is non-accidentally true, in the relevant sense, is a particular competence that explains the occurrence of a true belief.<sup>7</sup>

Sosa accordingly defines knowledge as follows: "Belief amounts to knowledge when [...] its correctness is attributable to a competence exercised in appropriate conditions."<sup>8</sup> John Greco similarly claims:

<sup>6</sup> Greco, *Achieving Knowledge*, 7.

<sup>7</sup> *Ibid.*, 23.

<sup>8</sup> *Ibid.*, 92.

“S knows that *p* if and only if S believes the truth (with respect to *p*) because S’s belief that *p* is produced by intellectual ability.”<sup>9</sup>

Sosa defines a “competence” as follows: “[A] competence is a disposition, one with a basis resident in the competent agent, one that would in appropriately normal conditions ensure (or make highly likely) the success of any relevant performance issued by it.”<sup>10</sup> Thus, according to Sosa, a capacity can be analyzed into two components: a disposition to bring about certain acts and a high rate of “successful” acts under “appropriately normal” circumstances, i.e., a high rate of acts that realize the end of the competence in question. Sosa accordingly describes the requisite criterion of success by saying that a disposition can have the status of a “competence” or a “capacity” if and only if the disposition in question “is sufficiently reliable, at least in its distinctively appropriate conditions.”<sup>11</sup> That is to say, a certain disposition counts as a competence of the relevant sort if and only if it is sufficiently reliable, under appropriate conditions, in bringing about acts that are in accordance with the end that defines the competence in question.<sup>12</sup>

According to this account, a capacity, such as the competence to bring about *y*-acts, can be analyzed into the following two components:

- (1) the disposition to bring about *x*-acts

and

- (2) a high rate of acts resulting from (1) that are in accordance with the concept of *y*-acts, under the conditions that are appropriate for such acts.

Sosa then brings this general characterization of capacities to bear on the idea of an epistemic competence and analyzes the latter in terms of:

<sup>9</sup>Ibid., 71.

<sup>10</sup>Sosa, *A Virtue Epistemology*, 29.

<sup>11</sup>Ibid.

<sup>12</sup>Ibid.

- (1) the disposition to form beliefs on the basis of “intellectual appearances”

and

- (2) a high rate of acts that spring from (1) that are in accordance with the concept of a true belief, under circumstances that are appropriate for such accord, i.e., a high rate of true beliefs that are formed on the basis of “intellectual appearances.”

The concept of “intellectual appearances” is meant to signify, for Sosa, a particular type of conceptual representation, which arises from sensible affection and which thereby forms a potential reasons for belief—such as the state of a subject who is enjoying the visual appearance that *p*.

Let us ask, then, what exactly it is, on this analysis, that qualifies a mere disposition to form beliefs on the basis of so-called “intellectual appearances” as an epistemic competence. Clearly such a disposition would need to provide an explanation for the success of certain acts. But what kind of explanation? One way to bring to light the upshot of this analysis is to note what is not required for a disposition to qualify as an epistemic competence, according to Sosa. For a disposition to qualify as an epistemic competence, it need not be such as to explain a belief in a way that rules out that a belief that is explained in that way could be false and hence, not knowledge. According to Sosa, the criterion of success that is definitive of competences is rather that someone who possesses an epistemic competence and actualizes it under appropriate circumstances meets, with sufficient frequency, with the sort of success that is definitive of that competence. This means that, on Sosa’s account, it is perfectly intelligible for someone to possess the relevant competence and exercise it under conditions that are appropriate for its exercise and yet fail to exercise it successfully. Such a case is intelligible because the idea of someone possessing a competence doesn’t mean that she is in possession of something that forecloses the possibility that an exercise, in the appropriate circumstances, can fail to be successful. It only means that she possesses something whose exercise, in the appropriate circumstances, is successful for the most part. Under

these so-called appropriate circumstances, two outcomes are logically possible: the exercise of the competence succeeds, or it does not.

However, this means that whenever the competence is exercised in the appropriate circumstances, its success or failure must be a matter of luck—a chance occurrence. The reason for this is that, given Sosa's description of the successful case, it is logically possible that there can be an unsuccessful case that is in no way different from the successful one, so that there is nothing we can appeal to in order to explain why the one case was successful and the other was not. Nor can we explain the successful case simply by appealing to the competence itself. For the competence is defined solely by the fact that its exercise mostly meets with success in the appropriate circumstances. So the competence, just as such, cannot provide an explanation for the successful case. Nor can we explain the successful case by pointing to specific circumstances that explain why the exercise met with success on this occasion. For, by definition, the same circumstances can also obtain in a case where the exercise of the competence fails. And if we cannot explain the success of the case either in terms of the competence or in terms of the prevailing circumstances, then its success can only be a matter of chance, a matter of luck. If that is so, however, the idea of epistemic competences or capacities, thus conceived, cannot solve the accidentality problem for which it was introduced.<sup>13</sup>

So far, we can glean the following general insight from this failure: Any epistemology that cannot construe an epistemic capacity as an explanation of acts of knowing that rules it out that something that is explained in that manner is not knowledge will inevitably fail to resolve the accidentality problem. That is to say, any epistemology that understands epistemic capacities as general characteristics of a subject

<sup>13</sup>Duncan Pritchard pursues a quite different argumentative route to arrive at the same verdict, namely that virtue epistemology cannot solve the accidentality problem. See Pritchard, "Virtue Epistemology and Epistemic Luck." But Pritchard is wrong to conclude from this failure that there is no reason to endorse an epistemology that treats the concept of epistemic capacities as fundamental. Virtue epistemology fails not because it treats epistemic capacities as fundamental but because it misunderstands the very idea of such capacities. For a more detailed critique of virtue epistemology, see my essay "Knowledge as a Fallible Capacity."

that do not explain a belief in a way that rules it out that a belief thus explained could be false, will not be able to solve the accidentality problem. An epistemic capacity must be a general characteristic of a subject that *guarantees* truth. That is, it must be a characteristic that rules out the formation of a false belief in the circumstances that are appropriate for its exercise.

At this point an obvious objection comes to mind—and it is presumably an objection along these lines that explains why virtue epistemology doesn't even consider the idea of a truth-guaranteeing capacity of a subject as a possible option. **The objection is that it is unreasonable to demand of someone who possesses an epistemic capacity that the beliefs produced through the exercise of that capacity are *always* true.** That simply demands too much of the cognizing subject. It seems more reasonable to demand that, in the appropriate circumstances, an epistemically competent person will *often* form a true belief, not that she will always form true beliefs. To demand the latter would be to demand epistemic infallibility. For it would mean that someone has the capacity to acquire knowledge only if her capacity rules out, as a matter of principle, that she can be mistaken. And that is a conception of knowledge that cannot sensibly be attributed to human beings. The thought that human beings are capable of knowledge cannot reasonably be cashed out in a way that would require us to be epistemically infallible.

Yet if we consider this objection more closely, it is easy to see that it rests on a misunderstanding. The objection mistakenly equates two thoughts that must be held apart. Namely:

- (1) The idea of an epistemic capacity implies that beliefs formed through exercises of that capacity in the appropriate circumstances are always true.
- (2) Bearers of an epistemic capacity are epistemically infallible.

How does one come to assimilate these two thoughts? For there is quite obviously another way of doing justice to the fallibility of beliefs without at all disputing thought (1)—a way that we have already developed in Part Three. It will be helpful to briefly revisit this alternative understanding of fallibility in the context of a capacity familiar to us

all: the capacity for speech. According to thought (1), if someone has the capacity for speech, then whenever the circumstances are appropriate for the exercise of that capacity, it is impossible for an exercise of the capacity to fail. Of course, it can happen that the appropriate circumstances obtain but the bearer of the capacity chooses not to exercise it. Thought (1) only rules out cases in which the capacity is exercised in the appropriate conditions and yet the exercise goes awry. But the thought that someone is able to speak is quite obviously not the same as the thought that it is impossible for her to make mistakes in speaking. Someone who can speak can also misspeak—she can make a grammatical error, misuse a word, and so on. So how do these thoughts go together? It is at least clear how we shouldn't interpret such cases—namely, as cases in which the appropriate circumstances obtain and yet the exercise of the capacity misfires. And this itself indicates how we ought to understand such cases: namely, as cases in which the exercise of the capacity misfires in some respect precisely because the circumstances appropriate for its exercise do not obtain. More precisely: because circumstances obtain that hinder or restrict the successful exercise of the capacity.

Once we construe cases of success and failure in this manner, we can satisfy the demand elaborated above, according to which we require a conception of an epistemic capacity as something truth-guaranteeing. This conception doesn't preclude but instead ensures that there can be conditions under which someone will be hindered from exercising her epistemic capacity "correctly," or "properly," or "successfully."

With this understanding of capacities in hand, we can easily reconcile the possibility of error with the idea of truth-guaranteeing capacity. If an epistemic capacity is a truth-guaranteeing capacity, then it explains a successful case in a manner that rules it out that a case that is explained in that manner could not have been successful. By contrast, any case that is not successful, but instead is defective in some way or other, can be explained only by invoking (in addition to the relevant capacity) unfavorable prevailing circumstances that explain why the exercise of the capacity failed in one way or another. We called such a capacity a fallible capacity. A fallible capacity is one that cannot be successfully exercised under all possible circumstances. It is one whose successful exercise depends on the presence of favorable circumstances. Yet, to

say that a capacity is fallible is not, as Sosa thinks, to say that its successful exercise under favorable circumstances is only very likely. It is to say that the favorable circumstances on which its successful exercise depends are part of what it means to exercise it successfully.

## 2. Rational Capacities as a Species of Teleological Causality: A Kantian Approach

In section 1 of this chapter, I claimed that the concept of epistemic capacities, as it has been invoked in recent virtue epistemology, is incapable of resolving the problem it is meant to address because it is blind to the idea of a truth-guaranteeing capacity. Why is contemporary virtue epistemology blind to the idea of a truth-guaranteeing capacity?

Virtue epistemology seeks to claim that the relevant concept of capacities that we require in epistemology contains the idea of an end. The conception of virtues and capacities that is relevant to epistemology, to quote Sosa again, is a “teleological conception.” It is striking, however, that contemporary virtue epistemology gives us no indication of what it takes an end to be. This is no arbitrary omission. For as we will see in what follows, virtue epistemologists do not feel the need to elucidate the concept of an end because their account of epistemic capacities hinges on the implicit presupposition that describing an epistemic capacity as the “cause” of an act involves a kind of causality that is independent of the end that defines this capacity. That the concept of an epistemic capacity is a concept that contains the idea of an end does not mean, on the contemporary conception, that it is the concept of something that has a special form of causality, distinct from the form of causality proper to those things whose concept does not contain the idea of an end. Now my aim in what follows is not to provide a general refutation of this conception of capacities. For we have already seen that this conception of capacities refutes itself when applied to the case of knowledge. My aim instead is to contrast this conception of capacities with the conception of capacities we have developed in Part Three, by unfolding the different understanding of the causality of capacities that this conception entails. That is, our task now is to make explicit the idea of causality that has been implicitly at work in our understanding

of capacities. We will do this by first looking at an alternative approach to capacities, which brings into focus the idea that the relevant concept of a capacity is a “teleological” concept, and which takes its point of departure from a reflection on precisely that feature of the relevant concept of capacities.

A paradigm example of such an alternative conception of capacities can be found, *inter alia*, in the work of Kant. Kant is particularly helpful in addressing our question, for two reasons. First, Kant argues that the idea of a “teleological causality” is a sort of causality that is *sui generis*. That is to say, Kant develops an account of the idea of a “teleological causality” by arguing that it contrasts—formally—with another kind of causality, which he calls “mechanical causality.” His highly abstract conception of “teleological causality” will provide our basis for understanding the teleological structure of capacities and, more narrowly, the teleological structure of a capacity for knowledge. This will allow us to conceive of the failure of contemporary virtue epistemology as the consequence of a misunderstanding about the kind of causality that is exhibited by capacities. Virtue epistemology conceives of capacities as instances of mechanical causality rather than as instances of teleological causality. The second reason Kant is helpful here is that the concept of a capacity for knowledge lies at the very center of Kant’s metaphysics of mind in a way that is true of few other philosophical systems. Though Kant does not himself explicitly develop the concept of a capacity for knowledge as the concept of a teleological kind of causality, this teleological understanding is implicit throughout Kant’s work. Moreover, some crucial aspects of his account of knowledge, as I will argue in what follows, can be made intelligible on the basis of this teleological conception and will likewise help us to deepen our account of knowledge in terms of a rational capacity we have developed so far.

It is a telling feature of the literature on Kant that, thus far, very little work has been done on Kant’s notion of a rational capacity, despite the fact that even a superficial reading of the text reveals it to be one of his key terms.<sup>14</sup> In what follows I will therefore provide some basic elements

<sup>14</sup> See, for example, the collection of papers in Perler, ed., *Faculties*, which give an historical overview of the role of the notion of faculties through the philosophical tradition. In his contribution “Faculties in Kant and German Idealism,” Johannes Haag

of a Kantian account of the idea of a capacity for knowledge by first considering Kant's abstract notion of an end or "telos". I will then apply this account of ends to the idea of a capacity that is defined by the end of knowledge.

In the third Critique, Kant first introduces the concept of an end in its most abstract sense. An end, he tells us, is "the object of a concept insofar as the latter is considered the cause (the real ground of the possibility) of the former."<sup>15</sup> The concept of an end is, accordingly, "the concept of an object, insofar as it [the concept] contains the ground of the possibility of that object."<sup>16</sup> When we characterize an object as an end, we are thus determining an object through a concept that serves to explain the reality of that object, in a particular sense.

To get an initial grip on this highly abstract definition, it is helpful to bring it into contact with other formulations where Kant elucidates the concept of an end by employing causal terminology. In §64 of the third Critique, for example, Kant elucidates the concept of an end by saying that to characterize an object as an end is to understand it as the effect of a cause "whose efficacy [*Vermögen zu wirken*] is determined through concepts."<sup>17</sup> What we need to understand in order to comprehend the

nicely brings out the widespread conception of the role that "faculties" play in Kant's account of knowledge. Kant's "transcendental approach"—as opposed to, for example, a psychologistic approach—brings with it, Haag argues, the assumption that the "metaphysical status of the faculties invoked in this type of reasoning no longer carries any importance" (199). This opposition between a "transcendental approach" and a "metaphysical project," however, is unfortunate, for it seems to suggest that the point of Kant's so-called transcendental approach is to merely justify the "introduction of a particular faculty" into our philosophical account without itself providing an understanding of what that thing is that is thus introduced. Kant's "transcendental approach" is an attempt to answer the question of how it is possible to have representations with objective purport. This brings him to "introduce" the idea of sensibility and understanding as the two stems of a receptive capacity for knowledge whose very idea he then seeks to unfold in the course of his inquiry. Now, if one thinks that the metaphysical status of this capacity "no longer carries any importance," one fails to appreciate the very task that Kant sets himself in the course of his inquiry, especially in the Transcendental Deduction: namely, to give an account of what a receptive capacity for knowledge is by showing us how we have to conceive of it in order for this idea to be so much as intelligible.

<sup>15</sup> Kant, *Kritik der Urteilskraft* [Power of Judgment], §10, AA 5:220.

<sup>16</sup> *Ibid.*, introduction, §IV, AA 5:180.

<sup>17</sup> *Ibid.*, §65, AA 5:369.

idea of an end, Kant tells us, is the idea of a special kind of cause: namely, a cause “whose efficacy is determined through concepts.” According to Kant, this means that we must distinguish between two irreducibly different kinds of causal connection. On the one hand, we can think of a causal connection among elements that involves a one-sided dependence of one element on another—i.e., a connection in which the element that is regarded as the effect cannot also serve as the cause of the element whose effect it is. We typically call this sort of causal connection “that of efficient causes,” or *nexus effectivus*.<sup>18</sup> Following Kant and traditional usage, we can also describe such a form of causality as “mechanical causality.” In addition to this kind of causal connection, however, we can also conceive of a further sort, which involves a reciprocal dependency between the relevant elements. This would be a causal connection “in which the thing which is at one point designated the effect nevertheless deserves [ . . . ] to be called the cause of the very thing whose effect it is.”<sup>19</sup> We can call this kind of causal connection “teleological causality,” or *nexus finalis*.

It is telling that Kant gives such an abstract explanation of what it means to represent something as an end. Obviously Kant understands this to be a purely formal characterization, which, as such, contains no indication of the sort of objects that can enter into such a causal connection as elements.

So let us consider Kant’s most prominent illustration of this account in the practical realm, while bearing in mind that we should understand it as an exemplary case of a more general phenomenon. The example has to do with building a house. Kant writes: “In practice (namely, in art) it is easy to find such connections, in which, for example, a house is the cause of the money that can be taken in as rent, while it is also the case that, conversely, the representation of this possible income was the cause of the house’s being built.”<sup>20</sup> Now in what sense does the thing that at one point is designated the effect also deserve to be called a cause? In this sort of causal connection—between the house and the income from rent—the thing that is at one point designated an effect

<sup>18</sup> Ibid., §65, AA 5:372.

<sup>19</sup> Ibid.

<sup>20</sup> Ibid.

(the rent) also deserves to be called a cause (of the house) in the sense that the representation of that effect is what explains the presence of the thing that is the cause of that very effect. The representation of the effect is what explains the existence of the house, in the sense that a rational creature who has this representation of a house is determined, through that representation, to perform precisely those actions that lead to the existence of a house. To give a complete account of the cause of the rental income, in this example, we would have to include some reference to a rational being who has a certain conceptual representation of the effect of the house, which representation leads her to build the house.

Thus, it is clear that when Kant describes a teleological causal connection as a cause “whose efficacy is determined through concepts,” what he has in mind, in this context, is a rational subject capable of intentional action. A rational agent embodies a cause whose efficacy is determined through concepts in the sense that such a being produces things precisely by acting in accordance with a conceptual representation of the things she produces (or is going to produce). The things that rational agents produce in this manner—namely, in accordance with a concept of the thing in question—accordingly stand in a special relation to that concept on account of the special manner in which they were produced. In particular, their agreement with the concepts that represent them is no mere accident but a matter of necessity. We can therefore say that an object that exists as an end is one that necessarily agrees with the concept of that end. This enables us to understand what Kant means by saying that an end is the object of a concept insofar as the concept contains the ground of the actuality of the object. On this Kantian conception of an end, it is logically ruled out that an object that exists as an end is merely accidentally in agreement with the concept of that end.

We can therefore express Kant’s formal characterization of the idea of teleological causality in the following minimal way: It is the idea of a causal connection in which the thing that is represented as the cause is logically dependent on the thing that is represented as the effect. Let’s now apply this account of the idea of a teleological causality to the idea of a capacity that is defined by the end of knowledge. We want thereby

to take into account that the capacity we are interested in cannot be exercised under all possible circumstances—that is, that there is a distinction between circumstances that are appropriate for its exercise and circumstances that are less than appropriate for its exercise. This gives us the following, preliminary understanding of the kind of cause that a capacity for knowledge is: *A capacity for knowledge is a cause that is logically dependent on its effect (viz., knowledge) in the sense that exercises of that capacity, as such, fall under the concept of knowledge as being in either perfect or imperfect agreement with the concept of knowledge.*

As we proceed we will develop a more determinate account that specifies the idea of teleological causality—in terms of a logical dependency of cause on effect—as it applies to a capacity for knowledge. Yet our preliminary characterization of the kind of cause that a capacity for knowledge is already provides a basis for getting clear about the contrast between this teleological conception of a capacity for knowledge and the conception of capacities presupposed by contemporary virtue epistemology. For contemporary virtue epistemology is defined by a conception of capacities that conflicts with the above characterization. In particular, it understands knowledge as the effect of a cause—namely, an epistemic capacity—that is logically independent of the end at which it is directed. The capacity *qua* cause is logically independent of its end (knowledge) in the sense that a state of knowledge is understood as the effect of a cause whose causality can be fully described without any employment of the concept of the thing at which its activity is directed *qua* end.

Recall that the fundamental characterization of an epistemic capacity that we find in contemporary virtue epistemology consists of two elements: (1) the disposition to produce beliefs on the basis of intellectual appearances, and (2) a high rate of agreement, under appropriate circumstances, between beliefs produced on the basis of intellectual appearances, and the concept of a true belief. While (1) describes the causality of this capacity as involving the production of beliefs, (2) describes the accord of beliefs thus explained with the end of that capacity, which Sosa characterizes in terms of true beliefs. The central feature of this account of an epistemic capacity lies in the fact that the first element of the account is logically independent of the second element. The sort of

causality that explains the occurrence of beliefs does not, as such, explain the agreement of these beliefs with the concept of a true belief. It follows that it is logically excluded, from the very outset, that one can explain an instance of a true belief just by appealing to the relevant sort of causality, i.e., the capacity.

We have followed Kant in calling such a form of causality “mechanical causality.” It is part of Kant’s central argument that an account that conceives of the causality of the relevant capacity as a form of mechanical causality is unable, for logical reasons, to understand the agreement between a particular act and the concept that designates the end of the capacity as a necessary agreement. Thus, any line of thought that seeks to understand a capacity for knowledge as analyzable into two logically independent elements—one that describes the causality of the capacity and another that describes the agreement of its acts with the end of the capacity—will be unable to lay claim to the idea of knowledge as non-accidentally true belief. The idea of a non-accidentally true belief remains unintelligible on such an account because it is impossible, in the context of such an analysis, to explain the truth of a belief through the causality of the capacity in question.

To get a clearer view of what it means to have a mechanistic conception of the causality of the relevant kind of capacities and how such a conception differs from a teleological conception, let’s look at how virtue epistemology analyzes the idea of a case that is successful, not just in one respect or other, but, as Sosa puts it, on all “levels of success.”<sup>21</sup> Consider, once again, the capacity for *y*-acts. On this view, a successful exercise of this capacity for *y*-ing consists in meeting the following three conditions:

- (1) The act agrees with the end of the capacity to *y*.
- (2) The act is a manifestation of the capacity to *y*.
- (3) (1) is true because (2) is true—i.e., the act agrees with the end of the capacity to *y* because the act is a manifestation of the capacity to *y*.

<sup>21</sup> Compare Sosa, *Knowing Full Well*, 1.

If an act fulfills condition (1), Sosa calls it “accurate.” If it fulfills condition (2), Sosa calls it “adroit.” And when it fulfills condition (3)—which implies the fulfillment of (1) and (2)—it is a successful exercise of the capacity, which Sosa would call “apt.”<sup>22</sup> It is thus an essential feature of this conception of a capacity that it is possible for two acts to be identical insofar as both fulfill (1) and (2) but for only one of them to fulfill (1) specifically *because* it fulfills (2). Thus, virtue epistemology holds that it is possible for an act to constitute a manifestation of an epistemic capacity *in the very same sense* that an act of knowing does, yet without itself being an act of knowing. This might be because it does not fulfill the end of the capacity, in which case the act is competent in the sense of being “adroit” but not “accurate,” or it might be because the act does fulfill the end of the capacity but not in virtue of its competence, in which case the act is “accurate” and “adroit” but not “accurate” because “adroit.”

This analysis of a successful exercise of a capacity illustrates that and how the causal efficacy of a capacity is, on this conception, logically independent of the end of that capacity. For on this conception, the causal aspects of our notion of a capacity—which we exploit in speaking of “manifestations,” “actualizations,” or “exercises”—can be understood quite independently of the question of whether a particular act agrees with the end of the capacity. The question whether an act is a manifestation of the capacity to *y* can be settled independently of determining whether it fulfills the end of that capacity. According to virtue epistemology, claims (1) and (2) are logically independent of one another. This is why claim (3) must be separately added to the account.

Now if we instead understand a capacity for knowledge as a form of teleological causality, we precisely deny this logical independence. For we then represent the causal efficacy of a capacity as logically dependent on its end. Whereas a mechanistic conception of capacities takes it to be possible to describe what it is for an act to constitute a manifestation of a capacity without thereby making reference to the *telos* of the capacity in question, a teleological conception of capacities demands

<sup>22</sup>Ibid.

that one refer to the telos of the capacity in order to so much as describe a given act as a manifestation of the capacity in question. It follows that characterizations of acts as manifestations of capacities that cannot be exercised under all possible circumstances must, accordingly, be understood disjunctively: namely, either as perfect manifestations of the capacity, which would then be identical with the perfect realization of its telos, or as manifestations of the capacity that are faulty in one way or another, which would amount to an imperfect realization of the capacity's telos. If a capacity for knowledge exhibits a teleological form of causality, then judging that a given act constitutes a manifestation of that capacity entails a judgment about the act's agreement with the concept of knowledge under which we bring it in characterizing it as a manifestation of a capacity for knowledge.

### 3. Kant's Refutation of the Idea of an "Implanted Subjective Disposition"

We argued that we need a teleological conception of capacities in order to adequately account for the idea of knowledge. A mechanistic conception of capacities will not do. However, a teleological conception of a capacity for knowledge is faced with a question that might threaten its very intelligibility. For the above account of a capacity for knowledge must explain how we can understand the possibility of a "cause" whose efficacy depends on the employment of the concept of knowledge. How can we account for such a "cause"? How, we have to ask, is such a "cause" even possible?

I will develop an answer to this question by taking as my starting point Kant's discussion of the same issue. Kant's most straightforward answer to this question can be found in §27 of the *Critique of Pure Reason*, where he summarizes the argument he has just given in the so-called Transcendental Deduction. For our purposes we do not have to worry about the details of the Deduction but only need to consider its most general ambition, as Kant presents it. The task of the Transcendental Deduction is to demonstrate the "objective validity" of the "pure concepts" that correspond to the forms of judgment. Kant takes himself

to have completed this task by the end of §26. One way to characterize his result, which brings it into contact with the question we now face, is as follows. By the end of §26, Kant takes himself to have shown that the idea of a subject capable of forming judgments about sensibly given objects is identical to the idea of a subject who possesses a capacity for knowledge about sensibly given objects. For the very concepts a subject must possess in order to make judgments about sensibly given objects—the “pure concepts”—are demonstrably valid, *a priori*, of any sensibly given object. However, this characterization of the result of the Transcendental Deduction seems to elicit a worry about the very possibility of such a capacity. Up to this point one might think Kant has managed to show that, and why a sensible being who possesses the concepts constitutive of the unity of a judgment is thereby in possession of a capacity for knowledge of objects of experience. But it is tempting to think that he has not yet addressed the question of how a capacity for knowledge of objects of experience itself is possible. How, one might ask, can there even be such a thing as a capacity for knowledge of objects of experience?

This question rests on the assumption that demonstrating the objective status of the “pure concepts” does not yet provide an answer to this question. The discussion in §27 is meant to show that a proper understanding of the status of the “pure concepts” already contains an answer to this question. Thus, §27 aims to ensure that the idea of “pure concepts” is understood in the right way. It does so by arguing that a certain account of the “cause” of a capacity for knowledge of objects of experience is incompatible with a proper understanding of such a capacity and that there is, in fact, only one way to adequately conceive of it.

The account that Kant wants to rule out as incompatible with a proper understanding of a capacity for knowledge of objects of experience is one that tries to answer the question we raised above in the following way. We can understand a capacity for experiential knowledge in an analogous way to how we understand artifacts. According to this argument, the idea of a capacity for experiential knowledge is a species of a genus of teleological causality whose fundamental understanding is provided by the idea of artifacts. Kant’s concern is to rule out this (mis)interpretation. In so doing, he brings into view an alternative

understanding of the kind of teleological causality that is exhibited by a capacity for knowledge, one that is, in a crucial respect, different from that exhibited by artifacts.

Let us therefore take a look at the case of artifacts and how to account for the kind of teleological causality they exhibit. One of Kant's paradigmatic examples of an artifact in the third Critique is a clock. The defining characteristic of a clock is its ability to tell time. We can express this by saying that, like a capacity for knowledge, a clock is a cause whose efficacy is dependent on the concept of its distinctive effect—namely, telling the time. For it is no accident that the clock has this ability. Rather, the concept of the clock's distinctive effect—namely, telling the time—is, in a certain sense, the cause of its being constituted in the particular way that it is. The concept of its effect is the cause of the clock in the sense that the concept of telling the time becomes efficacious in the actions of a rational subject by determining her actions “in the production and combination of [its] parts” through a conceptual representation of this effect.<sup>23</sup> The fundamental cause of the clock is thus a rational subject—the clockmaker—who produces the clock in accordance with a conceptual representation of its distinctive effect. And a clock that is thus produced in accordance with a concept of this effect is constituted precisely so as to generate this effect. Thus, an object that is produced in such a way will do things that necessarily agree with the concept of telling time.

Kant considers the question of whether a capacity for knowledge of objects of experience, which we want to know how it is possible to possess, can also be understood in this manner. He invites us to think of our cognitive faculty as a “subjective disposition for thinking implanted in us with our very existence [...] which is so ordered by our creator that its use is in precise agreement with the laws of nature in accordance with which experience proceeds (a sort of preformation system of pure reason).”<sup>24</sup>

Kant's proposal here is that we should try to understand our capacity for knowledge of objects of experience in just the same way that we understand a clock's capacity to tell time. We should, accordingly,

<sup>23</sup>Kant, *Kritik der Urteilskraft*, §65, AA 5:373.

<sup>24</sup>Kant, *Kritik der reinen Vernunft*, B167.

imagine that some creator has implanted in us a subjective capacity for thought and that this creator has instituted this capacity for thought in such a way that it agrees “with the laws of nature in accordance with which experience proceeds.” The idea that this subjective faculty for thought is ordered precisely so as to agree with the “laws of nature in accordance with which experience proceeds” means that our creator has equipped this faculty with precisely those concepts whose employment in thought leads to judgments that agree with objects of experience.

Kant formulates several objections to the mooted hypothesis, though he takes only one of them to be “decisive.”<sup>25</sup> The crucial objection is supposed to demonstrate that the mooted proposal is actually incompatible with the concept of knowledge. The objection is that this suggestion can only make sense of the “subjective necessity” of employing the concepts “implanted” in us but cannot account for their “objective necessity.” Kant’s argument runs as follows. The hypothesis can admittedly explain why we cannot help but make judgments about objects of experience by bringing them under one or another of the concepts that have been “implanted” in us. But if the employment of these concepts represents nothing more than “an arbitrary subjective necessity implanted in us” for “combining representations in accordance with such a rule governing their relations,” then it is impossible for us ever to make a judgment in which we are conscious that our judgment necessarily agrees with the object of experience.<sup>26</sup> In such a case, Kant says, one can only ever say: “I am so constituted that I cannot think these representations otherwise than as thus connected.”<sup>27</sup> And this outcome, Kant writes, “is precisely what the skeptic most desires. For then all our insight through the supposed objective validity of our judgments is nothing but sheer illusion, and there would be no shortage of people who would not admit this subjective necessity (which can only be felt) in their own case.”<sup>28</sup>

What Kant is saying here is that the idea of a creator who installs in us a faculty of thought in accordance with a concept of the agreement

<sup>25</sup> *Ibid.*, B168.

<sup>26</sup> *Ibid.*

<sup>27</sup> *Ibid.*

<sup>28</sup> *Ibid.*

between its judgments and the objects of experience is incompatible with the idea of knowledge. His argument is that, according to the hypothesis, we are unable to perform acts in which we combine concepts into the kind of unity that involves a consciousness of the “objective necessity” of this combination, i.e., the sort of unity that involves consciousness of the necessity of combining these concepts to a unity that is *in the object* and not just in us.<sup>29</sup> Yet performing a judgment of the form “a is F,” Kant argues, involves doing just that: namely, combining concepts into a kind of unity that involves a consciousness of the necessity of this combination as one that resides in the object—one that is therefore and in that sense represented as in necessary agreement with the object.

On the creator hypothesis, performing a judgment of the form “a is F” would simply be impossible. For it is conceivable, on this hypothesis, that there could have been, in fact, no agreement between our faculty of thought and the objects of our experience. This rules out the possibility to combine concepts in a way that involves a consciousness of their necessary agreement with the objects we experience. Thus, if this hypothesis were true, it would be impossible to perform judgments that exhibit this form.

Kant’s refutation of the idea that a capacity for knowledge of objects of experience might be considered an “implanted subjective disposition” entails a denial of the idea that the teleological causality exhibited by a capacity for knowledge of objects of experience is of the same sort as that exhibited in an artifact. In order to get clearer about this distinction between two different species of a teleological causality that is entailed in Kant’s argument, let us compare, once more, a capacity for knowledge with a clock’s capacity to tell time. In the latter case we can understand quite well how a thing can have a capacity whose causal efficacy is determined by a concept (of its proper effect): namely, by positing a creator, distinct from the thing in question, who has the rational capacity for producing things in accordance with a concept of their effects. The idea of a creator, distinct from the thing in question, who has the rational capacity for producing things in accordance with a concept of their effects, is not only not incompatible with

<sup>29</sup> Ibid., B167.

the thing's capacity. Rather, it constitutes its very explanation. By contrast, this mode of explanation, Kant wants to say, is unavailable to us in the case of a capacity for knowledge of objects of experience. Kant's argument, as we seen above, focuses on the possibility of a judgment in which one is conscious of one's judgment as being in necessary agreement with the object of experience. Now, being conscious of one's judgment as being in necessary agreement with the object of experience means being conscious of one's judgment as non-accidentally true. Yet if knowledge consists in non-accidentally true judgment, then being conscious of one's judgment as non-accidentally true means that one is conscious of one's judgment as in agreement with the concept of knowledge. That is, the idea of judgment Kant is concerned with is the idea of judgment as a self-conscious exercise of one's capacity for knowledge.

Thus, when Kant claims that the idea of an "implanted subjective disposition" undermines the very intelligibility of an act of judgment, his argument is not based on the idea that the concept of knowledge is distinct from the concept of telling time, in terms of its content. Rather, it is distinct in form. In contrast to the concept of telling time, Kant argues, the concept of knowledge is the concept of a *self-conscious* teleological cause. What fundamentally distinguishes a capacity for knowledge from a clock's capacity to tell time is that the activities of a clock, which manifest its capacity to tell time, are logically distinct from the acts that determine those activities as manifestations of the capacity to tell time. The clock tells the time—but it does not itself make judgments about its present activities as acts of telling the time. To be sure, the clock would not have the capacity to tell the time in the first place if there had never been a judgment that asserted an agreement between its activities and the concept of telling time. But a judgment that asserts an agreement between the clock's activities and the concept of telling time is a manifestation of a capacity that is different from the one manifested in the clock's own activities. Hence, in the case of an artifact there are two distinct capacities in play. On the one hand, there is the capacity for judgments about the agreement between certain acts and the telos of a certain capacity, e.g., telling the time. On the other hand, there is another capacity, distinct from the first, which the clock manifests

when it tells the time. The clock's capacity to tell time is characterized by the fact that its manifestations are logically distinct from actualizations of the first capacity. We can express this by saying that the clock's capacity is a non-self-conscious teleological cause.

Knowledge is different from telling time, Kant argues, in that it is a self-conscious telos. That is, unlike the clock's capacity, a capacity for knowledge is not simply one whose exercises consist in acts that fall under the concept of that capacity from some perspective or other. A capacity for knowledge is one whose exercises fall under the concept of this capacity from the perspective of the subject whose capacity it is. It is a capacity whose exercises consist in an employment of the concept of that capacity by the subject who possesses it. In the case of a capacity for knowledge, an act that manifests this capacity contains a representation of that act as being in (perfect or imperfect) agreement with the concept of the capacity. The representation of an act as being in agreement with the concept of the relevant capacity and the exercise of that capacity itself are not two acts stemming from different capacities but two aspects of one and the same act.

We have thus come to the following characterization of the kind of cause that a capacity for knowledge is: *A capacity for knowledge is a cause that is dependent on the concept of knowledge in the sense that exercises of the capacity are dependent on a subject's representing her act as being in agreement (whether perfect or imperfect) with the concept of knowledge.*

Kant's argument is thus that the hypothesis that a creator implanted in us a subjective capacity for thought—a capacity constituted in such a way that its acts agree with objects of experience—can offer no answer to the question of how a capacity for knowledge, in the above sense, is possible. Rather, the creator hypothesis is incompatible with self-conscious acts of knowledge. The sort of explanation that the mooted hypothesis offers renders impossible the very thing it is trying to explain.

It is worth noting that the idea of a creator, taken in its logical form, can come under various headings, all of which would still be subject to the same Kantian argument. For example, the creator hypothesis might come in the guise of the idea of "inborn capacities," i.e., in the guise of the idea that we happen to have a capacity for knowledge of objects of

---

experience as part of our natural endowment as a matter of luck. Or it might come in the guise of the idea that we happen to have such a capacity as part of our natural endowment, not as a matter of luck but as the result of evolution, given that only those beings who happened to possess a capacity for knowledge managed to survive various ecological pressures. By now it should be obvious that, when applied to the idea of knowledge as a self-conscious teleological cause, the evolutionary story operates in the very same logical framework as the deistic story. For the evolutionary account—just like the deistic hypothesis—explains the agreement between our judgments and their objects in a way that is external to the judgments thus explained. The one hypothesis attributes such agreement to an intelligent creator, the other attributes it to an evolutionary process. By “external” I mean that these accounts conceive of the capacity they want to explain as one whose exercise does not, as such, entail a judgment about its agreement with the object to which it refers. If the agreement is thus represented as external to exercises of the capacity, however, then what is explained is not a capacity for knowledge in the sense of a capacity that is self-consciously exercised in judgments. The most that we could explain in such a manner would be a subjective disposition to combine certain representations in a certain way. This disposition might be so strong that one cannot help but actualize it. But it would not be a capacity for combining concepts into the unity of a judgment that represents that combination as grounded in the object.

Any explanation that represents a capacity for knowledge as something whose agreement with the objects is external to it, in the above sense, undermines the very capacity it seeks to explain in the attempt to explain it. We should therefore understand Kant’s discussion of the creator hypothesis as a vivid example of a much more general class of explanations that are, logically speaking, in the same boat.

#### 4. Knowledge as a Self-Constituting Capacity

We asked how we can understand the possibility of a capacity for knowledge, e.g., the capacity for knowledge of objects of experience. Our result, so far, is negative. We have established how it cannot be understood: we cannot conceive of it as a subjective disposition implanted in us “with our existence”—whether by a creator, or by evolution, or by luck—a disposition in virtue of which it just happens to be the case that our thoughts agree with the objects of our experience. How then are we to understand it?

Kant discusses the creator hypothesis as a tempting apparent “middle path” between what he describes as the “only two ways” of understanding the possibility of a necessary agreement between “experience and the concepts of its objects” we employ in judging about them, and thus to understand the possibility of knowledge. Kant writes:

Now there are only two ways in which we can conceive [denken] of a necessary agreement between experience and the concepts of its objects: either experience makes these concepts possible, or these concepts make experience possible.<sup>30</sup>

Now, it is clear that the first way of conceiving the necessary agreement between experience and the concepts of its objects—namely, by thinking of the experience of the object as making the concept of the object possible—is not available to us here. For the present issue is to understand how there can so much as be a capacity whose concepts of objects are in necessary agreement with its experiences of objects. And one cannot explain how there can be such a capacity by supposing that experiences of objects make the concepts of those objects possible. Experiences of objects that make concepts of those objects possible can explain how there can be concepts whose use in judgments would be a manifestation of a capacity for experiential knowledge. But they cannot explain how there can be a capacity for experiential knowledge in the first place—i.e., how the capacity presupposed by this explanation is itself possible.

Kant therefore concludes that a capacity for knowledge of objects of experience must be conceived of in the second of the “only two ways.”

<sup>30</sup>Ibid., B166, see also B124–125.

We have to conceive of a capacity for knowledge of objects of experience as a capacity that contains concepts of objects that, in a certain sense, bring about the objects of experiential knowledge—and hence experiential knowledge itself. Such concepts do not bring about the objects of experiential knowledge in the sense that they generate the existence of the objects, “since,” as Kant emphasizes, “representation unto itself does not produce its object with respect to its existence.”<sup>31</sup> Instead, such concepts give rise to its objects *as* objects of possible experiential knowledge. To employ such concepts, Kant argues, means to have general representations of objects that necessarily agree with the objects of experiential knowledge, for they make the objects of experiential knowledge, as such, possible. Having such general representations of objects thus means to have a kind of *knowledge* of these objects. What kind of knowledge is this? Kant calls it “a priori knowledge.” It is *a priori* knowledge in a sense that contrasts with experiential knowledge. It is a kind of knowledge that cannot be acquired through an exercise of a rational capacity for experiential knowledge, for it is knowledge that explains how a rational capacity for experiential knowledge is possible in the first place. Now, given that this kind of knowledge consists in representations that make the objects of experiential knowledge, as such, possible, it is a kind of knowledge of objects that would be manifested in any act of experiential knowledge as that which explains any such act. Because it is not knowledge of this or that particular object, but knowledge of something general that characterizes any object of experiential knowledge, as such, it is a kind of *general* knowledge of objects of experience. Yet the idea of a capacity that contains general knowledge of objects of experience that would be manifested in any act of experiential knowledge as that which explains the possibility of any such act is nothing other than the idea of a capacity for experiential knowledge that contains *a priori* knowledge of itself as a capacity for experiential knowledge.

Kant’s answer to the question of how a rational capacity for experiential knowledge is possible thus is the following: What makes a rational capacity for experiential knowledge possible is a certain kind of knowledge—namely, *a priori* knowledge of itself as a capacity for

<sup>31</sup>Ibid., B125.

experiential knowledge.<sup>32</sup> According to Kant's argument, this is the only way to explain the possibility of a rational capacity for experiential knowledge. Now, to explain a rational capacity for experiential knowledge through an act of *a priori* knowledge of the capacity itself, does not explain the capacity in question through something that is different from what it explains. The act that explains how a rational capacity for experiential knowledge is possible already entails what it explains.

We might call a capacity that is explained through an act that already entails what it explains a self-constituting capacity. It is, as Kant argues, conceptually impossible to think of a rational capacity for experiential knowledge without conceiving it as a capacity that constitutes itself by employing concepts of objects that make objects of experiential knowledge possible in the first place and, hence, that constitutes itself through *a priori* knowledge of itself as a capacity for experiential knowledge.

<sup>32</sup> It is an advantage of our account of the Kantian position that we can allow ourselves to abstract from the specific Kantian distinction between the forms of sensibility and the forms of the understanding, because it allows us to liberate our account from the problems that affect the Kantian position due to this distinction. One of the deepest problems has been pointed out by McDowell in "Hegel's Idealism as Radicalization of Kant," where he brings out how the "brute fact" character of the forms of sensibility that go with the way in which Kant treats this distinction spoils the whole Kantian ambition to have succeeded in entitling himself to a position that is compatible with "empirical realism." This problem is due to his methodological starting point, which is to give an account of sensibility's contribution to knowledge in isolation from the understanding—which, as we discover at the end of the Transcendental Deduction, proves to be impossible. If one gives up this starting point, no such problem can even arise. However, to appreciate this problem does not entail that we have to ascribe to Kant a "two-capacity conception" of knowledge, according to which knowledge is the product of two capacities whose exercises can be conceived to be logically independent of one another. Although this is a widespread reading of Kant, it fails to do justice to what Kant actually achieves. I argue against such readings of Kant in "Spontaneity and Receptivity in Kant's Theory of Knowledge." That such a reading of Kant fails to appreciate the fundamental thought at which Kant arrives by the end of the Deduction is one of the crucial points of McDowell's *Mind and World* as well as of his criticism of Sellars's reading of Kant—see, e.g., McDowell, "Sellars on Perceptual Experience," "The Logical Form of an Intuition," and "Intentionality as a Relation."