



**CON-TEXTOS KANTIANOS.**  
**International Journal of Philosophy**  
**N.º 6, Diciembre 2017, pp. 138-164**  
**ISSN: 2386-7655**  
**Doi:**

## **The Case for Absolute Spontaneity in Kant's Critique of Pure Reason**

### ***La defensa de la espontaneidad absoluta en la Crítica de la razón pura de Kant***

**ADDISON ELLIS\***

University of Illinois at Urbana-Champaign, USA

#### **Abstract**

Kant describes the understanding as a faculty of spontaneity. What this means is that our capacity to judge what is true is responsible for its own exercises, which is to say that we issue our judgments for ourselves. To issue our judgments for ourselves is to be self-conscious – i.e., conscious of the grounds upon which we judge. To grasp the spontaneity of the understanding, then, we must grasp the self-consciousness of the understanding. I argue that what Kant requires for explaining spontaneity is a conception of judgment as an intrinsic self-consciousness of the total unity of possible knowledge. This excludes what have been called ‘relative’ accounts of the spontaneity of the understanding, according to which our judgments are issued through a capacity fixed by external conditions. If so, then Kant conceives of understanding as entirely active. Or, to put it another way, he conceives of this capacity as *absolutely* spontaneous.

#### **Keywords**

Spontaneity; self-consciousness; unity; apperception; idealism

---

\* PhD Candidate in Philosophy at University of Illinois at Urbana-Champaign, USA. E-mail: [acellis2@illinois.edu](mailto:acellis2@illinois.edu)

## §1 Introduction

Kant's *Critique of Pure Reason* is famous for its many distinctions: pure/empirical, intuition/concept, sensibility/understanding, just for a few examples. But the distinction that is perhaps the most deeply pervasive throughout Kant's critical philosophy as a whole, and the one underlying all the previous distinctions, is that between spontaneity and receptivity. In the first *Critique*, the faculty of the understanding (the capacity for theoretical knowledge) is characterized as spontaneous, which is to say that it is a faculty of "bringing forth representations from itself."<sup>1</sup> In the second *Critique*, our transcendental freedom ("absolute spontaneity") is taken to follow from the moral law – a "fact of reason".<sup>2</sup> For Kant, then, acts of both theoretical knowledge (from a *Vermögen der Erkenntnisse*<sup>3</sup>) and practical knowledge (from *Wille*) are acts to be characterized as spontaneous. The understanding is a theoretical capacity, or one whose object is given to it from elsewhere. The will is a practical capacity, or one whose object is brought about from itself.

What it means for these acts to be spontaneous has been the subject of much debate.<sup>4</sup> But it is relatively common practice for commentators to refer to spontaneity as a kind of freedom. McDowell writes that spontaneity or "conceptual activity" takes place in the "realm of freedom" (1994, 5). Allison (1996, 57) interprets Kant as holding that the understanding is "absolutely spontaneous", which (Kitcher 1990) in turn takes to mean "transcendentally free." And recently, (Kohl 2015) argues that Kant's notion of the spontaneity of the understanding is a notion of freedom in judgment. It is correct, in some sense, to think of spontaneity as a kind of freedom, if by 'freedom' we mean something suitably general. But it is important to distinguish between our theoretical freedom (the freedom of the understanding) and our practical freedom (the freedom of the will). Even though the first *Critique* concerns itself with what the Dialectic calls "transcendental freedom" or "absolute spontaneity," I believe that Kant's primary interest here is in a theoretical conception of spontaneity. To see this, consider that Kant introduces the concept of spontaneity in two basic ways:

(1) Initially, as the spontaneity of the understanding, or the faculty of "bringing forth representations from itself."

(2) Later, the spontaneity of transcendental freedom, which is the ability to begin new causal chains:

An absolute causal spontaneity beginning from itself a series of appearances that runs according to natural laws, hence transcendental freedom...<sup>5</sup>

---

<sup>1</sup> KrV, A51/B75

<sup>2</sup> KpV, AA 5:30

<sup>3</sup> KrV, B137

<sup>4</sup> In recent decades, this debate has taken place among (Sellars 1970) and (Kitcher 1990), (Allison 1990), (Allison 1996), (Pippin 1987), (Valaris 2013), (Kohl 2015), (Boyle 2016).

<sup>5</sup> KrV, A446/B474

A faculty of absolutely beginning a state, and hence also a series of its consequences.<sup>6</sup>

He also calls this “freedom in the cosmological sense”:

By freedom in the cosmological sense, on the contrary, I understand the faculty of beginning a state from itself, the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature.<sup>7</sup>

The spontaneity of understanding and the spontaneity of transcendental freedom must be related in some deep way (viz., they are both capacities of bringing about something from themselves), but it is not clear from what Kant has said that they are identical. In fact, I believe they are importantly distinct. Transcendental or cosmological freedom is, in a most general way, an ability for a causal series of appearances to begin from itself, to be unconditioned. The idea of transcendental freedom, in the Dialectic, arises from the question “can there be an uncaused cause?” But notice that such a question cannot be directed at the understanding if what we have in mind, as Kant does in the Third Antinomy, is the question whether what is caused is a chain of appearances. The exercises the understanding is responsible for bringing about from itself are exercises of cognition, not appearances or objects to be known. It follows from this consideration that when Kant discusses transcendental or cosmological freedom, he is discussing a specific way in which a capacity could bring about something from itself, though he does not seem to have in mind acts of cognition. From this we can gather two things: first, that transcendental freedom and the spontaneity of the understanding are, in the first *Critique*, not identical notions; second, that both notions are nevertheless instances of a general concept of an ability to bring something about from itself.

Now we can ask whether it is possible for the capacities of will and understanding to do this, and if so, in what sense they do it. In this paper, I focus on the capacity of understanding. Thus, I am focused on the following two related issues: whether a capacity whose object must be given to it from elsewhere can nevertheless bring about its exercises from itself; and if so, in what sense.

Specifically, at A51/B75, Kant writes that the spontaneity of cognition is “the faculty for bringing forth representations itself” and identifies this with the faculty of the understanding. Call this the *spontaneity thesis*. But the thought that our understanding can bring forth its own representations is one that needs to be unpacked further, and Kant says very little else in the way of clarification. In this paper, I aim to unpack this in a way that helps us to understand just what the nature and scope of this spontaneity amounts to.

One way to begin to grasp Kant’s thought about the spontaneity of the understanding is to note that it is an expression of a kind of control that subjects have over their representations. Because we exercise some control over our representations,

---

<sup>6</sup> KrV, A445/B473

<sup>7</sup> KrV, A533/B561

according to Kant, we are in some sense epistemic agents. In fact, this is one Kantian lesson that many contemporary philosophers have taken very seriously.<sup>8</sup> I will begin from this thought and then develop from it an account of the spontaneity of the understanding.

### §1.1 Two Senses of “Mine”

I am the thinker of my thoughts. This is to say that my thoughts belong to me; they are mine. This is in one sense trivial, but in another sense non-trivial. It is non-trivial if we understand the mineness of my thoughts to be a special kind of mineness. Here is one way of seeing this special kind of mineness: to say that my thought is mine is to say that I am the source of my act of thinking. This much seems to be captured in Kant's spontaneity thesis. His claim, that is, is that the faculty of understanding brings forth its own representations; it brings them about. This is one sense in which my representations are mine, according to Kant.

But Kant's spontaneity thesis also entails a second sense of mineness: that my representations are mine in the sense that I am responsible or accountable for them. When I judge that P, I take my judgment to be valid. In taking my judgment to be valid, in turn, I am taking myself to have a right to issue the judgment. And I can only take myself to have a right to issue my judgment on the grounds that I am accountable for it. To gloss it in Brandomian-Sellarsian terms, I judge in a space of giving and asking for reasons. When I judge, I make a claim in a way such that I am prepared to answer a challenge of justification, or a question “why?”.

To say that my thought is mine in the first sense, then, is to say that I am accountable for my thought because I issued it. To say that my thought is mine in the second sense is to say that I am accountable for my thought because I am responsible for it. The second sense of mineness also comes with a deeper sense of freedom: because I am prepared to justify my claim, I am free in a way that I am not when I possess, say, a knee-jerk belief. Furthermore, I am responsible for what I judge *through* my act of judging. This is because when I judge, I do it on the basis of some reason, of which I am necessarily conscious. So, through judging, we might say, I bind myself to what I judge. I am responsible for what I judge because I am self-conscious in my act of judgment.

If it is right to attribute these thoughts to Kant, then the spontaneity of judgment is the self-consciousness of judgment. So, to understand what spontaneity is, we have to understand what self-consciousness is. I will begin by spelling out the two basic interpretations of Kant's spontaneity thesis: a strong interpretation (*absolute spontaneity*) and a weak interpretation (*relative spontaneity*). Then, I will show that Kant gives us reasons to think that self-consciousness belongs to the understanding in a way that excludes the weak interpretation. I will end by considering a way around one type of worry that might accompany this conclusion.

---

<sup>8</sup> (Boyle 2011, 2); (McDowell 1994); (McDowell 2009); (Brandom 1994)

## §1.2 The Spontaneity of the Understanding as Self-determination

By identifying the ‘spontaneity of cognition’ with the faculty of ‘understanding,’ Kant tells us that spontaneity is a capacity to bring forth representations from itself. So, according to Kant, the understanding is a capacity that can be the source of its own representations. The understanding is the birthplace of certain representations (e.g., the categories, and the judgments that make use of them).<sup>9</sup> This birth takes place through “self-activity” (*selbsttätigkeit*),<sup>10</sup> which is to say that the representations are “self-thought” (*selbstgedachte*),<sup>11</sup> which is a notion left unexplained. Despite the lack of explanation, we can sketch a plausible account of what it is for certain representations, like judgments, to be self-thought.

Kant explains that judgment is an act of determination. So, because the understanding is the source of judgment, it is thereby also the source of acts of determination. We will say that this makes the understanding a *self-determining* capacity. The notion of self-determination can be understood in two different ways, only one of which I will concern myself with here. First, we can give a negative characterization of the self-determination of judgment: my judgment that the Cubs will win the World Series is not implanted in me or put into my head, so I judge it for myself. Second, we can give a positive characterization of the self-determination of judgment: my judgment that the Cubs will win the World Series is an act from an awareness of that act’s own grounds. To get a more complete sense of this kind of self-determination, we will need to examine the concepts ‘determination’ and ‘capacity.’ First, consider the concept of determination. According to Kant, to judge that S is P is to *determine* S as P, which in turn is equivalent to predicating P of S. Kant writes:

Anything one likes can serve as a logical predicate, even the subject can be predicated of itself; for logic abstracts from every content. But the determination is a predicate, which goes beyond the concept of the subject and enlarges it. Thus it must not be included in it already.<sup>12</sup>

So, to determine something to be the case is to fix to some subject concept a predicate that goes beyond the subject concept (is not already contained in it). Even more clearly, the pre-Critical Kant writes:

To determine is to posit a predicate while excluding its opposite. That which determines a subject in respect of any of its predicates, is called the ground.<sup>13</sup>

---

<sup>9</sup> KrV, A66/B90

<sup>10</sup> KrV, B130

<sup>11</sup> KrV, B167

<sup>12</sup> KrV, A598/B626

<sup>13</sup> PND, AA 01:391

Thus, the judgment that the table is brown is a *determination* of the table as something brown, which in turn is to *predicate* brownness of the table by *excluding* from it the predicate 'not brown.'

Now consider the concept of the understanding as a *capacity*. What the capacity does is judge, or determine something to be the case.<sup>14</sup> As such, a particular act of judgment is an act of determination – it says of a thing S that it is a P by fixing to S a predicate P, excluding not-P. And exclusion is non-accidental. To exclude, Kant says, is to determine a subject on the ground of some reason. To say that acts of determination are acts of exclusion is to say that they exclude *on some ground*. When I judge that the table is brown, I do so from the presumption that what I judge is true – the table *is brown*, not *not-brown*. And when I take myself to be judging truly, I take myself to be judging on sufficient grounds.

Accordingly, the determination of a judgment is determination from a principle. Just as we would say that the capacity to hit a baseball is governed by some *standard* for what counts as hitting a baseball, we would say that the capacity to judge is governed by some standard for what counts as judgment. By 'standard' I have in mind an explanatory principle. All capacities are aimed at some telos, which governs the exercises of the capacity in the sense that it *explains* its exercises. The standard of judgment is the telos of the capacity of judgment, which tells us when we have succeeded or failed at forming a judgment.<sup>15</sup> In addition to being a standard for what counts as successful judgment, the principle of the capacity to judge is to be thought of as the *source* of judgment. Kant writes:

I would therefore call a "cognition from principles" that cognition in which I cognize the particular in the universal through concepts. Thus every syllogism is a form of derivation of a cognition from a principle. For the major premise always gives a concept such that everything subsumed under its condition can be cognized from it according to a principle.<sup>16</sup>

In every syllogism, we derive cognition – the conclusion – from a principle. To judge correctly, then, is not only for that judgment to be held to some standard, but for that standard to be at the same time the origin of the judgment. It is from a consciousness of the principle that I judge, and thereby come to have knowledge.

In addition to the thought that judgment is governed by a principle that acts as a standard and a source of judgment, we must think of this principle as a principle of unity.

---

<sup>14</sup> Kant says that all acts of the understanding can be traced back to judgments, KrV, A69/B94. He later says that the *first pure cognition* of the understanding is the principle of the original synthetic unity of apperception (self-consciousness), KrV, B137. This means that while the understanding is a capacity to determine judgments, so that its *determinate acts* are acts of judgments, it determines through a pure cognition of self-consciousness, something self-determined.

<sup>15</sup> While the analogy helps explain what I mean by 'principle', the way in which we are governed by a principle in hitting a baseball is, of course, different from the way in which we are governed by a principle in judging.

<sup>16</sup> KrV, A300

Judgment is no arbitrary collection of representations, such as [{cat}, {mat}]. Rather, judgment is a non-accidental (i.e., *necessary*) unity of representations: {the cat is on the mat}. This tells us that judgments are governed by a rule for bringing representations together as one.

The principle of unity governing acts of judgment, as the principle of a capacity, is internal to the capacity. This is because capacities are given form by their principles. We identify a range of behaviors as falling under one act – say, hitting a baseball – by identifying a principle for distinguishing between successful and unsuccessful instances of hitting a baseball. This means that the capacity for hitting a baseball is constituted by its principle – likewise for the capacity of judgment.

By spelling out the understanding as a capacity for judgment in accordance with an inner principle, we can now see how it is possible to call the understanding a capacity of *self-determination*. The act of judging, for Kant, is an act of determination governed by a principle belonging to the subject’s own capacity to judge.

### §1.3 Kant on Two Kinds of Self-Determination

Now that we understand judgment as an act of a capacity governed by a principle of unity, we can look more closely at Kant’s own understanding of the different ways in which the principle of a capacity can belong to that capacity. From the *Metaphysik Pölitz*:

But now the transcendental concept of freedom follows; this means absolute spontaneity, and is self-activity from an *inner principle* according to the power of free choice. Spontaneity is either absolute or without qualification, or qualified in some respect. –Spontaneity in some respect is when something acts spontaneously *under a condition*. So, e.g., a body which is shot off moves spontaneously, but in some respect. This spontaneity is also called automatic spontaneity, namely when a machine moves itself according to an inner principle, e.g., a watch, a turnspit. But the spontaneity is not without qualification because here the inner principle was determined by an external principle. The internal principle with the watch is the spring, with the turnspit the weight, but the external principle is the artist who determines the internal principle. The spontaneity which is without qualification is an absolute spontaneity.<sup>17</sup>

Here we get two conceptions of spontaneity: (1) self-activity from an inner principle that is determined by an external principle [qualified spontaneity], and (2) self-activity from an inner principle according to the power of free choice [absolute spontaneity]. Clearly, in this particular passage, Kant has in mind a distinction between two kinds of *practical* spontaneity. But we can just as well apply this distinction to the activity of the understanding. As we have seen, Kant describes the understanding as “self-active” in the Transcendental Deduction.<sup>18</sup> Now we must ask whether the spontaneity of the

<sup>17</sup> V-MET-L1/Pölitz, AA 28:267-268

<sup>18</sup> KrV, B130

understanding is absolute or qualified. By 'qualified spontaneity' Kant means an activity governed by an inner principle which is itself determined by an external principle. By 'absolute spontaneity' Kant means an activity governed by an inner principle which is not determined by an external principle. If the understanding is spontaneous in a qualified sense, then we might say that it is self-determined in accordance with a *given* inner principle. If the understanding is spontaneous absolutely, then it is self-determined in accordance with a *self-acquired* inner principle.<sup>19</sup>

Following the qualified/absolute distinction, and a similar distinction drawn in the second *Critique*,<sup>20</sup> interpreters have divided themselves into two camps: those who argue that Kant takes the spontaneity of the understanding to be absolute, and those who argue that Kant takes it to be qualified or "relative."<sup>21</sup>

I have introduced the idea of a spontaneous capacity of understanding through the idea of the self-determination of judgment. But, we might now step back and look at the spontaneity of the understanding in an even more general sense. In fact, this will help us to evaluate the contemporary relative/absolute debate, since it is rarely put specifically in terms of judgment, but rather, more broadly, in terms of 'synthesis.' Kant holds that the understanding is active not only in issuing judgments, but, more broadly, in acts of the imagination and in intuition itself:

But insofar as its synthesis [imagination] is still an exercise of spontaneity, which is determining and not, like sense, merely determinable, and can thus determine the form of sense *a priori* in accordance with the unity of apperception, the imagination is to this extent a faculty for determining the sensibility *a priori*...<sup>22</sup>

The imagination, as spontaneity, is an act in accordance with the unity of apperception. The unity of apperception, in turn, is the highest act of the understanding.<sup>23</sup> This unifying act is also at work in intuition, as evidenced by Kant's remark that "the supreme principle of all intuition in relation to the understanding is that all the manifold of intuition stand under conditions of the original synthetic unity of apperception."<sup>24</sup>

It follows from what Kant says that the understanding, while a capacity for judgment, is even more broadly a capacity that brings objective unity to all our cognitive activity. For our purposes, then, we can define the relevant types of spontaneity as follows:

**Relative spontaneity (RS)** = cognitive activity that is self-determined according to an externally determined (i.e., given) inner principle.

---

<sup>19</sup> I take this to be what Kant calls "original acquisition" in "On a Discovery...", TP, AA 8:221

<sup>20</sup> KpV, AA 5:97, where Kant contrasts transcendental freedom with the freedom of a turnspit.

<sup>21</sup> Those in the 'absolute' camp include (Pippin 1987) and (Allison 1990); those in the 'relative' camp include (Kitcher 1990) and (Sellars 1970).

<sup>22</sup> KrV, B151-152; see also B136.

<sup>23</sup> KrV, B134n.

<sup>24</sup> KrV, B136



**Absolute spontaneity (AS)** = cognitive activity that is self-determined according to an internally determined (i.e., self-acquired) inner principle.

Now that we have spelled out two different interpretations of Kant's spontaneity thesis, we can ask which one Kant has reason to endorse. This means that we must explain what the nature of the understanding's principle is. To this I will now turn.

## **§2 The Principle of the Original Synthetic Unity of Apperception**

We have explained the capacity of the understanding as being governed by a principle, which serves as both a standard and a source of the understanding's activities. We have also explained that the most basic activity of the understanding is the activity of determining the unity of cognition generally. But what is the principle that governs this activity? §17 of the Transcendental Deduction is entitled "The principle of the synthetic unity of apperception is the supreme principle of all use of the understanding." Here, Kant tells us that the principle of the understanding is the *original synthetic unity of apperception*:

The supreme principle of all intuition in relation to the understanding is that all the manifold of intuition stand under conditions of the original synthetic unity of apperception.<sup>25</sup>

Two questions must be answered. First, what is the original synthetic unity of apperception (OSUA)? Second, why is this "the supreme principle" of the understanding, even in its relation to intuition? I'll answer each in turn.

The understanding is a spontaneous capacity. It is the capacity to bring forth representations from itself. But Kant also writes, as we have just seen, that this is the capacity to bring all intuition under the unity of self-consciousness. That is, the principle of spontaneity is that the subject's representations are brought under the OSUA.<sup>26</sup> Kant has in mind a single capacity of understanding. He must therefore think that the two characteristic functions of this capacity are one and the same. The act of bringing all representations under a unity of apperception is the same as the capacity bringing forth representations from itself. Turning to the relation between these two ideas will shed light on what it means for the principle of the understanding to be the OSUA.

First, we must consider Kant's complicated thought that the understanding is a capacity for bringing representations under the unity of self-consciousness:

Thus all manifold of intuition has a necessary relation to the I think in the same subject in which this manifold is to be encountered. But this representation is an act of spontaneity, i.e., it cannot be regarded as belonging to sensibility. I call it the pure apperception, in order to distinguish it from the empirical one, or also the

---

<sup>25</sup> KrV, §17 of the Deduction

<sup>26</sup> KrV, B136; and at B134n. Kant writes that the faculty of self-consciousness is the understanding itself.

original apperception, since it is that self-consciousness which, because it produces the representation I think, which must be able to accompany all others and which in all consciousness is one and the same, cannot be accompanied by any further representation.<sup>27</sup>

The representation 'I think,' the absolute subject of every judgment,<sup>28</sup> is an act of spontaneity, and thus of the capacity of the understanding. When I judge that S is P, my judgment has the form "I think S is P"; and the representation 'I think' is, because spontaneous, self-thought. The only way 'I think' could be self-thought is through a consciousness of what I am thinking – that is, through *self-consciousness*.

How are we to understand this complicated thought: that it is through self-consciousness that I have the representation 'I think'? It might seem more natural to say that 'I think' *precedes* self-consciousness, which in turn would be a more complex representation, such as 'I think of myself thinking.' Kant gives us a clue to this puzzle by pointing out that the self-consciousness at issue cannot be empirical – that is, it is not a second-order observation of my first-order conscious states. This is because all thought is already accompanied by 'I think'.<sup>29</sup> Because all thought has the form 'I think,' all thought is already self-conscious. It does not become self-conscious through a second act. This is why all empirical self-consciousness presupposes yet a higher kind of self-consciousness: because self-consciousness in its most basic form is already contained in any act of judgment. So, Kant says, the 'I think' is a representation that cannot be accompanied by any further representation. This "pure self-consciousness," then, is no second-order representation of a first-order thought, since first-order thought already contains an act of (pure) self-consciousness.

So far, Kant has established that all thoughts are already purely self-conscious acts, and so all thoughts presuppose an original self-consciousness. Furthermore, this pure self-consciousness is the unity under which all representations must stand:

I also call its unity [the unity of pure apperception] the transcendental unity of self-consciousness in order to designate the possibility of *a priori* cognition from it. For the manifold representations that are given in a certain intuition would not all together be my representations if they did not all together belong to a self-consciousness; i.e., as my representations (even if I am not conscious of them as such) they must yet necessarily be in accord with the condition under which alone they can stand together in a universal self-consciousness, because otherwise they would not throughout belong to me.<sup>30</sup>

Pure self-consciousness is a condition of the possibility of all a priori knowledge, because no representation at all could belong to me as a thought without already belonging to self-

---

<sup>27</sup> KrV, B132

<sup>28</sup> Kant calls the 'I' the "absolute subject of all my possible judgments", KrV, A348

<sup>29</sup> KrV, B131-132

<sup>30</sup> KrV, B132, brackets mine

consciousness. Through intuition we are given a manifold of representations – singular, immediate representations of objects. These representations must belong, together, to a single self-conscious subject (a universal self-consciousness), because nothing at all could be thought by me independently from an awareness of myself thinking it. Thus, Kant takes himself to have established that all knowledge belongs to a self-conscious *unity* – that is, one self-consciousness.

To bring cognition under the unity of self-consciousness (under one self-consciousness) is, in turn, to recognize that all acts of cognition are acts belonging to me, the ‘I think’, the absolute subject of judgment.<sup>31</sup> If this is right, then we can see why Kant says that the understanding is both a capacity to bring forth representations from itself *and* a capacity to bring cognition under a unity of self-consciousness. For, to bring representations forth from itself is something the understanding does *through* bringing those representations under the unity of self-consciousness. All my judgments are representations that necessarily belong to a self-conscious unity. And my judgments belong to this unity by themselves being necessarily self-conscious. So, it is through self-consciousness that I form judgments.

The self-consciousness through which I judge contains, moreover, a double-faceted unity. On the one hand, when I am given an intuition I am conscious of the manifold of representations given to me as belonging to one ‘I’ – this is what Kant calls the *analytic unity* of apperception.<sup>32</sup> On the other hand, my consciousness of an identical ‘I’, to which all my representations must attach, presupposes a synthesis of representations – a *synthetic unity* of apperception:

Therefore it is only because I can combine a manifold of given representations in one consciousness that it is possible for me to represent the identity of the consciousness in these representations itself, i.e., the analytical unity of apperception is only possible under the presupposition of some synthetic one.<sup>33</sup>

The consciousness of my representations as belonging to one ‘I’, therefore, is at one and the same time a consciousness of the ‘I’ as common to all my representations *and* a consciousness of a synthesis of representations belonging to one and the same ‘I’. Since in order to be conscious of an ‘I’ common to all my representations I must be conscious of a synthetic unity of representations that has this ‘I’ in common, this synthetic unity “is the highest point to which one must affix all use of the understanding, even the whole of logic and, after it, transcendental philosophy.”<sup>34</sup>

To say that the principle of all use of the understanding is the OSUA, then, is to say that all acts of the understanding are acts of *my* synthesizing a manifold of representations through a consciousness of their *belonging to me* – that is, self-consciously.

<sup>31</sup> KrV, B134

<sup>32</sup> KrV, B132

<sup>33</sup> KrV, B133

<sup>34</sup> KrV, B134n.

So far we have characterized the OSUA as a principle of synthetic unity by saying that I synthesize *a manifold* of representations self-consciously. But it would be misleading to leave off here, for Kant does not see the representation of a synthetic unity as just a unity of any arbitrary collection of representations. For Kant, the representation of synthetic unity in apperception is a representation of a single body of possible knowledge, or a *totality*. To grasp this, we can look at two passages – one from the B Deduction, the other from the A Deduction.

First, the B Deduction passage. Kant says of the relation between the act of combining (synthesis) and unity:

Combination is the representation of the synthetic unity of the manifold. The representation of this unity cannot, therefore, arise from the combination; rather, by being added to the representation of the manifold, it first makes the concept of combination possible.<sup>35</sup>

What we take from this is the idea that in combining representations we already represent a synthetic unity (it “cannot, therefore, arise from the combination”). This is to say that the representation of a synthetic unity precedes any *non-arbitrary* act of combination. As long as it is no accident that I combine A with B, then the act of combining A with B is done from a consciousness that A and B really do belong together in one representation. To know that A and B really do belong together as one representation is to already represent a unity, prior to actually bringing them together as one.

This passage makes a point that Kant had already made, in a slightly different way, in the A Deduction:

But that empirical rule of association, which one must assume throughout if one says that everything in the series of occurrences stands under rules according to which nothing happens that is not preceded by something upon which it always follows – on what, I ask, does this, as a law of nature, rest, and how is this association even possible? The ground of the possibility of the association of the manifold, insofar as it lies in the object, is called the affinity of the manifold. I ask, therefore, how do you make the thoroughgoing affinity of the appearances (by means of which they stand under constant laws and must belong under them) comprehensible to yourselves?<sup>36</sup>

Kant calls the ground of the possibility of laws of association – laws that say “when A, represent B” – the “affinity” of the manifold of appearances (objects of possible knowledge). Through their affinity, appearances “must belong” and “stand under constant laws.” Clearly, then, the affinity of appearances at the very least entails that they belong together in a unified way. So, all objects of possible knowledge belong – prior to any act of

---

<sup>35</sup> KrV, B130-131

<sup>36</sup> KrV, A112-113

combination – to a synthetic unity. Kant goes on in the next paragraph to make this point very clearly:

On my principles it is easily comprehensible. All possible appearances belong, as representations, to the whole possible self-consciousness.....All appearances therefore stand in a thoroughgoing connection according to necessary laws, and hence in a transcendental affinity, of which the empirical affinity is the mere consequence.<sup>37</sup>

All appearances stand together under necessary laws, which is to say that they belong together in a unified body of possible knowledge. If they didn't belong together in this way, then we would only be capable of possessing a "swarm of appearances", but not experience. Knowledge, as distinct from mere perception, involves an act of determination, of taking something for true.<sup>38</sup> And in taking something for true, we must implicitly take our representations to be synthesizable, which in turn requires that we take all appearances to belong together, much like pieces of one large puzzle.

This is a complicated series of thoughts to which full justice can only be done in a much lengthier project, but we can capture Kant's basic insight with a single thesis. Call it the *internality thesis*.

*The internality thesis:* in the act of determining S as P, I am implicitly conscious of S and P as belonging to a total synthetic unity.

From this, Kant's double-faceted unity of apperception can be reconstructed. In the act of determining S as P, I am already conscious of myself as judging that S is P in the sense of being conscious of the thoroughgoing identity of the 'I' to which my representations are attached. I am simultaneously already conscious of myself as judging that S is P in the sense of representing a total synthetic unity of representations, making it possible to determine S as P.

The question we now must consider is whether the OSUA is a *given* principle or a *self-thought* principle. That is, we must ask whether RS is true or AS is true. I argue that RS fails to capture Kant's internality thesis, and thus fails to provide a plausible explanation of the capacity of the understanding. To show this, I will first examine two versions of RS that Kant rejects in the Transcendental Deduction of the *Critique of Pure Reason*. Then, I will show that the contemporary version of RS put forth most famously by Sellars and Kitcher also fails in the same way.

## §2.1 RS and the Internality Thesis

In the Transcendental Deduction, Kant rejects what he calls a 'preformationist' account of the understanding, according to which we come equipped with pre-determined

---

<sup>37</sup> KrV, A113-114

<sup>38</sup> Prol, AA 04:296

("implanted") rules for judging.<sup>39</sup> The preformationist account of the understanding would not be sufficient to account for the necessity and universality of our judgments, because it would make our judgments merely *subjectively* universal and necessary – that is, I would only be able to say that I must judge that *P* because I happen to be so constituted as to judge that *P*. But we are capable of objectively universal and necessary judgments – that is, we are capable of saying that *P* must be judged by all rational thinkers.

The empiricist's account of the rules of judgment is also insufficient, since it too must hold that if there are such rules, they are merely habits of association, and thus only subjectively necessary and universal.

According to both preformationism and empiricism, judgment is determined in accordance with a principle that is itself determined externally. By rejecting these views, we are left with the thought that judgment is self-determined – that is, determined by the understanding itself, and thus through a self-determined principle.

Those who support RS could attempt to explain this self-determination of judgment in one of several ways, but will usually hold in any case that RS is not empiricism of the sort that Kant criticizes. Both Sellars and Kitcher, who have famously defended RS, hold that they are being true to Kant's project of overcoming the limitations of Hume and the classical empiricists.<sup>40</sup> They take RS to be a more sophisticated position, according to which the activity of a distinct capacity of understanding is necessary in the determination of judgment. If RS is true, then while the principle of the understanding is not ultimately self-determined, the determination of judgment must be considered a cooperative affair of the internal activity of the understanding *alongside* an external constraint or guide. According to RS, judgment *is* in a way self-determined for just this reason. But, as Sellars himself notes, the activity of judging is not pure activity, but an active/passive hybrid of sorts:

...the spontaneity of which we are conscious is, though not *sheer* passivity, nevertheless *a* passivity in that the inner development is set in motion by a foreign cause and follows a routine. In the awareness of noumenal activities of synthesis we would encounter simply another example of a cause the causality of which is caused.<sup>41</sup>

So, in line with Kant's early definition of spontaneity, RS holds that the spontaneity of the understanding is a *spontaneitas automatica*, or one whose inner principle of determination is itself externally determined.<sup>42</sup>

Setting aside the rejected preformationist account, if we wish to understand Kant's claims about the spontaneity of the understanding, then we must take the more

---

<sup>39</sup> KrV, B167

<sup>40</sup> (Sellars 1997) famously rejects the "Myth of the Given," which might be ascribed to classical empiricism, and (Kitcher 1990) takes her account of the understanding to show how Kant improves upon Hume's psychology.

<sup>41</sup> (Sellars 1970, 23-24)

<sup>42</sup> V-MET-L1/Pölitiz, AA 28:267

sophisticated version of RS seriously. I will first look at two versions of RS that Kant already hopes to have refuted, both of which may be associated with classical empiricism: reductive causal-mechanism and natural teleology. We will see that what Kant takes to be the central flaw in both is that they rule out the internality thesis. By seeing how this is the case, we will later be able to see how it is also the case for the more sophisticated version of RS supported by Sellars and Kitcher.

### §2.1.1 Habit and Self-Consciousness

One version of RS is the view that judgment is the result of habits of association. According to this view, judgments are formed on the basis of a principle of association: when you represent A, you represent B. This is a possibility that Kant considers and rejects as inadequate for understanding the self-determination of judgment.<sup>43</sup> Kant first describes the need for a “subjective ground” or principle of association:

Since, however, if representations reproduced one another without distinction, just as they fell together, there would in turn be no determinate connection but merely unruly heaps of them, and no cognition at all would arise, their reproduction must thus have a rule in accordance with which a representation enters into combination in the imagination with one representation rather than with any others. This subjective and empirical ground of reproduction in accordance with rules is called the association of representations.<sup>44</sup>

If experience is to be more than “unruly heaps” of representations, those representations must be able to be reproduced in accordance with rules. This much was recognized by Hume, who says that such rules are acquired through habits. In turn, habits are themselves developments in accordance with some principle. Kant’s interpretation of Hume says that through habits of association, we learn to combine certain representations with some necessity.<sup>45</sup> Hume’s explication of his own view includes remarks in the *Treatise* that we learn such habits of association through the relative degrees of *enlivenment* afforded to us through particular acts of perception.<sup>46</sup> That is, according to Hume, perceptions of constancy, resemblance, and regularity are enlivening perceptions – ones that make us more likely to reproduce them in the future. Now we might ask why it is enlivening to perceive constancy, resemblance, and regularity. It is enlivening either through brute causal-mechanistic force or through agreement with the perceiving subject. First consider the latter view.

---

<sup>43</sup> At KrV, B5 and B19-20, he tells us that Hume’s associationism fails to account for the strict universality of judgments like those in mathematics; at B123n.b, he tells us that a subjective or implanted necessity (habit) would not prove the strict necessity of the relation between cause and effect; at B127-28, he similarly suggests the need to see beyond the limited perspective of Hume, i.e., beyond mere subjective necessity derived empirically through habit.

<sup>44</sup> KrV, A121

<sup>45</sup> KrV, B5

<sup>46</sup> (Hume 1978, 1, 86)

Call this account the *natural teleological* account of judgment (NT):

**NT:** The rational human, a kind of animal, develops habits of association (and thus judgments) through a recognition of what is good for it.<sup>47</sup>

Thus, the habits that account for the association of representations in a judgment would be developed through a kind of self-consciousness – a consciousness of what it is good for one to do or think. This, in turn, is a consciousness of something given to us – our animal nature. It might seem that this is enough to capture Kant's internality thesis. That is, NT claims that a kind of self-consciousness is already internal to acts of judgment.

But while Kant agrees with Hume that there is a subjective ground of the principle of association, he explains that the principle of association is incapable of accounting for objective judgments (and it is therefore insufficient for knowledge):

But now if this unity of association did not also have an objective ground, so that it would be impossible for appearances to be apprehended by the imagination otherwise than under the condition of a possible synthetic unity of this apprehension, then it would also be entirely contingent whether appearances fit into a connection of human cognitions...

...For only because I ascribe all perceptions to one consciousness (of original apperception) can I say of all perceptions that I am conscious of them. There must therefore be an objective ground, i.e., one that can be understood *a priori* to all empirical laws of the imagination, on which rests the possibility, indeed even the necessity of a law extending through all appearances, a law, namely, for regarding them throughout as data of sense that are associable in themselves and subject to universal laws of a thoroughgoing connection in reproduction. I call this objective ground of all association of appearances their affinity.<sup>48</sup>

Our ability to combine representations into an objectively valid judgment – that is, one capable of predicating something of the object (*viz.*, *what is* of it) – is necessarily governed by a principle that goes beyond the merely subjective and empirical ground of the principle of association. To be sure, we can associate representations in the Humean manner, and thereby combine them through habit in accordance with how we have learned to bring representations together in the past (thus entirely empirically, subjectively, and with only “comparative universality”<sup>49</sup>).

But this would not be judgment, and the reasoning goes something like this. It is no accident that appearances – the undetermined objects of experience – “fit into” or have a

---

<sup>47</sup> I do not intend to interpret Hume as rejecting a merely causal-mechanistic account of cognition in favor of some form of natural teleology. I merely understand NT as one manifestation of RS that can be constructed from the materials given to us in Hume's *Treatise* together with Kant's treatment of empiricism in the Transcendental Deduction. Additionally, I regard NT as a charitable reconstruction of a Humean empiricism. For more on a teleological interpretation of Hume, see (Baier 1991).

<sup>48</sup> KrV, A122

<sup>49</sup> KrV, A92 and B3-4; see also “judgments of perception” vs. “judgments of experience” in the *Prolegomena*.



place in a single unity of knowledge. If they did not fit into one unity of knowledge, then they could not all belong to one consciousness. To put it another way, if appearances did not necessarily fit into one unity of knowledge, then we would thereby confess that while some appearances may be objects of my knowledge, others are objects of my knowledge\*, where '\*' indicates a distinct unity. But we would not say that *I* know that *P* unless it were in agreement with all *my* knowledge as *one* unity. The very fact that we have this ability suggests that all representations that could be thought by me must thereby belong to one unity of consciousness. Indeed, Kant famously states:

The I think must be able to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible or else at least would be nothing for me.<sup>50</sup>

If I consciously represent an object – that is, if I think it – I necessarily represent that object as belonging to a single unity of thought. And indeed, when I think objects, I represent them as agreeing with all the other possible objects of knowledge. I could not hold that some objects that I know fit together or agree with some other objects of my knowledge, while some objects that I know *don't*, for this would entail a contradiction. This is why Kant says that in order to explain a priori knowledge – knowledge of what is necessarily and universally true – I must also think that what is known fits together with everything else that can be known, not in a hodgepodge manner, but necessarily. But if this is the case, it fits *for some reason*. Thus, we need a principle that explains this unity, and that principle is what Kant calls pure apperception.

If all we had at our disposal were a principle of association, then while some representations might belong to one unity of knowledge, others might belong to an entirely distinct unity of knowledge. This would make knowledge subjective, as “I would have as multicolored, diverse a self as I have representations of which I am conscious.”<sup>51</sup> It would make knowledge subjective, that is, because it would allow knowledge to be bound to particular subjects (or kinds of subject). I might know that *P* here and now, and I\* might know that *Q* there and then. But then what is known – what is the case – is carved up into what-is-the-case-for-I, what-is-the-case-for-I\*, and so on. This sort of view is not incoherent, but it could not account for objective a priori knowledge.

### §2.1.2 Process Accounts of Self-Consciousness, pt. 1: RS as Causal-Mechanism

We have looked at one empiricist account of self-consciousness that Kant rejects. Let us briefly turn to another kind of empiricist account of self-consciousness that Kant has equal reason for rejecting.

<sup>50</sup> KrV, B131-132

<sup>51</sup> KrV, B134

Assuming that the world behaves in accordance with causal-mechanistic laws, it is natural to look to causal-mechanism as a possible explanation of self-determination in judgment. The causal-mechanistic account, understood broadly, states that what I judge is causally necessitated by my antecedent mental states together with the laws of nature (hereafter CM).<sup>52</sup> Below I examine a problem for CM.

As some commentators have pointed out, our naïve understanding of judgment poses a problem for CM.<sup>53</sup> In order for CM to provide an account of self-determination in judgment, it must be possible for us to attribute to ourselves a causal explanation of judgment. But, in order to be able to conceive of myself as mechanically *caused* to think (even if indirectly), I must have the capacity of identifying a mechanical cause and attributing it to my capacity for judgment. That is, I must be able to think “I am caused to judge that *P*.” But in making such a causal attribution, I am already aware of myself as deciding that the causal attribution is valid. The capacity of self-consciousness is already presupposed by any attempt to provide oneself with a purely causal-mechanical explanation of how one judges. So, a causal explanation will inevitably fail to fully explain the activity of judgment.

This argument may sound suspicious. For, it may sound as though the suggestion is as follows: because we necessarily think of ourselves as judging independently of causal-perceptual impingements on the senses, we cannot be caused to judge in this manner. But this would be akin to arguing that, for example, because I necessarily see the stick in the water as bent, it cannot be the case that the stick is not actually bent. This is clearly a bad argument. I think, however, that the argument can be shown to be convincing once it is more carefully examined.

As Kant has argued, for a representation to be *my* representation is for it to be available for my use in accordance with my capacity for judging. To judge that *P* is to judge in accordance with a principle or rule for forming the judgment. But to judge in accordance with a rule, according to CM, is an act that we can understand as external to consciousness. By its own lights, this must be true for CM. For, while consciousness may accompany the causally-related mental states, the causal relations themselves do not have to be understood through any kind of consciousness.

CM may allow us to explain rule-following of a sort – for example, it may be that CM can account for the way in which a ball is governed by a mechanical rule when it breaks the glass. There is perhaps a sense in which the ball accords with, and thus “follows” a rule. But it does not allow for an explanation according to which something consciously acts *from* a rule. To allow for this possibility, CM would have to build into its account of judging some non-causal condition, a condition of self-consciousness which is not couched in causal terms.

Self-consciousness, according to CM, is therefore something that merely *accompanies* the power of mechanical causation. Kant's internality thesis, however,

---

<sup>52</sup> RS is sometimes taken to be a version of CM. See, e.g., (Allison 1996) interpretation of Sellars and Kitcher.

<sup>53</sup> (Pippin 1987, 46-47)

requires that acts of judgment are always already self-conscious acts – that is, *internally* self-conscious.

### §2.1.3 Two Kinds of Internality of Self-Consciousness

What we have seen is that a plausible account of judgment rests on a plausible account of self-consciousness in judging. Furthermore, while NT’s account of judgment does make judgment internally self-conscious *in some sense*, it is not in the way we require for rational subjects.

I think that there are two distinct kinds of internality for self-consciousness, only one of which can capture Kant’s thought. First, consider what appears to be Kant’s view, based on our discussion above:

(1) Cognitions can only be judgments if they are already self-conscious of one unity of knowledge (the understanding).

According to this view, judgment and the single unity of self-consciousness are inseparable, and self-consciousness is thereby *intrinsic* to judgment as such. Now consider the view of NT:

(2) Cognitions are judgments when they are already self-conscious of a subjective unity of what is good for me.

According to NT, judgment is intrinsically self-conscious, but only *relative to* subjects or kinds of subjects (viz., particular forms of animal life).

(1) and (2) map onto two kinds of internality:

(1) to **Objective Internality**: the principle of unity for forming judgments (self-consciousness) is internal to the subject through the concept of *what is the case* – i.e., internal to valid judgment as such, and thus objectively.<sup>54</sup>

(2) to **Subjective Internality**: the principle of unity for forming judgments (self-consciousness) is internal to the subject relative to its given nature.<sup>55</sup>

<sup>54</sup> Because I say “through the concept of what is the case” it might be asked whether the theory of absolute spontaneity I attribute to Kant is a conceptualist one. The conceptualism vs. non-conceptualism debate is a very difficult and important one, and I won’t be able to say anything convincing here. Tentatively, my reading is somewhere in between conceptualism and non-conceptualism as they are ordinarily conceived. I hold that every act of the understanding involves the “first pure cognition” of the understanding (KrV, B137) – the OSUA – through which all objective representation takes place. So, the OSUA is, in a sense, the representation through which all representation is objective for us. If this makes the OSUA conceptual, then I am a conceptualist. But, I suspect this is not the standard notion of conceptual. If so, then I very well may be considered a strange kind of non-conceptualist about the spontaneity of the understanding.

<sup>55</sup> In turn, we can say that objective and subjective internality correspond to different kinds of unity. Objectively internal self-consciousness is a seamless unity – i.e., not unified out of some antecedent elements or parts, but prior to them; subjectively internal self-consciousness is an aggregated unity – i.e., unified out of some antecedent elements or parts. This is not unlike the distinctions drawn recently by (Conant 2017) and (Boyle 2016) between additive and constitutive accounts of rationality. The additive account of rationality is an account of the unity of rationality that takes it to be aggregated out of separate elements, and the

Given Kant's explanation of the internality of self-consciousness, Kant's internality thesis requires us to think of the rational subject *not* as one with a given principle of thought – a given nature – but as judging in accordance with a self-determined principle. If it were the case that we judged in accordance with a given principle, then it would be inexplicable how we judge through a concept of *what is* rather than a concept of what is for me now, what is for me\* then, etc.

#### **§2.1.4 Process Accounts of Self-Consciousness, pt. 2: RS as Functionalism**

I have outlined Kant's arguments against two versions of empiricist explanations of self-determination in judgment. He rejects the view according to which self-consciousness (our self-conscious unity of knowledge) is merely subjectively internal to the subject. Now I will argue that even a contemporary version of RS, a functionalist account of the understanding, falls victim to the same sort of criticism.

#### **Sellars and Kitcher on RS**

The version of RS put forth by Sellars and Kitcher can be understood as a form of what contemporary philosophers of mind call 'functionalism' – the view according to which the basic elements of cognition can be understood in terms of their functional roles within a system.

In Sellars's well-known contribution to Kantian scholarship, "...This I or He or It (the thing) which thinks...", he compares the activity of synthesis to the processes of a computer. That is, we can understand judgment to be the combining of representations in accordance with rules given to the judging subject, triggered into action by an external input. According to this model of judgment, the judgment is produced by combining previous states of the system into a new one.<sup>56</sup>

Kitcher understands Kant in basically the same way, and dubs him a functionalist:

What Kant offers...is an account of judgmental states remarkably like that defended by contemporary functionalists. Functionalism is the theory that the identity conditions for mental states are given in terms of their causal connections to stimuli, responses, and other (internal) states.<sup>57</sup>

Both Sellars and Kitcher thus support a kind of functionalism about the activity of the understanding (judging), which we can call RSF:

---

constitutive account of rationality is an account of the unity of rationality that takes it to be seamless or prior to its parts.

In a previous version of this paper, I took NT (and all forms of teleological explanation) to account for a merely aggregative unity of cognition. I have since changed my mind. Teleological accounts of cognition can indeed account for seamless unity in judgment, which is why it is so important not to reduce all empiricist models of cognition to CM. In many ways, NT is an enormous improvement on CM. I hope this version of my paper exhibits this change.

<sup>56</sup> (Sellars 1970, 23)

<sup>57</sup> (Kitcher 1982, 66)

**RSF:** judgments are mental states produced out of antecedent mental states in accordance with certain rules of production.

This is a version of RS, according to our definition, because the principle of synthesis (here, rules of production) is not a self-thought principle. As noted earlier, the defenders of this view make it clear that they take themselves to be giving a plausible account of Kant's view of spontaneity, and so one that reaches beyond the forms of empiricism rejected by Kant in the Deduction. In order to accomplish this, RSF must be a better account of self-consciousness in judgment than CM and NT.

Kitcher takes RSF to be superior to the empiricist accounts of judgment because it accounts for the causal connections between our states. Indeed, for Kitcher, judgment cannot be the product of a subject who merely tends to associate representations on the basis of a habit. The empiricist is committed to such a view because the empiricist begins by looking at the impact of various impressions on a subject – i.e., with experience. But Kitcher, through Kant, begins with the conditions of the possibility of experience. Such conditions include among them a principle of causation, according to which the mind must be taken to be capable of causally organizing its perceptions in accordance with rules for forming judgments. Kitcher takes this to be Kant's way of overcoming the empiricism of Hume.<sup>58</sup>

Sellars takes RSF to be superior to the empiricist accounts of judgment because it avoids the Myth of the Given. By assigning to the mind a set of dispositions *through which* perceptual inputs are taken up and informed, we do not have to think of impressions ("experience" for the empiricist) as a kind of basic point of departure for judgment. Rather, experience is always already informed by the dispositions through which perceptual inputs are taken up. This, like Kitcher's project, helps to highlight a central insight of Kant's – namely, that it is only *through* some internal activity of the mind that we can make sense of what is given in experience.<sup>59</sup>

Despite moving past the empiricism of Hume, RSF is criticized by Henry Allison and Robert Pippin on the grounds that it fails to do justice to the epistemic role of the 'I think' in Kant's doctrine of pure apperception.<sup>60</sup> We can break their shared worry down into the following basic point: if it were the case that synthesis were merely combinations of antecedent mental states in accordance with given principles of cognition, then judgment would not necessarily be an act of what Allison calls "taking as." As Allison says:

Reducing a long and complex story to its barest essentials, to judge is just to take some intuitively given item or set thereof as a determinate something...the main point is that [in] all cases of taking as, no matter how complex, the mind must not

---

<sup>58</sup> (Kitcher 1990, 97)

<sup>59</sup> (Sellars 1997)

<sup>60</sup> Op. Cit., Pippin & Allison

only combine its representations in a single consciousness, it must also be conscious of what it is doing.<sup>61</sup>

If judgment is “taking as”, then it is explanatorily insufficient to define judgment in purely functionalist terms, because those terms leave out the necessary first-personal component of judgment.

As Pippin has put it, even if we understand judgment as a mere process of combination, it would still be necessary for me to first-personally *take* my representation to be synthetically connected to others. So, no functionalist account of judgment is complete. I am sympathetic with these concerns. But, to fully appreciate them, we need to see how they fall into the same category of anxieties that Kant had about the empiricist accounts of the understanding. Self-consciousness, according to RSF, remains *subjectively internal* to the judging subject. Now I will spell out more fully why I take this to be true.

### **The Subjective Internality of Self-Consciousness in RSF**

One thing we have seen in the explication of Kant's view is that pure self-consciousness is a condition on all a priori knowledge because it grounds necessity and universality. Nothing, that is, could be taken as *objectively* necessarily or *objectively* universally known unless it were related to pure self-consciousness, since pure self-consciousness is a consciousness of what I must think that is independent of my particular subjective constitution. And NT failed to capture this thought because it was only capable of explaining a subjective necessity and a subjective universality in judgment.

Likewise, RSF only guarantees a subjective necessity and universality in judgment. This is because it does not conceive of pure self-consciousness as an awareness of a single unity of knowledge. Instead, it conceives of pure self-consciousness as an awareness of the contentual interdependence of my actual states of judging. I will show why these are two different thoughts, and why Kant's internality thesis relies on the former.

Kitcher takes Kant to be saying that when I judge, I am aware of the fact that my representations are all synthetically connected, for this is exactly how it is possible for me to judge in the first place (i.e., by bringing the content of my representations [e.g., {red} + {ball}] together in various ways in order to produce judgments). The model Kitcher uses to explain the unity of self-consciousness is, as we have seen, a functionalist model along the following lines:

States M1 and M2 are combined in accordance with a rule to produce state M3,  
which is the synthetic product (a judgment).

On Kitcher's view, Kant needs only to be able to say that the unity of self-consciousness – the single awareness of all my judgments as belonging together – is the result of a synthetic process. But a synthetic process such as the one just described does not require an awareness of a totality. Kitcher writes:

---

<sup>61</sup> (Allison 1995, 346)

Kant's contention was that subjects accepted representations as the basis for objective judgments just in case they could be fitted into their existing beliefs in a particular way: they were consistent with (or extended or coherently revised) their beliefs about the basic constituents of reality and their causal interrelations. Bringing a representation to the "objective unity of self-consciousness" would be a matter of determining its coherence with existing beliefs along the categorial dimensions of substance and cause, that is, "according to the principle of the objective determination of all representations."<sup>62</sup>

So, according to Kitcher's functionalist interpretation of self-consciousness in judgment, the unity of self-consciousness that Kant speaks of is a unity of existing beliefs or judgments. The picture, then, is one according to which self-consciousness emerges with judgment: as I judge, I become aware of my (actual) judgments as cohering with my other (actual) judgments.

Now we can recall the motivation for Kant's principle of pure self-consciousness: that all objects of knowledge necessarily fit, like pieces of a puzzle, into one body of knowledge. To be purely self-conscious is to be aware, in all acts of judgment, that my representations belong to me and all my other possible acts of knowing. We can also recall the problem faced by NT: that their conception of self-consciousness in judgment is a merely subjective conception, or one that says what we are aware of is how I must judge *given my particular constitution* as a subject belonging to a particular form of animal life.

But this is *not* the thought that we have attributed to Kant, which is that all judgments presuppose an awareness of a single unity of knowledge. The reason why Kant takes pure self-consciousness – an awareness of a single unity of knowledge – to be the ground of objectivity in judgment is that it is an awareness of a totality. Or, in other words, the reason why Kant takes pure self-consciousness to be the ground of objectivity in judgment is that it is an awareness of my judgment as valid for any thinker, because it is a judgment of the way things are. But this requires a consciousness of a totality that precedes the combination of my actual judgments.

To see this more clearly, consider another objection to the functionalist. Melnick criticizes functionalist interpretations of Kant by suggesting that Kant could not have held that self-consciousness is the product of anything *sub-personal*.<sup>63</sup> The functionalist interpretation, he thinks, is sub-personal in two senses: first, because it holds that the unity of our various states and functions is not literally the subject itself; and second, because it holds that these states and functions do not consciously belong to the subject. While it does not seem to me that the functionalist necessarily believes the latter, the former does express a worry similar to mine. My worry might be seen as an expansion on Melnick's. I hold that even if RSF gave up on the view that pure self-consciousness were an emergent product of sub-personal processes – that is, even if it conceded that it is no product at all, but that self-

---

<sup>62</sup> (Kitcher 1999, 374)

<sup>63</sup> (Melnick 2009, 70)

consciousness is always already attached to every act of judgment – the account would still be unacceptably sub-personal in an important sense. While the functionalist might hold that self-consciousness and judgment are inextricably bound to one another, she still holds that it is through acts of synthesis that self-consciousness becomes intelligible. But, if I am right, then for Kant it is the other way around: it is through self-consciousness that acts of synthesis become intelligible.

Finally, I want to suggest that it is no accident that RSF holds the view that pure self-consciousness is merely a consciousness of the unity of my actual judgments. It is no accident because it is a process account of self-consciousness. The product of a process of combination is a unity through its parts. We understand M3 by first understanding M1 and M2, the states that produced M3. If self-consciousness is intelligible through synthesis, as it is for functionalists, then the character of self-consciousness is determined by an actual combination of states. So, it makes sense for the functionalist to hold that the self-consciousness internal to judgment is a self-consciousness *relative* to the actual judgments of an individual subject.<sup>64</sup>

### **§3 Conclusion: Avoiding Constructive Idealism**

I have attributed to Kant the thought that self-consciousness is absolutely internal to judgment. Based on this thought, we have seen that the various “relative spontaneity” approaches to explaining the understanding will fail. In particular, even though the functionalist version put forth by Sellars and Kitcher is intended to truly account for what Kant is up to in the Transcendental Deduction, it fails for the same basic reason that all empiricist accounts of judgment fail – that is, it fails because it makes self-consciousness only subjectively internal to judgment. If I am right, then Kant holds (for good reason) that the understanding is absolutely spontaneous, so governed by a self-thought internal principle.

Now, some might worry that the absolute spontaneity of the understanding would, contrary to what I have claimed, make the objectivity of knowledge impossible. After all, they will point out, if the understanding is absolutely spontaneous, then its valid exercises – objectively valid judgments – are entirely determined by itself. But, if so, then how could we avoid what Kant already takes to be a philosophically undesirable conclusion – namely, that knowledge is a subjective construction? I will briefly spell out this anxiety and then gesture at a way to relieve it.

I have discussed objectively valid judgment in quite general terms. Thus far we have understood it to consist of a necessary and universal form that excludes judgment from being a haphazard or arbitrary association of representations. This perhaps gives us a

---

<sup>64</sup> I should point out that I do not deny that our cognition *can* be merely relatively spontaneous. For example, some of our judgments are only subjectively necessary, such as those that Kant in the *Prolegomena* calls judgments of perception. The mere subjective necessity of my associating the representations [sun], [stone], and [warmth] can have a natural teleological basis or even a causal-mechanistic basis. What I deny, however, is that the paradigm case of judgment – what Kant in the *Prolegomena* calls a judgment of experience – can be relatively spontaneous. Judgments of experience, since objectively valid, are acts of rational self-determination.



first pass at what it means to distinguish between subjective and objective validity, but more needs to be said about what kind of rational norm objective validity is.

Kant interpreters generally take objective validity to be a kind of objective *purport* in judgment – that is, they take it to refer to the ability of a judgment to be successfully applied to the world. According to this interpretation, objectively valid judgments need not be true judgments. Robert Hanna draws the distinction between subjective and objective validity as follows:

**Subjective validity** = apparent meaningfulness and apparent truth for an individual rational cognizer

**Objective validity** = “empirical meaningfulness” of judgment, intuition, or concept<sup>65</sup>

In turn, Hanna argues that objective validity is a “necessary but not sufficient condition of truth...for false judgments are also objectively valid (A58/B83). In this way the objective validity of a judgment is equivalent to its *propositional truth-valuedness*, but not equivalent to its propositional truth”.

So, according to Hanna, we can understand objectively valid judgment as judgment that could indeed come to be true. This, together with the thought that objective validity is the same as empirical meaningfulness, expresses the idea that objectively valid judgment is objective purport as I have described it. Clinton Tolley also argues that objectively valid judgment is judgment with objective purport. He notes that while “Kant’s use of the term fluctuates to some degree”, nevertheless he takes false judgments to be objectively valid.<sup>66</sup> However, it is also possible to read Kant as holding that objectively valid judgment *is* true judgment. Stephen Engstrom, for instance, has argued for this view.<sup>67</sup>

If objective validity is, as Tolley believes, a merely possible relation to the object, then it will be important for me to demonstrate that this relation is self-determined. If objective validity is truth, as Engstrom believes, then it will be important for me not only to demonstrate that truth is self-determined, but to show that the self-determination of truth is not the same as a “creation” of truth. Thus, there is a lingering concern about identifying objective validity with truth. To do so would be to hold that all objectively valid judgment is true judgment. But, as a related matter, if all objectively valid judgment is true judgment, and objective validity is a thoroughly *internal* matter, then how could we avoid the conclusion that objectively valid judgment is an internal creation, or an artifact? If objectively valid judgment is a mere construction, then it entails a *constructive idealism*. I will conclude this paper by suggesting a possible way out for Kant.

On the one hand, the truth interpretation holds that objective validity in and of a judgment amounts to truth in judgment. Coming to know the truth, or becoming an objectively valid judgment, presupposes the original synthetic unity of apperception. In

<sup>65</sup> (Hanna 2017)

<sup>66</sup> (Tolley 2011, 9-11)

<sup>67</sup> (Engstrom 2017, 4, footnote 5); (Engstrom 2016, 19)

virtue of its being original, it is a fully self-determined unity. Thus, knowledge is an absolutely spontaneous inner activity.

On the other hand, the objective purport interpretation holds that objective validity in and of a judgment is the judgment's empirical meaningfulness or counterfactual relation to an object – i.e., a judgment is objectively valid when it could be either true or false or when it is actually applicable to objects in the world. Coming to be truth-valued, or empirically meaningful, or empirically applicable, also presupposes the original synthetic unity of apperception, because any actual judgment is secondary to the *unity of judgment*. Thus, knowledge is an absolutely spontaneous inner activity.

Thus far what I have argued is that on either interpretation, objectively valid judgment entails absolute spontaneity for Kant. But what about the worry that this also entails a constructive idealism? As I have argued, the principle of self-determination in judgment is for Kant the original synthetic unity of apperception. Furthermore, the original synthetic unity of apperception is absolutely internal to judgment, so self-thought. If this is correct, then the self-determination in judging is not an act external to the judgment itself – that is, to self-determine in judgment is not for the subject to act *on* judgment, but for the judgment to emerge as an activity of thought. The internality of self-determination in judgment is enough to show that to judge is not to act as a craftsman acts on his creations. Now, in addition to eliminating the craftsman metaphor, it is also important to see that in both cases, while objectively valid judgment is absolutely spontaneous inner activity, it rests on the conditions of sensibility – i.e. of being given objects through the senses. Thus, unlike the spontaneity of transcendental freedom in Kant's practical philosophy, the spontaneity of the understanding is a spontaneity in relation to something given to the thinking subject.

But I want to suggest that *resting on the conditions of sensibility* is not the same as being *determined by* sensibility. A capacity of receptivity is necessary for the absolute spontaneity of knowing through judgment. As (Engstrom 2006) argues, there is a sense in which the spontaneity of human cognition is limited: we do not cognize through intellectual intuition, as God would. So, we are *in some sense* limited because our cognition requires a capacity of receptivity. But this alone does not entail that it is the reception of objects that determines our acts of cognition.

### **Bibliography**

- Allison, Henry. 1996. *Idealism and Freedom*. Cambridge: Cambridge UP.
- . 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge UP.
- Allison, Henry. 1995. "On Naturalizing Kant's Transcendental Psychology." *Dialectica* 49 (2/4): 335-351.
- Baier, Annette. 1991. *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, MA: Harvard UP.
- Boyle, Matthew. 2016. "Additive Theories of Rationality: A Critique." *European Journal of Philosophy* 24 (3): 527-555.

- Boyle, Matthew. 2015. "Die Spontaneität des Verstandes bei Kant und einigen Neokantianern." *Deutsche Zeitschrift für Philosophie* 63 (4).
- Boyle, Matthew. 2011. "'Making up Your Mind' and the Activity of Reason." *Philosophers' Imprint* 11 (17): 1-24.
- Brandom, Robert. 1994. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge: Harvard UP.
- Conant, James. 2017. "Kant's Critique of the Layer-Cake Conception of Human Mindedness in the B Deduction." In *Kant's Critique of Pure Reason: A Critical Guide*, by James R. O'Shea, 120-139. Cambridge: Cambridge UP.
- Engstrom, Stephen. 2017. "Knowledge and its Object." In *Kant's Critique of Pure Reason: A Critical Guide*, by James R. O'Shea, 28-45. Cambridge: Cambridge UP.
- Engstrom, Stephen. 2016. "Self-Consciousness and the Unity of Knowledge." *International Yearbook of German Idealism* 25-48.
- Engstrom, Stephen. 2006. "Understanding and Sensibility." *Inquiry: An Interdisciplinary Journal of Philosophy* 49 (1): 2-25.
- Hanna, Robert. 2017. *Kant's Theory of Judgment*. <https://plato.stanford.edu/entries/kant-judgment/>.
- Hume, David. 1978. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge and P. H. Niddich. Oxford: Clarendon.
- Kant, Immanuel. 1998. *Critique of Pure Reason*. Translated by Paul Guyer and Allen W. Wood. Cambridge: Cambridge UP.
- Kitcher, Patricia. 1999. "Kant on Self-Consciousness." *The Philosophical Review* 108 (3): 345-386.
- Kitcher, Patricia. 1982. "Kant on Self-Identity." *The Philosophical Review* 91 (1): 41-72.
- . 2011. *Kant's Thinker*. New York: Oxford UP.
- . 1990. *Kant's Transcendental Psychology*. New York: Oxford UP.
- Kohl, Markus. 2015. "Kant on Freedom of Empirical Thought." *Journal of the History of Philosophy* 53 (2): 301-326.
- McDowell, John. 2009. *Having the World in View: Essays on Kant, Hegel, and Sellars*. Cambridge: Harvard UP.
- . 1994. *Mind and World*. Cambridge: Harvard UP.
- Melnick, Arthur. 2009. *Kant's Theory of the Self*. New York: Routledge.
- Pippin, Robert B. 1987. "Kant on the Spontaneity of Mind." *Canadian Journal of Philosophy* 17 (2): 449-475.
- Pippin, Robert B. 2005. "Leaving Nature Behind." In *The Persistence of Subjectivity: On the Kantian Aftermath*. Cambridge: Cambridge UP.
- Sellars, Wilfrid. 1970. "...this I or He or It (The thing) which thinks..." *Proceedings and Addresses of the American Philosophical Association* 44: 5-31.
- . 1997. *Empiricism and the Philosophy of Mind*. Cambridge: Harvard UP.
- Sellars, Wilfrid. 1974. "Metaphysics and the Concept of a Person." In *Essays in Philosophy and Its History*. Dordrecht, Holland: D. Reidel Pub.

Tolley, Clinton. 2011. "Kant on the Content of Cognition." *European Journal of Philosophy* 22 (2): 200-228.

Valaris, Markos. 2013. "Spontaneity and Cognitive Agency: A Kantian View on the Contemporary Debate." *Kant Yearbook* 5 (1).

