

# 5

## Kant, Bolzano, and the Formality of Logic

Nicholas F. Stang

### §1 Introduction

In §12 of his 1837 *magnum opus*, the *Wissenschaftslehre*, Bolzano remarks that “In the new logic textbooks one reads almost constantly that ‘in logic one must consider not the material of thought but the mere form of thought, for which reason logic deserves the title of a purely formal science’” (*WL* §12, 46).<sup>1</sup> The sentence Bolzano quotes is his own summary of others’ philosophical views; he goes on to cite Jakob, Hoffbauer, Metz, and Krug as examples of thinkers who held that logic abstracts from the matter of thought and considers only its form. Although Bolzano does not mention Kant by name here, Kant does of course hold that “pure general logic”, what Bolzano would consider logic in the traditional sense (the theory of propositions, representations, inferences, etc.), is formal. As Kant remarks in the Introduction to the 2nd edition of *Kritik der reinen Vernunft*, (pure general) logic is “justified in abstracting – is indeed obliged to abstract – from all objects of cognition and all of their differences; and in logic, therefore, the understanding has to do with nothing further than itself and its own form” (*KrV*, Bix).<sup>2</sup>

In recent work, both John MacFarlane and Sandra Lapointe have argued that this ‘formality thesis’ is original to Kant; according to them, no one in the pre-Kantian, Leibnizian logical tradition held that logic

---

<sup>1</sup> References to the *Wissenschaftslehre* (*WL*) are to Bolzano (1837); it is cited by section number and page.

<sup>2</sup> Cf. A55/B79, A56/B80, A70/B95, A131/B170. See the end of the essay for a list of abbreviations of Kant’s work, and editions and translations used. While I have mainly used the Cambridge translations, I have occasionally modified them in small ways.

is about the form of thinking.<sup>3</sup> As MacFarlane points out, the claim that logic is formal is now so widespread that it is often simply asserted without argument. So in criticizing the formality thesis in these post-Kantian figures (whom Lapointe aptly dubs ‘Kantian logicians’) Bolzano is really targeting one of Kant’s most influential ideas in the philosophy of logic.

Bolzano raises two distinct objections to the formality thesis. First, he objects that the formality thesis, as formulated at least by Kant and the Kantian logicians, assumes that logic is about thoughts [*Gedanken*],<sup>4</sup> which Bolzano takes to be, roughly, mental states with propositional content. Logic is not about thoughts so conceived, Bolzano argues throughout the *WL*, or anything else psychological, but about ‘propositions in themselves’ [*Sätze an sich*], abstract bearers of truth-values that stand in entailment relations, and ‘representations in themselves’ [*Vorstellungen an sich*], the mind-independent concepts that constitute propositions ‘*an sich*’. But this objection arises from Bolzano’s very general opposition to psychologism; it is not specifically a problem for the formality thesis. For the defender of the formality thesis could, in principle, claim that logic concerns propositions ‘*an sich*’ in respect of their form.

I want to focus, instead, on Bolzano’s second objection, which is specifically about the notion of ‘form’ in the formality thesis. Bolzano begins by pointing out that “this explanation [that logic is about the form of thought] appears to be, in part, insufficiently clear and, in part insufficiently determinate” and proceeds to explain one clear sense in which logic might be said to concern the *forms* of propositions, which Bolzano finds relatively unproblematic. Logic concerns not individual propositions, except perhaps as examples, but whole sets of propositions that share structural (formal) features. For instance, in logic we regard <All men are mortal> and <All dogs are mammals> as having the same form because they are instances of the same scheme: <All As are B>.<sup>5</sup> Two propositions *p* and *q* are said to have the same form if by consistent substitution of terms in *p* we can obtain *q*. This, Bolzano remarks, is a

---

<sup>3</sup> MacFarlane (2002) and Lapointe (2012).

<sup>4</sup> “All of the scholars who have made the claims quoted above tacitly assume that the totality of things that constitute the subject-matter of logic fall under the concept of *thought*, i.e. that all of them, whatever else they might be, must be *thoughts*” (*WL* §12, 47).

<sup>5</sup> Though Bolzano would not take this to be their form, for, on his view, the fundamental form of the proposition is <A has b>.

perfectly coherent notion of form, and in this sense logic can correctly be said to concern the *forms* of propositions.

If ‘formality’ in the formality thesis means something other than this, Bolzano claims, then Kant and the Kantian logicians have failed to explain what it is, to clearly demarcate the forms of propositions from their ‘matter’, or to explain why logic should be restricted to the formal study of propositions, in this other, specifically Kantian sense of ‘form’.<sup>6</sup>

The idea that logic is formal has proved very durable, and, furthermore, Bolzano’s ‘substitutional’ conception of the formality of logic has entered the mainstream of philosophical thought. Alfred Tarski’s definition of the formality of logic in “The Concept of Logical Consequence” (1936) in particular is similar to Bolzano’s, as Lapointe (2011) points out.<sup>7</sup> This raises an important question: was Bolzano right about Kant’s version of the formality thesis? Is there more to Kant’s formality thesis than what Bolzano’s substitutional conception of formality allows, and, if so, can this additional Kantian conception of formality be made clear (contra Bolzano) and defensible?

In this essay I am going to argue that the answer to both questions is ‘yes’. There is more to the formality of logic in Kant’s sense than Bolzano’s ‘substitutional’ understanding of formality, and this Kantian conception of formality can be rendered reasonably clear and defensible. In Section §2 I explain Bolzano’s substitutional understanding of the formality of logic. In Section §3 I argue that Kant does think that logic is formal in Bolzano’s substitutional sense. However, in Section §4 I argue that, on Kant’s view, the fact that logic is substitution-formal is a *consequence* of the fact that logic is formal in Kant’s sense of ‘formal’; it is not what the formality of logic, for Kant, consists in. The formality of logic for Kant is closely tied up with the matter/form distinction and I distinguish several different classes of logical items to which Kant applies that distinction: concepts, judgments, and inferences [*Schlüsse*] are all said to have a matter and a form. I argue that close attention to Kant’s texts shows that the formality of logic ultimately consists in the fact that logic concerns the *form* of the capacity for thought itself, not merely the *form* of its products (concepts, judgments, inferences). The form of the capacity explains the form of its products. That logic is about the

---

<sup>6</sup> Bolzano makes a similar objection to Kant’s theory of analyticity in terms of conceptual ‘containment’ – that it is a metaphor that Kant has not cashed out in literal terms. See *WL* §148, 87.

<sup>7</sup> However, others have argued that Tarski’s conception is quite different than Bolzano’s. See Siebel (1996) and (2002); and Rusnock and Burke (2010).

form of thought explains why logic is only concerned with the forms of propositions, hence, why logic is substitution-formal in Bolzano's sense. In Section §5 I explore what Kant could mean by talking of the 'form' of a faculty (a capacity) and I propose that the form of a faculty is what Aristotle would call the 'formal cause' of that faculty: it is that which makes the faculty the faculty that it is. In claiming that logic concerns the form of understanding, Kant is claiming that logic discovers laws that apply to the faculty of understanding *in virtue of being the faculty that it is*. I then give a Kantian characterization of what it is to be the capacity for thought, and why this entails that certain laws (the laws of logic) apply to the activity of that capacity. I conclude by arguing that this explains why the products of this capacity (in particular, inferences) have the features that Bolzano identifies as their formality: validity under substitution.

## §2 The substitutional account of formality: Bolzano

In *WL* §12 Bolzano explains the sense in which logic can correctly be said to be formal, which I have called the 'substitutional' conception of the formality of logic:

For these reasons logic (at least in its doctrines – it can be otherwise in the examples) never considers a fully determinate proposition [*Satz*], i.e. one in which the subject, predicate, and copula are fully specified, but, rather, a whole class [*Gattung*] of propositions, i.e. all propositions collectively, which, if some of their components are completely fixed, the remainder can be read in this or that way. [...] If one calls such classes of propositions general *forms* of propositions [...] then one can say that logic concerns only the forms of propositions, never individual propositions. If one calls that which is left indeterminate in such a class of propositions, such as **A** and **B** in the previous example, the *matter* of these propositions [...] then one can also say that logic concerns only the form, not the matter, of propositions. (*WL* §12, 48)

Although he does not directly state the definition of the form of a proposition, it is not difficult to reconstruct such a definition that would be acceptable on Bolzano's lights. I will proceed in two steps: first I will define what it is for two propositions to be formally equivalent, then I will define what the form of a proposition is. At each stage I follow Bolzano's characteristic procedure of defining the notion (being formally equivalent, having this or that logical form) relative to a designated set of terms (i, j, k...) and then defining the absolute notion.

Propositions, for Bolzano, are structured complexes of representations. To define formal equivalence of propositions we need, first, to define what is a ‘substitution’. Where  $p$  is a proposition and  $i, j, k, \dots$  is an ordered sequence of representations in  $p$ , a substitution for  $i, j, k, \dots$  in  $p$  is a function  $f$  from that sequence to a sequence, of the same length, of representations of the same syntactic type,  $i^*, j^*, k^*$ , etc. The substitution of  $f$  in  $p$  refers to the proposition that is obtained from  $p$  by replacing  $i$  with  $i^*$ ,  $j$  with  $j^*$ ,  $k$  with  $k^*$ , etc. Two propositions  $p$  and  $q$  are said to be formally equivalent with respect to representations  $i, j, k, \dots$  where  $i, j, k, \dots$  are representations in  $p$ , if and only if there is a substitution  $f$  that, when applied to  $p$ , produces  $q$ . For instance, the two propositions

- (1) <Men have ears>
- (2) <Cows have ears>

are formally identical with respect to <men> because (2) can be obtained from (1) by substituting <cows> for <men>. Proposition (1) is formally equivalent to

- (3) <Dogs have fleas>

with respect to <men> and <ears>, because (3) can be obtained from (1) by substituting concepts for concepts: <dogs> for <men> and <fleas> for <ears>. We can represent all three propositions (1)–(3) as substitution instances of the abstract schema:

- (4) <A have b>

We will say that two propositions  $p$  and  $q$  are absolutely formally equivalent in Bolzano’s sense if there is a substitution  $f$  of all of the non-logical terms in  $p$  that, when applied to  $p$ , produces  $q$ .<sup>8</sup> Propositions (1)–(3) are absolutely formally equivalent.

---

<sup>8</sup> As Bolzano himself points out later in the *WL*, this definition relies on being able to distinguish the logical terms from the non-logical terms. We need to restrict the range of terms that can be substituted to ‘non-logical’ or ‘non-formal’ features, or else we will obtain the absurd result: < $P \ \& \ Q$ > has the same form as < $P \ \rightarrow \ Q$ > because we can obtain the latter from the former by substituting  $\rightarrow$  for  $\&$ . The plausibility of this definition of logical form depends upon drawing a principled distinction between the ‘formal’ (or logical) terms in propositions and their ‘material’ (or non-logical) terms, a problem philosophers of logic have discussed down to the present day. Cf. Woods (forthcoming) for further discussion.

Having defined a notion of formal equivalence, we can now define the logical form of a proposition as the class of all propositions with which it is formally equivalent. The form of a proposition  $p$  with respect to its terms  $i, j, k, \dots$  is the class of all propositions that are formally equivalent to  $p$  with respect to  $i, j, k, \dots$ . We can adopt the following notational convention: the form of a proposition  $p$  with respect to  $i, j, k, \dots$  will be indicated by replacing  $i, j, k, \dots$  with variables. Any substitution instance of that scheme will be a proposition that is formally equivalent to  $p$  with respect to those terms. For instance, the form of (1) with respect to <men> can be represented as

(5) <A have ears>

Similarly, we can define the absolute logical form of a proposition  $p$  as the class of all propositions that are absolutely formally equivalent to  $p$ .<sup>9</sup>

When Bolzano claims in §12 of the *WL* that logic is formal he means that logic concerns itself with forms of propositions in this precise sense. Logic does not study individual propositions, but whole classes of propositions that share a logical form. Logic studies classes of propositions that share a logical form because propositions have many of their most important logical properties in virtue of their logical form. Some caution is required here. It would be misleading to say that, on Bolzano's view, *all* of logic is formal in this substitutional sense. Bolzano uses 'logic' in both a wide and a narrow sense. In the wide sense, 'logic' refers to the philosophical discipline of '*Wissenschaftslehre*' itself (*WL* §6, 22), the general theory of any scientific knowledge whatsoever. For instance, logic-as-*Wissenschaftslehre* includes in its field of study the distinction between *a priori* and *a posteriori* knowledge, and how the former is possible. Since the property of being *a priori* knowable is clearly not

---

<sup>9</sup> Trivially, absolute formal equivalence is reflexive (since every proposition is equivalent to itself under the trivial substitution that maps every representation to itself). Likewise, if there is a substitution  $f$  that takes  $p$  to  $q$ , then there is a substitution that takes  $q$  to  $p$ , so absolute formal equivalence is symmetric. Finally, if there is a substitution that takes  $p$  to  $q$ , and one that takes  $q$  to  $r$ , there is a substitution that takes  $p$  to  $r$ , so absolute formal equivalence is transitive. Consequently, the absolute logical form of a proposition is its equivalence class under the relation of absolute formal equivalence. It follows that two propositions  $p$  and  $q$  are absolutely formally equivalent if and only if they have the same absolute logical form. The logical form of propositions (1)–(3) can be represented as (4).

shared by propositions in virtue of their common logical form, Bolzano cannot mean that logic in the broader sense is formal. However, Bolzano also has a narrower conception of logic on which logic is the science of the relations between propositions, representations, inferences, etc., the domain of logic as traditionally conceived. On this narrower conception, 'logic' is roughly equivalent to what Kant would call 'pure general logic', although of course they have quite different conceptions of what that logic is and how it should be done. Henceforth, when I talk about 'logic' in Bolzano I mean logic in the narrow sense.

We can now formulate the Bolzano substitutional formality thesis (SFT) more precisely:

- (SFT) Property F is a proper object of study for logic if and only if propositions have property F in virtue of their logical form.

One corollary of this claim is the following:

- (C) Property F is a proper object of study for logic if and only if, for any propositions  $p$  and  $q$  that share the same form (they are formally equivalent to one another),  $p$  has property F if and only if  $q$  has property F.

I will illustrate this point in the case of Bolzano's theory of logical '*Ableitbarkeit*' because this corresponds most closely to what has come to be seen as a (if not the) central concept in logic, logical consequence, and will facilitate the comparison with Kant's theory of logical inference.<sup>10</sup>

Since Bolzano's own definition of logical *Ableitbarkeit* is formulated in his own, not always very natural, technical notation, I will quote, with my own slight modifications, Sandra Lapointe's gloss on this definition in *Bolzano's Theoretical Philosophy* (2011):

Let  $P = \{p_1, p_2, p_3, \dots\}$  and  $Q = \{q_1, q_2, q_3, \dots\}$  be sets of propositions. The  $Q$  propositions are logically *ableitbar* from the  $P$  propositions if and only if:

- (i) There is a substitution for the non-logical terms in  $P$  and  $Q$  that makes all of the propositions in  $P$  and all the propositions in  $Q$  true ( $P$  and  $Q$  are said to be 'compatible').

---

<sup>10</sup> This discussion of *Ableitbarkeit* is heavily indebted to the discussion in Lapointe (2011), Chapter 6.

- (ii) For every substitution of the non-logical terms in  $P$  and  $Q$ , if all of the propositions in  $P$  are true under this substitution, then all of the propositions in  $Q$  are true under this substitution. (Lapointe 2011, 73–74; *WL* §155)<sup>11</sup>

Since I will only be discussing Bolzano's notion of 'logical *Ableitbarkeit*' (rather than the more general notion), I will drop the modifier and refer to this as *Ableitbarkeit simpliciter*. Let  $P$  and  $Q$  be as in the definition and let  $P^* = \{p^*_1, p^*_2, \dots\}$  be a set of propositions, where each  $p_n$  is absolutely formally equivalent to  $p^*_n$ . A short argument proves that  $Q$  is also logically *ableitbar* from  $P^*$ . By clause (i) of the definition of logical *Ableitbarkeit*, there is at least one way of substituting for the non-logical terminology in the  $P$  propositions and the  $Q$  propositions that make all the propositions in both sets come out true. But for every substitution of the non-logical vocabulary in  $p_n$  there is a corresponding substitution of the non-logical vocabulary in  $p^*_n$  such that these substitutions will produce the same proposition; they will not in general be the same *substitution*, for  $p_n$  and  $p^*_n$  will not have the same non-logical vocabulary. So there is at least one substitution of the non-logical terminology in the  $P^*$  propositions and the  $Q$  propositions that make all the propositions in both sets come out true. In Bolzano's terminology, these two sets of propositions are compatible. This satisfies clause (i) of the definition of the logical *Ableitbarkeit* of  $Q$  from  $P^*$ . Next, we will prove that clause (ii) obtains as well (for  $P^*$  and  $Q$ ). Each  $p_n$  is absolutely formally equivalent to the corresponding  $p^*_n$ . By definition, this means that for every substitution of the non-logical terms in  $p^*_n$ , there is a corresponding substitution of non-logical terms in  $p_n$  such that these two substitutions will produce the same proposition. So for any substitution of non-logical terms that makes all of  $P^*$  true, there is a corresponding substitution of non-logical terms that makes all of  $P$  true. By assumption, for any substitution of non-logical terms in  $P$  and  $Q$ , if all of the  $P$  propositions are made true, all of the  $Q$  propositions are made true. So, for any substitution of the non-logical vocabulary in  $P^*$  and  $Q$ , if all of the  $P^*$  propositions are true, all of the  $Q$  propositions are true. So the set of propositions  $Q$  is *ableitbar* from the set of propositions  $P^*$ . Similar reasoning will show that if  $P$  is *ableitbar* from  $Q$  then  $P^*$  is *ableitbar* from

---

<sup>11</sup> I have modified Lapointe's presentation slightly; rather than first define *Ableitbarkeit* relative to a set of terms, then define the absolute notion of *Ableitbarkeit*, I have defined the absolute notion of *Ableitbarkeit* from the beginning.



Q as well. In other words, relations of *Ableitbarkeit* are shared by sets of propositions that have the same logical form. Or, to put it another way, relations of *Ableitbarkeit* are formal features of sets of propositions. Given the substitution-formality thesis (SFT) from above, this explains why *Ableitbarkeit* is a proper object of study for logic.

### §3 The substitutional account of formality: Kant

In this section I will argue that, on this point, Kant agrees with Bolzano: Kantian pure general logic is formal in Bolzano's substitutional sense. Kant accepts the SFT from above. This is not the only coherent sense in which logic can be said to be formal, according to Kant, but my argument for that must wait until the next section. Since I have focused on the formality of *Ableitbarkeit* in Bolzano's logic, I will focus on the closest corresponding notion in Kant's logic, his theory of logical inferences [*Schlüsse*]. First, though, we must explore Kant's conception of the logical form of the judgments that figure in such inferences.

Kantian judgments are unified wholes composed of concepts. The structure or unifying principle of a judgment is called its logical form. The table of logical functions in judging characterizes the logical forms of judgments according to four 'moments': quantity, quality, relation, and modality. Under each 'moment' of the table of logical functions there are three types of judgment:

	<u>1. Quantity</u>	
	Universal	
	Particular	
	Singular	
<u>2. Quality</u>		<u>3. Relation</u>
Affirmative		Categorical
Negative		Hypothetical
Infinite		Disjunctive
	<u>4. Modality</u>	
	Problematic	
	Assertoric	
	Apodictic <sup>12</sup>	

Every judgment has a quantity, a quality, a relation, and a modality. This means that every judgment is either universal, particular, or

---

<sup>12</sup> *KrV* A70/B95.

singular in quantity; affirmative, negative, or infinite in quality, etc. For reasons I do not have the space to explore here, Kant claims that within pure general logic we cannot distinguish between the third judgmental form and the first two in the moments of quantity and quality.<sup>13</sup> This means that, for the purposes of pure general logic, the Kantian logic with which we are concerned here, we can ignore singular and infinite judgments. Likewise, the modality of a judgment, according to Kant, does not concern the form of the judgment itself, but the relation of that judgment to the mind;<sup>14</sup> consequently, the modality of judgments does not affect their role in syllogistic inference, and I will forgo further discussion of it here.

The logical form of a judgment is uniquely characterized by its logical quality, quantity, relation, and modality. Two judgments have the same form if and only if, through consistent substitution of concepts, one can be transformed into the other. For instance, consider the two judgments:

- (1) <Some Greeks are Athenian>
- (2) <Some dogs are Alsatians>

Through substituting <dogs> for <Greeks> and <Alsatian> for <Athenian>, (1) can be transformed into (2); (1) and (2) have the same logical form: they are particular (in quantity), affirmative (in quality) categorical (in relation) judgments (ignoring for the moment their modality). If we refer to the concepts of which these judgments are composed as their matter, then we can say: judgments with the same form can differ in matter, or, equivalently, a judgment is a complex whole of some matter structured according to one of the possible logical forms.

Kant claims that in the moments of quantity and quality the third type of judgment cannot be distinguished from the first type by pure general logic (for Bolzano, 'logic' in the narrower sense), but only by 'transcendental logic', the philosophical discipline that "concerns the origin of our cognitions of objects insofar as that [cognition] cannot be ascribed to the objects" (A55–56/B80). In this essay, I am concerned exclusively with Kant's pure general logic, so I will ignore the positive claim that transcendental logic can distinguish the third logical

---

<sup>13</sup> *KrV* A71-72/B96-97. See Stang (2012) for further discussion of why infinite judgment cannot be distinguished within pure general logic.

<sup>14</sup> *KrV* A74/B99-100.

function of quantity and of quality and the difficulties surrounding the notion of ‘transcendental logic’<sup>15</sup> but I will briefly comment on what it means that, within pure general logic, we cannot distinguish universal judgments from singular (quantity), or affirmative judgments from infinite (quality). Very roughly, Kant’s idea is that *with respect to inferential relations* singular judgments behave just like universal judgments (the moment of quantity) and infinite judgments behave just like affirmative judgments (the moment of quality). So from within pure general logic, the science tasked with studying the inferential relations among judgments, we cannot distinguish these two kinds of judgments. I will illustrate the point in the case of affirmative judgments and infinite judgments;<sup>16</sup> the case of singular and universal judgments brings in additional complications.<sup>17</sup>

Consider two kinds of judgments that can be distinguished within general logic, affirmative and negative judgments, for example.

- (1) Socrates is Athenian. [Affirmative]
- (2) Socrates is not Athenian. [Negative]

In Kant’s logic, these judgments have, respectively, the logical forms

- (1\*) A is B
- (2\*)  $\sim(A \text{ is } B)$

The negation in (2\*) takes wide scope (over the whole judgment) because it applies to the copula;<sup>18</sup> (2\*) says that the predicative relationship asserted in (1\*) fails to obtain. However, we might also negate, not the copula as in (1) itself, but the predicate, as in:

- (3) Socrates is non-Athenian [Infinite]

whose corresponding logical form is (3\*):

- (3\*) A is ( $\sim B$ )

Judgment (3) asserts that Socrates falls within the infinite sphere of everything outside the sphere of the concept <Athenian>. That pure

---

<sup>15</sup> See, however, Stang (2012) for an explanation of why infinite judgments are distinct from affirmative judgments in transcendental logic.

<sup>16</sup> The main discussion of this in the *KrV* is A72/B97.

<sup>17</sup> See A71/B96.

<sup>18</sup> Cf. Ak. 24:758.

general logic cannot distinguish (3\*) means that, from the point of view of pure general logic, (3) is an affirmative judgment; it affirmatively predicates the concept <non-Athenian> of Socrates. Consequently, it has the inferential role, with respect to other judgments containing the concepts <Socrates> and <non-Athenian> that any affirmative judgment does. For instance, the following is a valid instance of the schema for categorical syllogisms:

Socrates is non-Athenian (3) [X is A]  
 All non-Athenians are warlike. [All A are B]  
 ∴ Socrates is warlike. [X is B]

However, (3) has a different inferential role from (1), because (1) and (3) are affirmative judgments with distinct and *unrelated* (from the point of view of general logic) predicates. Nor is the inferential role of (3) the same as the inferential role of (2), for (at least within pure general logic), substituting (2) for (3) in the syllogism above renders it invalid. The difference between (1) and (3) concerns not the form of the judgment *itself* but the form of the predicate, <~A> as opposed to <A>. Pure general logic cannot account for the difference between these predicates; to do so, we need to bring in considerations of how our *a priori* concepts obtain content (relation to an object), and this is a subject for transcendental logic.<sup>19</sup>

Kant's theory of inference, following the Aristotelian tradition, is a theory of syllogistic inference. A syllogism contains two premises (a major and a minor) and a conclusion. He divides syllogisms into three kinds, based on the logical form of their major premise: categorical, hypothetical, and disjunctive. I will begin with the best developed part of Kant's theory of inference, the theory of categorical syllogisms, which is encoded in this table:<sup>20</sup>

MP	PM	MP	PM
SM	SM	MS	MS
SP	SP	SP	SP

<sup>19</sup> This is Kant's 'official' view about infinite judgments in pure general logic. However, there are reasons to worry that he is not completely consistent on this point. See below for more.

<sup>20</sup> *JL* §68, Ak. 9:125.

'S' stands for the subject-concept of the conclusion, 'P' for its predicate, and 'M' for its middle term. Each column represents a categorical syllogistic 'figure', that is, a kind of a logically valid categorical syllogism. In each figure, the top line represents the major premise, the middle the minor premise, and the line on the bottom, below the black line, represents the conclusion. In each judgment, the left-hand variable indicates the subject-concept, and the right-hand variable indicates the predicate concept. For instance, the third syllogistic figure, the third column from the left, indicates that syllogisms of this figure have the following form:

	<i>Subject</i>	<i>Predicate</i>
<i>Major premise:</i>	M	P
<i>Minor premise:</i>	M	S
<i>Conclusion:</i>	S	P

The logical quantity and quality of the premises and conclusions are not specified, though, because each syllogistic figure includes under it several sub-cases that fully specify the logical forms of its constituent judgments.

Kant analyses the validity of syllogisms of the other three figures by showing how they can be reduced to syllogisms of the first-figure by certain conversion rules. First, I will show that the validity of a categorical syllogism of the first-figure is substitution-formal in the sense defined earlier. Then I will show that the conversion rules that transform syllogisms of the other three figures are themselves substitution-formal. I will then argue that the validity of *all* categorical syllogisms is substitution-formal. In fact, it is a trivial (though laborious) exercise to enumerate every possible example of a valid categorical syllogism and manually check that they are all substitution-formal.

A syllogism is of the first-figure just in case (i) the concepts in its major, minor and conclusion can be obtained by substitution from the scheme in the above table, (ii) the major premise is universal (either negative or affirmative), and (iii) the minor premise is affirmative (either universal or particular). This gives us four possible sub-cases of the first-figure:

	<u>I.1</u>	<u>I.2</u>
<i>Major:</i>	All M are P	All M are P
<i>Minor:</i>	All S are M	Some S are M
<i>Conclusion:</i>	All S are P	Some S are P

	<u>I.3 (invalid)</u>	<u>I.4 (invalid)</u>
<i>Major:</i>	Not all M are P	Not all M are P
<i>Minor:</i>	All S are M	Some S are M
<i>Conclusion:</i>	Not all S are P	Not all S are P

Since I.3 and I.4 are clearly invalid, there are only two valid syllogistic forms of the first-figure: I.1 and I.2.

When we get to the other figures, things get slightly more complicated. First, we have to introduce a set of conversion rules, rules that allow us to convert premises in syllogisms into forms that allow the application of a syllogistic form of the first-figure. These conversion rules are:

1. No X is Y    ↔    No Y is X
2. Some X is Y    ↔    Some Y is X
3. All X are Y    ↔    Some X are Y (↔ Some Y are X)<sup>21</sup>

Kant describes rule (3) as the rule that universal affirmative judgments can be converted (predicate and subject switched) *per accidens*, which means that they can be converted *only* in one direction, to the corresponding logically weaker particular affirmative judgment. Rule (3) can be thought of as the logical product of the traditional Aristotelian rule of sub-alternation<sup>22</sup> and rule (2), which converts particular affirmative judgments. The rule of sub-alternation has the effect of ruling out vacuous universal affirmative judgments. If the concept <A> has no objects in its extension, then no true affirmative judgments can be made with it as the subject-concept.

A syllogism is valid if and only if it can be transformed, via the conversion rules, into a categorical syllogism of one of the two kinds listed above (I.1 & I.2). Since the two kinds of valid first-figure categorical syllogisms are substitution-formal (I take it that this is obvious by inspection), and the conversion rules are substitution-formal (ditto), it follows that whether a syllogism is valid is substitution-formal. Although, I take it, the premises of this argument are obvious by inspection, it may be less obvious that the conclusion follows. Let S and S\* be two syllogisms which are formally equivalent in the following sense: by consistent substitution of concepts, the premises and conclusion of S\* can be transformed into those of S. Note, though, that logical forms of judgments

<sup>21</sup> *JL* §53, Ak. 9:119.

<sup>22</sup> Kant discusses sub-alternation at *JL* §46, Ak. 9:116.

are not concepts; otherwise, we could get the absurd result that <All A are B> has the same logical form as <Some A are B>. Assume that S is a valid syllogism. Then S is either already a syllogism of the first-figure, or it can be transformed into one by a series of application of conversion rules. If S is a first-figure syllogism, then so is S\*, so S\* is valid. If there is a series of applications of conversion rules that transforms S into a valid first-figure syllogism, then, the same series of applications of conversion rules will transform S\* into a valid first-figure syllogism. So S\* is a valid syllogism. I conclude that the property of being a valid categorical syllogism is substitution-formal.

It is even clearer that the property of being a valid hypothetical or disjunctive syllogism is substitution-formal, for Kant's theory of those syllogistic forms is simpler than in the case of the categorical. All hypothetical syllogisms come in one of two forms:

	<u>H.1</u>	<u>H.2</u>
<i>Major</i>	If $p$ then $q$	If $p$ then $q$
<i>Minor</i>	$p$	$\sim q$
<i>Conclusion</i>	$q$	$\sim p$

where  $p$  and  $q$  are any judgment whatsoever and ' $\sim p$ ' denotes the negative judgment in which the *copula* in  $p$  (not its predicate, as in infinite judgments – see above) is negated. Disjunctive syllogisms have one of two forms:<sup>23</sup>

	<u>D.1</u>	<u>D.2</u>
<i>Major</i>	A is $(B_1 \vee B_2 \vee \dots \vee B_n)$	A is $(B_1 \vee B_2 \vee \dots \vee B_n)$
<i>Minor</i>	A is $B_j$	$\sim(A \text{ is } (B_1 \vee B_2 \vee \dots \vee B_{j-1} \vee B_{j+1} \vee \dots \vee B_n))$
<i>Conclusion</i>	$\sim(A \text{ is } (B_1 \vee B_2 \vee \dots \vee B_{j-1} \vee B_{j+1} \vee \dots \vee B_n))$	A is $B_j$

As is clear from these syllogistic forms, Kant thinks of disjunctive judgments in terms of what we would now call the 'exclusive or'. Since disjunctive judgments can be either universal or particular in quantity, we should distinguish two different forms of the major premise, and thus multiple different kinds of disjunctive syllogism. However, for the sake of simplicity, I will just assume that the major premise is universal

---

<sup>23</sup> *JL* §77.

and the minor is particular, and the conclusion retains the quantity of the major premise:

	<u>D.1U</u>	<u>D.2U</u>
<i>Major</i>	All As are $(B_1 \vee B_2 \vee \dots \vee B_n)$	All As are $(B_1 \vee B_2 \vee \dots \vee B_n)$
<i>Minor</i>	Some A is $B_j$	$\sim(\text{Some A is } (B_1 \vee B_2 \vee \dots \vee B_{j-1} \vee B_{j+1} \vee \dots \vee B_n))$
<i>Conclusion</i>	$\sim(\text{All A is } (B_1 \vee B_2 \vee \dots \vee B_{j-1} \vee B_{j+1} \vee \dots \vee B_n))$	All As are is $B_j$ <sup>24</sup>

Once again, by inspection, it is clear that the property of being a valid hypothetical syllogism and the property of being a valid disjunctive syllogism are substitution-formal.

We could conclude, on Kant's behalf, that the property of being a valid syllogism is substitution-formal because being a valid syllogism means being either a valid categorical, hypothetical, or disjunctive syllogism. However, this would be too quick for two reasons. Although Kant does not make this explicit, he should also count as mixed inferences of reason those syllogisms in which one premise immediately entails a premise that, together with the other premise, entails the conclusion by one of the syllogistic figures. For instance, consider this syllogism:

All M are P  
 If (some M are P) then (some R are Q)  
 $\therefore$  Some R are Q

Intuitively, this is a valid syllogism because the first premise immediately entails (according to rule 3 above) that some M are P, which, together with the second, constitutes a valid hypothetical syllogism. However, the original syllogism is not a valid syllogism of any of the officially recognized forms. So I propose the following expansion of our notion of a valid syllogism:

A pair of premises and a conclusion ( $P_1, P_2, C$ ) are a valid syllogism if and only if  $P_1, P_2,$  and  $C$  instantiate one of the valid syllogistic forms, or if  $P_1$  or  $P_2$  immediately entail one or more judgments that, together

---

<sup>24</sup> This is how I interpret Kant's remark that "we infer either (1.) from the truth of one member of the disjunction to the falsehood of the others, or (2.) from the falsehood of all members but one to the truth of this one" (*JL* §77, Ak. 9:130).



with  $C$ , instantiate such a form, or if  $P_1, P_2$  and  $C^*$  instantiate such a figure for some judgment  $C^*$ , where  $C^*$  immediately entails  $C$ .

But even taking this amendment on board, Kant has only given an account of entailment that holds between either a single premise and single conclusion (the immediate conversion rules from earlier, or “immediate inferences of reason”), or between a pair of premises and a single conclusion (syllogisms, “mediate inferences of reason”). For Kant’s logic to be adequate, it would need to account for general entailment relations among arbitrary (perhaps arbitrarily large finite) sets of premises and conclusions. Here is a sketch of how that might go:

*Defn.* The deductive closure of  $P_1 \dots P_n$  is defined iteratively in stages:

- (i) At stage 0 we add  $P_1 \dots P_n$ .
- (ii) At stage  $n + 1$  we add all immediate consequences of every judgment added at stage  $n$  or earlier, and all judgments  $C$  such that judgments  $A$  and  $B$  (where  $A$  and  $B$  are added in stage  $n$  or earlier) and  $(A, B, C)$  instantiate a valid syllogistic form.

*Defn.*  $P_1 \dots P_n$  entail  $C_1 \dots C_n$  if and only if the deductive closure of  $P_1$  through  $P_n$  includes each of  $C_1 \dots C_n$ .

Therefore, despite the serious defects in Kant’s logic, his account of entailment is formal in precisely Bolzano’s sense: it is substitutional. Or, to put it as Bolzano does in §12 of *WL*, the part of pure general logic that studies inferences of reason does not study individual ordered triples of judgments (propositions), but sets of such ordered triples: sets of ordered triples of judgments that share a logical form, meaning they can be obtained through consistent substitution of concepts from some abstract schema. This does not show that *all* of Kantian pure general logic is substitutional; it shows, however, that a significant portion of that logic, the theory of inference, is substitutional, and this is the part of Kant’s logic that most closely corresponds to Bolzano’s theory of *Ableitbarkeit*, which, in the previous section, we saw to be substitution-formal as well.

#### **§4 Form of the capacity versus form of the product of the capacity**

Now I am going to argue that, while Kant does think that logic, paradigmatically the part of pure general logic that studies inference, is substitution-formal, there is a more basic sense in which logic is formal, and that more basic sense explains why logic is substitution-formal.

Understanding the precise meaning of Kant's formality thesis can seem a daunting task, given the sheer variety of ways in which Kant characterizes the formality of logic: logic is said to concern the form of thought, the form of reason, the form of the understanding, the formal laws of the understanding, as well as the more familiar claims that it concerns the form of concepts, the form of judgments, or the form of inferences. The different forms that are relevant to the formality of logic can be roughly separated into two classes: forms of faculties (e.g. the form of understanding) and forms of products or activations of those faculties (e.g. the form of concepts, judgments, and inferences).

As numerous scholars have pointed out, Kant applies the matter/form distinction across a wide swath of his philosophy; concepts, experience, the will, etc. are all said to have a form and a matter.<sup>25</sup> In some contexts, by distinguishing between the form and the matter of some item, Kant is claiming that the item in question (e.g. a judgment) is a complex structured whole and distinguishing within that whole between its parts (its matter) and the structure that obtains among those parts (its form), in virtue of which those parts compose that very item. To apply the matter/form distinction, in these paradigm cases, requires distinguishing a class of items called *matter* from a class of structural relations called *forms* such that, when instances of the appropriate kind of matter are structured by an appropriate kind of form, a complex entity of the relevant kind exists. For instance, to talk of the matter/form distinction with respect to judgments requires distinguishing *concepts*, the parts out of which judgments are made, from *logical forms*, the structures that when applied to these concepts produce complex wholes, *judgments*. A consequence of the matter/form distinction, when applied in paradigm cases like these, is that the very same matter could be structured by a different form, producing a different complex whole. For instance, the concepts <Greek> and <Athenian> could be structured in the judgment <All Athenians are Greek> or in the distinct judgment <Some Greeks are Athenian>. It is also a consequence of the matter/form distinction, when applied in paradigm cases like this, that the same form can produce distinct complex wholes by structuring different matter. For instance, the logical form of universal affirmative judgment structures the distinct judgments <All Athenians are Greek> and <All squares are equiangular>, because these judgments are made of different concepts (their matter is different). In these paradigm cases of the matter/form distinction, the

---

<sup>25</sup> My thinking about the matter/form distinction in Kant is indebted to an unpublished paper by Matt Boyle, "Kant's Hylomorphism".

distinction is ultimately between the parts (matter) of some complex whole, and the structure (form) of that complex whole. I will call these cases the ‘standard hylomorphic cases’, because the distinction between matter (*hyle*) and form (*morphe*) can be applied here in a reasonably straightforward way. I take it that Kant’s application of the matter/form distinction to judgments and inferences are standard hylomorphic cases. I have already explained this in the case of judgments. The distinction between the matter and form of syllogisms follows directly from the theory of syllogistic inference in the previous section: the form of the syllogism is the syllogistic figure, while the matter is the particular judgments (which themselves have a further form).<sup>26</sup> Kant also applies the matter/form distinction to concepts, but I will not explore that here, because it involves the complex doctrine that the form of concepts is “universality”.<sup>27</sup>

One thing that leaps out about Kant’s ubiquitous use of the matter/form distinction is that some of the things to which he applies the distinction are not standard hylomorphic cases. Some of things that are said to have a form and a matter are not very naturally thought of as complex structured wholes with parts. Very roughly, we can divide the cases into two kinds (with citations to relevant texts in parentheses):

<u>Standard hylomorphic cases</u>	<u>Non-standard hylomorphic cases</u>
Judgments (Ak. 9:101)	Understanding (Bx)
Inferences (Ak. 9:121)	Reason (Bx)
Intuition (A23/B37)	Sensibility (A20/B34)
Maxims (Ak. 5:28)	Will (Ak. 4:436)
Concepts (?) (Ak. 9:91)	Thought (A59/B84)

From this table we can see that the items that are not naturally thought of as complex structured wholes (the non-standard hylomorphic cases) are faculties or capacities, while the items that are standard hylomorphic cases are products of such faculties. For instance, a judgment, according to Kant, is an activation of the faculty of the understanding. A judgment

---

<sup>26</sup> In *JL* §59 Kant claims that the form of a syllogism is the conclusion “insofar as it contains the *consequentia*” (Ak. 9:121). I take this to mean: the conclusion qua consequence of those premises. But if the conclusion is a valid consequence then it is a consequence in virtue of it and the two premises instantiating a syllogistic figure. So the point in the main text remains: the form is the syllogistic figure.

<sup>27</sup> *JL*, Ak. 9: 91.

is a complex structured whole, but it is very unnatural to think of the understanding *itself* as having parts and a structure. Concepts are an intermediate case because, while they are products of a capacity, they are not *standard* hylomorphic cases since, while they might be thought to consist in parts (marks) organized according to a form, Kant refers to the form of concepts, not as the logically complex manner in which they are arranged, but as their *universality*; whatever the ‘universality’ of concepts is, I take it, it is not a matter of their parts being related in some way, for even simple concepts are general and thus possess the form of universality.

Applying the hylomorphic analysis to faculties immediately leads to difficulties. What are the parts of the understanding? Are they other faculties? If so, why is understanding one of the basic faculties (rather than the faculties out of which it is composed)? If the parts of the understanding are not faculties, what could they be? I am going to take this as sufficient reason to at least *entertain the hypothesis* that when Kant talks about the *form* of the understanding (or of any other faculty) he does not mean the structure that obtains among the parts of the understanding (or of that other faculty). What, then, does he mean?

Before we address that question, though, we must first note some variation in Kant’s claims about which faculty’s *form* logic is about; although he typically describes logic as concerning the form of *thinking* (e.g. Bxxiii, A21/B36, A55/B79, etc.), he sometimes describes it as concerned with the form of *understanding* (e.g. Bix, A59/B84), sometimes with the form of reason (e.g. Bxi), and sometimes with the form of understanding *and* reason (e.g. A53/B77, A796/B824). One thing to note is that *thinking* is not a faculty in its own right, according to Kant. Kant’s lists of faculties vary somewhat from text to text, but in the first *Kritik* he seems to admit only four basic faculties: sensibility, understanding, (theoretical) reason, and the faculty of desire, practical reason (which, only briefly discussed in the first *Kritik*,<sup>28</sup> forms the central topic of the second). In the third *Kritik* he supplements this with a fifth faculty: the power of (reflecting) judgment [*Urtheilskraft*].<sup>29</sup>

---

<sup>28</sup> The main discussion of practical reason in *KrV* is in the Canon der reinen Vernunft, A797/B825–A820/B848.

<sup>29</sup> In the A Deduction Kant appears to claim that imagination is one of “three original sources which contain the conditions of the possibility of experience and cannot themselves be derived from any other faculty of the mind” (A94; cf. A115). However, already in the A Deduction (A119) and even more so in the B Deduction (B153), the synthesizing role of the imagination seems to be attributed to the understanding itself.

Thinking is not a separate faculty; it is a very general description that applies to the activities of both the understanding and reason because they are discursive, or conceptual.<sup>30</sup> Thinking is discursive cognition in general, a point Kant makes in passing: “since merely formal logic, so conceived, abstracts from all content of cognition (whether it be pure or empirical), and concerns itself with the form of thinking (of discursive cognition) in general” (A131/B170).<sup>31</sup> That understanding and reason are discursive/conceptual (and hence instances of thinking) means that activations of these faculties constitutively involve concepts: concepts themselves (both empirical and *a priori*), judgments (in the case of the understanding) and inferences (in the case of reason). In claiming that logic concerns the form of *thinking* Kant is claiming that it concerns the form of discursive representation in general, of which understanding and reason are instances. In order to mark this distinction between thinking and the particular discursive faculties (understanding, reason) I will refer to thinking as a ‘capacity’: we possess the capacity to think in virtue of possessing the more determinate faculties of understanding and reason.

To return to our earlier question, one possibility would be to interpret Kant’s talk of the form of a faculty as shorthand for talk of the form of the products of its activity (which *are* standard hylomorphic cases). So when Kant claims that logic concerns the form of thinking, what he means is that logic deals with concepts, judgments, and inferences (the products of the discursive faculties, understanding and reason) solely in virtue of their forms, that is, solely with features of, and relations among, these items that they possess in virtue of having the structure they do, and which they would share with any item that shared such a structure. On this reading, the formality of logic, for Kant, is very close to the formality of logic for Bolzano: logic does not concern itself with individual judgments or arguments, but with whole classes of

---

<sup>30</sup> I take it that the ‘power of judgment’ [*Urtheilskraft*] is also conceptual (though not wholly conceptual — see Ak. 5:214–217), in the minimal sense that its activities (reflecting judgments) are expressed with concepts, and hence is a faculty for thought normatively subject to logical laws (e.g. a logically contradictory aesthetic reflecting judgment is impossible). However, since by 1787 Kant had not even formulated the project of a *Kritik* of the faculty of reflecting judgment, and whether logic applies to reflecting judgment is not thematized in the *Kritik der Urtheilskraft*, this would be hard to substantiate from the text of either the first or the third *Kritik*.

<sup>31</sup> See also B93 and A230/B283.

structurally (formally) indistinguishable judgments or arguments. There are passages that suggest this reading, for instance:

General logic abstracts, as we have seen, from all content of cognition, i.e. from any relation of it to the object, and considers only the logical form in the relation of cognitions to one another, i.e. the form of thinking in general. (*KrV*, A55/B79)

One reading of this passage is that Kant is claiming that logic (a) abstracts from, or ignores, the matter (constituent concepts) of judgments, what their objects are, whether they even have objects, etc.; (b) considers only the relations that judgments have to one another when this content is abstracted from; and (c) he identifies the formality of logic with (b). This suggests that by 'formal' here Kant merely means that logic is concerned with relations (e.g. inferential relations) that hold between judgments in virtue of their logical form ("the logical form in the relation of cognitions to one another"). This 'reductive' account of the form of thinking (i.e. reducing it to the form of its product) might be further encouraged by Kant's claim that "we can trace all acts of the understanding back to judgments, so that the understanding in general can be represented as a faculty [*Vermögen*] for judging" (A69/B94). Kant seems to be claiming that we can understand what the understanding is by understanding what it does: judge. Consequently, it might be argued, talk of the form of the understanding is shorthand for talk of the form of its most basic activity, judgment. Likewise, on this interpretation, talk of the form of thought is shorthand for talk of the form of its most basic activities/products: concepts and judgments (understanding) and inferences (reason).

I think we should be sceptical of this 'reductive' reading of Kant's talk of the 'form of thinking' because it leaves unexplained why Kant does not *only* talk about logic as being about forms of judgments, inferences, etc. (or about judgments, inferences, etc. in virtue of their form) but consistently and repeatedly claims that logic is about the form of thinking itself. If Kant only meant to convey that logic is about the forms of the products of this capacity why does he consistently and repeatedly claim that logic concerns the form of the capacity itself? In general, Kant does not restrict the hylomorphic analysis to the products of faculties, but consistently and repeatedly speaks of the forms of various faculties (see table above); if he meant the latter only to be a shorthand for the former, this is at least a very confusing way for him to express his point. This is a *prima facie* reason to look for another explanation of Kant's talk of the form of a faculty.

In several places in his lectures on logic Kant employs a more general notion of form:

In every cognition we must distinguish matter, i.e. the object, from form, i.e. the manner in which we cognize the object. (Ak. 24:510)

In logic one abstracts from objects and regards only the form of the understanding, i.e. *modus cognoscendi*. (Ak. 24:616)

The second quote is from the *Busolt Logik*; a few pages later Kant distinguishes between matter and form and adds this parenthetical qualification: “form (*modus cognoscendi*)” (Ak. 24:621). I call this a more ‘general’ notion of form because it encompasses the more strictly hylomorphic notion of ‘form’: the *modus cognoscendi* of, e.g., a judgment is its logical form (in the hylomorphic sense) because the manner in which judgments cognize objects (form as *modus cognoscendi*) is by combining concepts according to that (hylomorphic) form. So in these texts at least Kant identifies form *in general* with manner (or mode) of cognition. This is not, I think, a deviant or minor usage on Kant’s part. For even in the *Jäsche Logik* Kant refers to the *form* of concepts as universality (Ak. 9:91), and to an improvement in the distinctness of a cognition as a change in its logical *form* (Ak. 9:64).<sup>32</sup> As I mentioned earlier, the *form* of concepts (universality) cannot naturally be thought of in terms of the hylomorphic notion of form: universality is not a structural relation among the parts of a concept (its marks) because those are concepts and thus possess universality in their own right. Likewise, Kant describes the distinctness of a cognition, the degree to which its parts are consciously apprehended by the subject, as a difference in the form of the cognition. This cannot be understood on the hylomorphic analysis of form because *by definition* the distinct and the indistinct cognition have the same parts organized in the same manner; the difference does not consist in the parts, or their manner of arrangement, but in the subject’s ability to consciously apprehend and differentiate them.

If we take this on board as Kant’s most general notion of form then we can say: logic concerns the *form* of thinking because it concerns the manner of cognition involved in thought. This can sound extremely vague, since thinking can be done in many ways (many ‘modes’): judging, inferring, etc. The *modus cognoscendi* of thinking is most naturally understood as the mode of cognition common to thought as such, or, what all

---

<sup>32</sup> Cf. Ak. 24:418, 512, 538, and 28:229.

thinking has in common in virtue of being thinking. If it referred merely to the accidental mode of our thinking (how we happen to think) then the study of the form (*modus cognoscendi*) of thought would be what Kant calls “applied logic” (a kind of empirical psychology), rather than “pure general logic”, the science of the absolutely necessary rules of thinking as such. So logic concerns the form of thought in this sense: it concerns the mode of cognition all thinking has in virtue of being thinking. This is exactly what Kant says about logic elsewhere:

[pure general logic] contains the absolutely necessary rules of thinking, without which no use of the understanding takes place. (*KrV*, A52/B76)

All rules according to which the understanding operates are either *necessary* or *contingent*. The former are those without which no use of the understanding would be possible at all, the latter those without which a certain determinate use of the understanding would not occur. (*JL*, Ak. 9:12)

So the claim that logic concerns the form of thinking means that logic concerns the manner of cognition of thinking as such, which means it concerns the rules which thinking necessarily obeys, the rules without which thinking is not thinking. This, I take it, is what Kant means when he describes the laws of logic as the “essential” laws of the understanding (Ak. 16:44, 24:526), or as the understanding’s “own laws” (24:824). Logic does not concern the *modus cognoscendi* thinking merely happens to have; it characterizes the laws that characterize the manner of cognition thinking has in virtue of being the capacity (of discursive representation) that it is. Since understanding and reason are essentially discursive capacities, logic also studies laws that describe the manner in which these faculties cognize their objects in virtue of being the faculties they are.

Earlier, we saw that Kant claims that logical laws are necessary for thought. In the *Jäsche Logik* he makes clear that this necessity is normative not descriptive: “in logic, the question is not about *contingent* but about *necessary* rules; not how we do think, but how we ought to think” (*JL*, 9:14).<sup>33</sup> It is not that we necessarily *do* obey logical laws, but that we ought to. So if the form of thinking is the *modus cognoscendi* of thinking

---

<sup>33</sup> Cf. *Pöhlitz Logik* 24:503; *Dohna-Wundlacken Logik* 24:693; *Wiener Logik* 24:792; Refl. 1627.



itself, and this consists in laws that are normative for thought itself, then one thing the formality of logic might mean is: logic studies the normative laws that apply to thinking in virtue of being the capacity that it is. This, I take it, is what Kant means by claiming that “logic is to teach us the correct use of the understanding, i.e. that in which it agrees with itself” (*JL* 9: 14), a claim echoed in his reference in the *Kritik* to “general logical rules for the agreement of cognition with itself” (B116). Logic brings the understanding (or more generally, thinking) into agreement with itself because it teaches the normative laws that apply to that capacity in virtue of being the capacity that it is.

Since ‘law’ carries the normative connotation of guiding one’s action (or, in this case, one’s thought) when I talk about ‘laws’ I mean the normativity of logic; e.g. the logical *law* of non-contradiction is that one ought not to *judge* that  $\langle A \text{ is } \sim A \rangle$  is true. When I talk about logical ‘rules’ I will mean the descriptive logical claims (e.g. that no judgment of the form  $\langle A \text{ is } \sim A \rangle$  is true) that are transformed into normative claims in logical laws. Kant himself draws a similar distinction between laws and rules:<sup>34</sup> rules become laws when they are represented as necessary. Although Kant has descriptive natural laws in mind, we could adapt this to the logical context: logical *laws* are logical rules thought as normatively necessary.

I have outlined an alternative reading of one thing that Kant means by claiming that logic studies the form of thinking: it studies the *modus cognoscendi* of thinking as such and thus discovers normative laws that apply to thinking in virtue of being thinking. I have also discussed the ‘reductive’ interpretation of Kant’s talk of the form of thought, according to which Kant’s claim that logic studies the form of thinking means that logic studies the products of the capacity for thought – concepts, judgments, inferences – in virtue of their form. It is a short step to understand the formality of logic’s study of these products in terms of substitution-formality: for instance, logic only studies properties of inferences that are preserved under substitution of non-logical concepts. This reductive reading would make Kant’s claim that logic is formal very close to Bolzano’s claim that logic is substitutional. In the remainder of this section, I will explain why I think the reductive reading is untenable.

---

<sup>34</sup> See A126, Ak. 5:184, and Refl. 5414. Cf. Kant’s claim in the *Grundlegung* that “only law brings with it the concept of an *unconditional* and objective and hence universally valid *necessity*” (Ak. 4:416). Kant has practical laws in mind here, but I think the same could be said of logical laws.

Consider the following passage, one of the longest sustained discussions of the formality of logic in the *Jäsche Logik*:

If we now put aside all cognition that we have to borrow from *objects* and merely reflect on the use of the understanding as such, we discover those of its rules that are necessary without qualification, for every purpose and without regard to any particular objects of thought, because without them we would not think at all. Thus we can have insight into these rules *a priori*, i.e., *independent of all experience*, because they contain merely the conditions for the use of the understanding in general, *without distinction among its objects*, be that use *pure* or *empirical*. And from this follows at the same time that the universal and necessary rules of thinking concern merely its form and never its matter. Therefore the science that contains these necessary and universal rules of thinking in general is merely a science of the form of our cognition through the understanding, or of thought. (*JL*, 9:12)<sup>35</sup>

Kant begins by claiming that if we ignore the particular concepts involved in judgments we can discover the rules that necessarily apply to any thinking, no matter about what object or with what kinds of concepts. These rules are normative rules, logical laws. But Kant goes on to claim that we can have *a priori* insight into these normative rules. ‘Insight’ is a technical term for Kant. It means knowledge from an explanatory ground; insight requires not merely knowledge that something is the case, but knowledge *why* it is the case.<sup>36</sup> So Kant is claiming that because the laws of logic are normative for thought as such we can have *a priori* insight into them, i.e. we can know *why* those normative rules of thought apply to all thinking as such.

He then concludes from this that these rules are *formal*. Now the reductive reading faces a dilemma: what does Kant mean when he claims that “the universal and necessary rules for thinking concern merely its form”? *If* he means that logic is substitution-formal, or that it concerns only properties of judgments, inferences, etc. that are preserved under substitution of non-logical concepts (matter), then in this passage Kant is

<sup>35</sup> For a very different interpretation of this passage, see MacFarlane (2002), 46.

<sup>36</sup> See *Jäsche Logik* (Ak. 9:65), *Pölitiz Logik* (Ak. 24:539), *Dohna-Wundlacken Logik* (Ak. 24:730), Refl. 1866 (Ak. 16:141), Refl. 1955 (Ak. 16:169) and Refl. 2394 (Ak. 16:342), *Kritik der praktischen Vernunft* (Ak. 5:27, 46, 47), and *Kritik der Urtheilskraft* (Ak. 5:83).

*inferring* the formality of logic (in this sense) from the *normativity* (necessity) of logic for all thinking as such. But we have already seen that Kant sometimes uses ‘form’ in a very general sense to refer to the *modus cognoscendi* of some faculty, so this claim (the normativity of logical laws for thought as such) is a kind of formality claim (for Kant): logic studies the *modus cognoscendi* of thinking as such and thus its essential normative laws. On this reading, Kant is inferring substitution-formality from its normative-formality. On the other hand, *if* the reductive reader of Kant interprets ‘formal’ in the second half of this passage as a restatement of logic’s normativity for thinking as such (normative-formality) then they have to admit that by the formality of logic Kant sometimes means what I take it to mean: it studies normative laws essential to thinking as such. I read the passage in the first way: Kant begins by asserting that logic studies the normative laws that apply to thinking as such (form in the sense of *modus cognoscendi*) and concludes from this that it studies the products of the capacity for thinking by abstracting from their differences in concepts and objects. Substitution-formality follows from normative-formality.

This passage from the *Pölitiz Logik* poses a similar interpretive dilemma:

The rules without which one would not think at all are necessary. The contingent rules that depend upon a particular object of explanation, are as numerous as the objects. Because the understanding, which acts according to rules, is present in every thinking there must be rules are common to all thinking, regardless of the object, and which lie a the basis of every use of the understanding, without which it would not be possible; and these are necessary. They therefore contain the form of thinking. (Ak. 24:502)

Kant here infers from the fact that logic is normative (necessary) for all thinking, regardless of its object, that logic is formal. Does formal mean substitution-formal here? If so, Kant is inferring substitution-formality from normative-formality, which is precisely my thesis. If not, then Kant is not drawing a conclusion from the normative-formality in the final line but restating that thesis in new terms: it is formal (in the normative sense). This, again, confirms my claim that the formality of logic, in its most fundamental sense, means that logic concerns the form of thinking, that is, the laws that apply to thinking in virtue of being the capacity that it is, its “essential laws”.

Readers who favour the reductive interpretation might want to object at this point: granted that Kant sometimes infers the (substitution)

formality of logic from its normativity for thought as such, why should we think that the latter (normativity) claim is ever what he means by claiming that logic studies the form of thinking? Two points in reply. First, as we have seen, Kant repeatedly refers to the forms of faculties, but this notion of form is not well captured in hylomorphic terms (because faculties, standardly conceived, do not have parts) and he does have a more general notion of form: *modus cognoscendi*. Thus, my reading gives us an explanation of why Kant talks of logic as studying the form of a faculty not just the forms of its products. Second, Kant sometimes talks about the formality of logic in a way that naturally invites my reading, but which is at best awkwardly interpreted on reductive lights: “logic is thus a self-cognition of the understanding and of reason, not in regards to their object, however, but merely in regard to form” (*JL*, Ak. 9:14). By now, it should be clear how I interpret this claim; logic is genuine self-knowledge on the part of the understanding because logical knowledge is knowledge of what norms the activity of the understanding is bound by in virtue of being the faculty it is (a discursive faculty). The reductive reader is forced to read it as follows: logic is self-knowledge of the understanding and of reason in that it is knowledge of normative laws that apply to the products of these faculties in virtue of the forms of those products. My objection is simply: this does not look like self-knowledge on the part of understanding or of reason, but of knowledge of their products. This is an instance of a wider problem with the reductive reading: where Kant talks about (the form of) a faculty, the reductive reading has to interpret him as meaning the (form of) the products of that faculty.

The connection to Aristotle, from whom the matter/form distinction ultimately derives, can help us understand why Kant might have talked about the ‘form’ of thinking in this way. While Aristotle introduces matter and form in the *Physics* in such a way that the notion of form (*eidos*) seems closely tied to the shape (*morphe*) or structure of some complex whole,<sup>37</sup> in the *Metaphysics* it is a wider notion. In its most general use, the form of a thing is its essence, the answer to the question, *what is it?*<sup>38</sup> Consequently, things without matter can nonetheless have *forms* (immaterial substances). In other words, the form of a thing makes it the thing it is. So, going back to Kant, if logic concerns the

---

<sup>37</sup> *Physics* 190b15, 191a10, and 193a30. Cf. *On generation and corruption* 335b6 where he seems to use *morphe* and *eidos* interchangeably.

<sup>38</sup> *Metaphysics* 1032b1–2, 1035b33–1036a2.

form of thinking, then logic concerns *what thinking is*, or, to put the point more straightforwardly, whatever rules logic uncovers are rules that are grounded in *what it is to think*. In other words, the rules of logic apply to thinking in virtue of *being the capacity it is*. They are grounded in the essence (form) of this capacity.<sup>39</sup> To bring this back to the point I put aside above, when Kant claims that logic “teaches us the correct use of the understanding, i.e., that in which it agrees with itself” I take this to mean: the rules of logic describe how the understanding operates when it is operating correctly as the faculty that it is. This also explains how we can have *a priori* insight into the rules of logic. *A priori* insight into the rules of logic is not merely knowledge that they are valid, but knowledge of *why*. We can know *why* the rules of logic are valid by understanding what capacity the capacity for thinking is and why, when it operates correctly as the capacity that it is, it obeys these rules.

The conclusion of this line of thought is that to understand what the formality of logic means for Kant, we must understand (1) what the capacity for thought is, and (2) what, if any, normative laws apply to it in virtue of what it is. If those normative laws are the laws of logic, then we will have made significant progress in understanding what it could mean that logic concerns the form of thinking.

## §5 Logic and the unity of apperception

Kant’s claims that logic studies the form of thought, the form of the understanding, and the form of reason means: each of these is a faculty of discursive representation and logical norms apply to discursive representation in virtue of its form (in virtue of being the faculty it is). Why does it lie in the nature of discursive faculties that their activities are normatively subject to the laws of logic? We might expect Kant’s answer to have this structure: an explanation of why logical norms apply to discursive representation as such, which would entail that logical norms

---

<sup>39</sup> I am not claiming that ‘form’ generally means essence, for Kant; as we have seen, a cognition changes its *form* as it becomes more distinctly understood by a subject, but it makes no sense to say its essence has changed. The most general notion of form in logical contexts is that of *modus cognoscendi*; since thinking is a capacity, a capacity to cognize (represent) in a certain way, its *modus cognoscendi* characterizes what it is to *be* that capacity, its essence. This is not true in general of things that had a *modus cognoscendi* (e.g. a concept had by a particular subject, which can have varying degrees of distinctness).

apply to the activity of the understanding. However, this is not what Kant does. He explains why logical norms apply to the activity of the understanding. He leaves it up to us to generalize from this to an explanation of why logical norms apply to activities of discursive representation in general. So I will first give a reconstruction of Kant's explanation of why logical norms apply to the activity of the understanding (in virtue of what that faculty is), and then I will explain how this generalizes to discursive representation in general (thinking).

Some readers might be surprised at my claim that Kant "explains why logical norms apply to the activity of the understanding". Where, they will ask, does Kant do anything of the sort? At three places in the Transcendental Deduction (two of which occur in footnotes!) he claims that the unity of apperception explains the possibility of logic. In each case, Kant's remarks are brief and cryptic, so it is up to us to reconstruct how they might constitute an explanation:

[...] the logical form of all cognition necessarily rests on the relationship to apperception as a faculty. (*KrV*, A117n)

And thus the synthetic unity of apperception is the highest point to which one must attach all use of the understanding, even the whole of logic and, after it, transcendental philosophy; indeed this faculty is the understanding itself. (*KrV*, B133–134n)

The category presupposes combination. We must therefore seek this unity... someplace higher, namely in that which itself contains the ground of the unity of different concepts in judgments, and, with it, the possibility of the understanding, even in its logical use. (*KrV*, B131; underlining in all three quotes by NS)

I have argued that logic, for Kant, is a normative science. In these passages, Kant claims that the unity of apperception explains how this normative science is possible. I take this to mean two things: (i) the unity of apperception explains how it is possible to cognize logical laws; and (ii) the unity of apperception explains how it is possible that (normative) logical laws apply to us. If (ii) were false, the unity of apperception would not explain the possibility of logic: it would fail to explain why the claims of logic (claims that certain norms apply to our thinking) are true. If (i) were false, logical norms might apply to us, but we could not cognize this (nor could we cognize why they apply). I take this to entail that logic, the normative science, would not be possible. However, I will focus on (ii); the Kantian account of how logical cognition (cognition

of logical laws as normatively binding on our thought) is possible lies outside the scope of this essay.

Kant does not tell us explicitly how the unity of apperception does (i) or (ii). He leaves it up to us to reconstruct his reasoning. We need to explain why logical norms apply to us, e.g. why I stand under the norm that if I judge that  $p$ , that  $p \rightarrow q$ , and that  $\sim q$  then I ought to revise or reject at least one of my judgments. But now ask yourself, why am I bound by that norm? Why am I normatively required not to have inconsistent judgments? Consider Kant's remark at B64 that:

For although a cognition may be in complete accord with logical form, i.e., not contradict itself, yet it can still always contradict the object. The merely logical criterion of truth, namely the agreement of a cognition with the general and formal laws of understanding and reason, is therefore certainly the *conditio sine qua non*, thus the negative condition of all truth; further, however, logic cannot go, and logic cannot discover by any touchstone the error that concerns not form but content. (*KrV*, A59/B84)

If Kant is claiming that logic concerns the negative principles about truth (i.e. no thought that violates logical laws can be true), this does not get us any closer to understanding why logic is normative for us. If logic describes a set of principles that my thoughts cannot violate and still be true, we do not get an explanation of why logic is normative for us *unless* we assume that truth (or at least non-falsehood) is the end of our thinking. This opens the space for a reconstruction of Kant's reasoning that the unity of apperception explains the possibility of logic, where logic studies the norms that apply to the understanding in virtue of the faculty the understanding is (logical laws). The explanation is: the faculty of understanding, because of the faculty it is, aims at non-falsehood. To put it another way, among the ends of the faculty of the understanding, one of them is non-falsehood. Logic describes the negative requirements that must not be violated if our thinking is going to satisfy one of its constitutive ends, non-falsehood.

Before continuing, I want to briefly explain what I mean by claiming that thought as such aims at non-falsehood rather than truth. In the *Kritik* Kant defines truth as "the agreement of cognition with its object" (A58/B82).<sup>40</sup> The definite description "its object" presupposes that the

---

<sup>40</sup> This definition of truth is echoed throughout the logic lectures: Ak. 24:391, 525, 718, and 823, as well as Refl. 2162 and 2177.

cognition has an object. But not all thoughts are cognitions; some thoughts do not have objects. These thoughts are not *false*. They do not fail to agree with their object; they simply lack objects and are thus not even *apt* to be true or false (unless an object is provided for them).<sup>41</sup> This is confirmed by Kant's corresponding account of falsity in the *Philippi Logik*:

Falsity is either: *materialis* or *formalis*. Formal falsity is when a cognition contradicts itself, or does not agree with itself [*sich nicht paßt*]. (Ak. 24:391)

'Material' falsity I take to refer to a cognition that does not agree with its object. But then what are we to make of a consistent (not formally false) thought that does not even have an object? It is not materially false, because it does not even have an object. Nor is it true, for the reasons given above.

The conditions under which an object can be given for a thought are specified in transcendental logic;<sup>42</sup> in particular, the thought must involve concepts of objects in space and time, because these are the forms in which objects are given to us.<sup>43</sup> Consequently, if thought as such aimed at *truth* then the logic that studies the form of thought (and thus uncovers why logical norms apply to thought) would have to be transcendental logic.<sup>44</sup> But it is not; it is pure general logic. Logic is concerned with the laws that no thought can violate if it is going to be non-false. Obeying the laws of logic is obviously not sufficient to make a thought *true* (for it might still lack an object),<sup>45</sup> but it isn't

---

<sup>41</sup> By denying the "truth-aptness" of objectless thoughts I am not talking about the contemporary notion of "truth-aptness" (that is involved in, for instance, debates about the truth-aptness of moral judgments). As I am using the term, it merely refers to the fact that such thoughts fail to satisfy the presupposition of Kant's definition of truth (agreement with their object).

<sup>42</sup> A55/B80.

<sup>43</sup> In my reading, giving an object for thought (e.g. giving an object for a concept) is making that object available for thought in a way that (among other things) makes it possible to cognize that there is such an object. So it is possible that there *are* objects of some of our concepts but which cannot be *given* to us (we cannot intuit them), so we do not cognize these objects. For instance, there is a God (as Kant argues in the second *Kritik*), so the concept <God> has an object, although that object cannot be given to us (intuited). I do not have the space here to defend this interpretation.

<sup>44</sup> A62/B87.

<sup>45</sup> Kant makes this point at A59/B84 and again and again in the logic lectures.



even sufficient to make a thought truth-apt; a thought might obey all of the logical norms and still be objectless. For instance, the principle of contradiction (<No A is  $\sim$ A>) is the highest principle of logic, but not every substitution instance of it is *true*; if A is an objectless concept, <No A is  $\sim$ A> is neither true nor false. But no substitution instance of this principle is *false*.<sup>46</sup>

Nor is it the case that all discursive faculties constitutively aim at truth (where truth is understood as the agreement of a cognition with its object). The faculty of reason, as we have seen, is a discursive faculty, but reason's constitutive end is not *truth* but to find a condition for any given conditioned object and, in pursuit of this end, the faculty of reason forms the concept of an *unconditioned* condition.<sup>47</sup> Since there are a variety of different relations of condition to conditioned (substance–accident, cause–effect, and totality–limitation), this generates several different representations of unconditioned objects: the unconditioned thinking substance (the object of rational psychology), the unconditioned cause of effects in space and time (the object of rational cosmology), and the unconditioned ground of all possibility (the object of rational theology).<sup>48</sup> But none of these representations can be given an object in intuition, the only way objects can be given to us. Consequently, these concepts cannot be involved in cognition. The constitutive activity of reason – the activity in which reason manifests itself as the capacity it is – is not about any object that can be given to us. So not only does reason *not* constitutively aim at truth, it constitutively aims at representations which are not even truth-apt, because they violate the conditions under which any representation can be a cognition (a representation to which an object is given) and thus be either a true or false representation of its object.<sup>49</sup>

It is relatively clear that Kant thinks that (theoretical) reason has a constitutive end: to find a condition for any given conditioned object. It is more controversial to claim that understanding (*Verstand*) has a constitutive end, but this is one of the central claims of the third *Kritik*. In both the published and the unpublished ('first') Introductions to that work, he argues that the 'formal purposiveness' of nature is an *a priori* principle of reflecting judgment. He defines formal purposiveness as follows: "the correspondence of a thing with that constitution of things

---

<sup>46</sup> Cf. A258/B314.

<sup>47</sup> A307/B364.

<sup>48</sup> A327/B384.

<sup>49</sup> A329/B385.

that is possible only in accordance with ends is called the **purposiveness** of its form" (Ak. 5:180). To say that X is purposive for Y is to say that X is only possible if X has an end, and that end furthers the end of Y. A thing is represented as formally purposive if its form (in this context, structure) is represented as being only possible if it has some end. That reflecting judgment represents nature as formally purposive means that reflecting judgment represents nature's systematic form (e.g. its division into species and genus) as possible "only in accordance with ends", that is to say, possible only under the assumption that this system of nature has some end. For what (or whose) end does reflecting judgment represent the systematic form of nature as purposive? Our capacity for cognition (*Erkenntnisvermögen*), in other words, the understanding.<sup>50</sup> To represent nature's systematic form as purposive for our *Erkenntnisvermögen* means representing that systematic form as possible "only in accordance with ends", namely ends that further the ends of our cognitive capacity. So to represent nature's systematic form as purposive for our cognitive capacity presupposes that this cognitive capacity has *ends* and that the systematic form of nature furthers those ends. This is what Kant means when he writes:

since the lawful unity in a combination that we cognize as in accordance with a necessary aim (a need) [*Absicht (einem Bedürfnis)*] of the understanding, but yet at the same time as contingent in itself, is represented as a purposiveness of the objects (in this case, of nature), thus the power of judgment [...] must think of nature with regard to the latter in accordance with a **principle of purposiveness** for our faculty of cognition. (Ak. 5: 184)

The systematic form of nature is contingent for our understanding because the understanding leaves open the possibility that the multiplicity of species and genres in nature might be so great that we could never discover them (or empirical laws governing them). This systematic unity is not provided by the understanding, but it furthers a necessary aim, an end, of the understanding: to cognize empirical objects under laws (which requires species and genus concepts). This systematic form, therefore, is represented as purposive for the understanding: it serves the ends (*Zwecke*) of the understanding. That the systematic form of nature is represented by reflecting judgment as *purposive* for our understanding is

---

<sup>50</sup> Ak. 5:174.

repeated throughout the third *Kritik*.<sup>51</sup> Something can be represented as purposive for any cognitive faculty only if that cognitive faculty has an end, so our understanding has an end: to cognize empirically given objects under empirically given concepts/laws. Clearly, understanding cannot achieve that end unless it forms logically consistent concepts and judgments. So part of understanding's constitutive aim is: non-falsehood.<sup>52</sup>

When I say that 'thought as such aims at non-falsehood' I mean that non-falsehood is a constitutive end of the capacity for thought (discursive representation in general). I have argued that (theoretical) reason and understanding have non-falsehood among their constitutive ends (more precisely, that non-falsehood is a component of their constitutive ends). Since understanding and reason are the two discursive faculties, I take this to be sufficient evidence that non-falsehood is among the ends of discursive thought as such. It does not *entail* this conclusion; it is compatible, strictly speaking, with these facts that, although understanding and reason do aim at non-falsehood, this is not part of what makes them discursive faculties. I will ignore that possibility in what follows.

I have begun reconstructing Kant's explanation of how the unity of apperception explains the bindingness of logical laws upon discursive representation by arguing that discursive representation (thought) as such has a certain end – non-falsehood – and that logical laws relate to the fulfilment of this end. But in order to understand how some normative laws might apply to a certain representational activity with a given end (or 'aim') we must understand the relation between those normative laws and that end. This is perhaps the place where Kant leaves the most to his readers' reconstruction. I am going to assume what I take to be the following minimal conditions: the negative rules our thought must satisfy to achieve one of its constitutive ends (non-falsehood), described at A59/B84, become normatively binding on us (they become logical *laws*) only if (i) we are capable of following or not following them (ought implies can),<sup>53</sup> and (ii) capable of representing ourselves as

---

<sup>51</sup> Cf. Kant's repeated references to the *a priori* principle of the "purposiveness of nature for our cognitive faculty (*Erkenntnisvermögen*)" (e.g. Ak. 5:182, 184, 185, 186).

<sup>52</sup> It is not merely that non-falsehood is a means to the end of the understanding; the understanding aims to cognize empirical objects, that is, to cognize them *accurately* and logical consistency (non-falsehood) is a constituent of that end (see above), not merely a means towards it.

<sup>53</sup> A principle to which Kant appears to be committed at A548/B576.

either following or not following them, and (iii) allowing the capacity mentioned in (ii) to guide the activity of the capacity mentioned in (i). For a normative law to apply to the activity of some capacity, it not only has to be something we both can satisfy and can fail to satisfy, but we must be able to represent ourselves as failing to do so, and our representation of the law has to be part of the explanation of why we do or do not fail to satisfy the law. In other words, our representation of our activity as satisfying or failing to satisfy the law cannot simply be *passive*; we must be able to actively bring our activity into agreement with the law *because* of our representation of the law. If we are merely capable of *observing* whether our activity agrees with the putative law, but not of modifying that activity in light of its agreement or non-agreement with the putative law, then that law lacks normative force for our activity (it is not a 'law' in the sense defined above). Kant, as I have interpreted him, claims that the unity of apperception explains how logical laws can be normatively binding on us. Although he does not spell it out for us, it is not difficult to find a plausible story in the Transcendental Deduction about how the unity of apperception does this.

Kant states at the beginning of §16 of the B Deduction that "the *I think* must be able to accompany all of my representations".<sup>54</sup> To 'accompany' a representation with the 'I think' is to explicitly attend to that representation as *my* representation. This means that my representations have a certain kind of unity, albeit in a very minimal sense of 'unity' (although the unity required for self-consciousness will be strengthened significantly over the course of the Deduction): I can explicitly attend to them as *my* representations, while I cannot do so to your representations (although you can). Kant calls this 'unity of apperception'. Clearly, unity of apperception involves unity of consciousness. I can become explicitly conscious of any of my representations (accompany them with the 'I think'); I can direct my awareness at any of my representations. However, unity of apperception means more than that. For if that were all there is to the unity of apperception, I might have unity of apperception even though I am in general totally unaware of my representations but could direct my attention to them at will. Unity of apperception requires that, even when I am not consciously attending to my representations, I am still implicitly aware of them. In other words, it requires

---

<sup>54</sup> There is an extensive literature on the B Deduction, the notion of apperception, and, indeed, this very sentence. I do not have space in this essay to engage with it. However, I have learned much from Longuenesse (1998) and the discussion of apperception in Van Cleve (1999).

that were I to explicitly attend to one of them, that explicit attention would be an activation of some awareness of that representation that was already in place.

In §17 Kant extends this analysis by claiming that the representational manifold in an intuition (an immediate singular representation of an object) must have the unity of apperception; so far, this is a direct consequence of the claims of §16. He goes further, though, in claiming that if I am conscious of a unity of representations constituting an intuition of an object – as I must be if that manifold stands under the unity of apperception – then I must ‘unite’ those representations in a concept of an object. In other words, if I am consciously intuiting an object, and that intuition is composed of a manifold of representations, then I must be conscious of those representations (though perhaps not explicitly) and ‘unite’ them (though perhaps not explicitly) under a concept by thinking of them as a manifold of representations that compose an intuition of an object of a particular kind. For instance, if I consciously visually perceive a dog running across my yard, then I am conscious of that perception as composed of overlapping perceptions of the dog at different temporal intervals, and I am conscious of those overlapping perceptions as united under the concept ‘dog’. This conceptual unity among a manifold of representations is something my mind does (though I may not be explicitly attending to it). I do not passively receive this information from the senses; the fact that my total perceptual experience includes as part of its content that this manifold of representations is a manifold of representations of a *dog* is due to my mental activity.<sup>55</sup> Two points are crucial here. First, because I unify this manifold of representations under a *concept* (a general representation) I can do so incorrectly. Second, because this unity is effected by an act of my mind, I can be subject to normative criticism if I make an error. If I did not unify this manifold through a general representation (a concept) but only by some essentially indexical or demonstrative representation (e.g. *that-there*) then I could not in principle go wrong: for whatever I am intuiting is *that-there*. By unifying the manifold through a general representation, though, I open my act to the possibility of *error* because I represent the object of the intuition as having the general features constitutive of the concept; if the object turns out not to have one of those features, I have incorrectly unified the manifold. Similarly, if my senses (without any input from the understanding or apperception) represented the object

---

<sup>55</sup> B129–130.

of the intuition as a dog, then if the object turned out not to be a dog, I could be said to have misperceived, but normative criticism would be out of place. It would be wrong to say I should *not* have seen it as a dog if my senses simply deceived me. As it is, I might represent a manifold of representations of an object under the concept 'dog' but then upon receiving more information see that it is not a dog ('dog' was the wrong concept under which to unify those representations), for the object is actually a wolf. The objective unity of apperception, the unity a manifold of conscious representations is to have if it is going to be about an intuited object, brings in the possibility of normative evaluation of our following of conceptual rules.<sup>56</sup>

For our purposes, the key text here is §19 of the B Deduction, "The logical function of all judgments consists in the objective unity of the apperception of the concepts contained therein." After an initial paragraph in which he rejects the "explanation that the logicians give of a judgment in general" Kant explains his own view in the second paragraph, beginning with the crucial claim that "a judgment is nothing other than the way to bring given cognition to the **objective** unity of apperception" (B141f). The objective unity of apperception was originally introduced by Kant as a unity among manifold representations in an intuition, but here he extends it to a manifold of concepts in a judgment. This means that Kant's claim that "a judgment is nothing other than the way to bring given cognitions to the **objective** unity of

---

<sup>56</sup> Contrary to appearances, this is not only compatible with Kant's repeated claim that the senses do not *judge*, but in fact expresses the very same point (from a different direction). One finds repeated throughout the lectures on logic a similar story about the possibility of error: the senses by themselves do not err (because they do not judge, a point made in the *Kritik* at A293/B350) and the understanding by itself would not err (e.g. Ak. 24:526–527, 720, 824). The origin of error is in the interaction of sense and understanding. But if we take seriously the claim of the Transcendental Deduction that "all manifold, insofar as it is given in one empirical intuition, is determined in regard to one of the logical functions of judgment" (B143) then some function of judging is active in any perceptual consciousness of an object that can belong to the unity of apperception, i.e. be a representation this "for me" (B132). I take this to mean that conscious perceptual awareness of objects in our environment is the joint-product of passive reception of sensory materials (sensibility, which does not judge, and thus does not err) and the synthetic activity of the understanding. If this is correct, then given that Kant locates the possibility of error in the interaction of sensibility and understanding, it follows that consciousness perceptual awareness of objects in our environment can *err* (unlike mere passive receipt of sensory matter).

apperception” means that a judgment is nothing other than the way to unite concepts together so that they possess the objective unity of apperception. Just as the manifold of an intuition is consciously united in a concept of the intuited object, Kant claims, concepts are consciously united in a judgment that asserts a relation between their domains. This is what distinguishes a judgment from mere ‘association’: when I judge  $\langle \text{All } S \text{ are } P \rangle$  I am not merely having an associative episode involving the concepts  $\langle S \rangle$  and  $\langle P \rangle$ . I am making a claim about objects.

The natural question is, what in the case of judgment plays the unifying role played by the concept in the case of intuitions? Kant’s answer is “that is the aim of the copula is in [judgments]: to distinguish the objective unity of given representations from the subjective” (B141–142). Kant’s reference to ‘is’, however, is disappointing, for we will look in vain among the table of logical functions of judgment for ‘is’. I think Kant’s point, though, is that all judgments *as such* are unified by the copula, and the different logical functions of judgment are different ways of determining the copula.<sup>57</sup> But just as in the case of the synthesis of a manifold of representations in an intuition, two points are crucial: in judging, our minds are active, not passive; and the ‘copula’ by which we unify concepts into a judgment introduces the possibility of *error*. Combining those two points, the fact that judgments are acts of the mind and that we can incorrectly unify concepts into judgments opens the possibility of normative criticism of our judgmental acts. That judgment is an *act* means merely that our judgmental representations (our representations whose content is judgmental) are not passively received by the senses: if the impact of the world on my sensory organs by itself made it the case that I believe that there is a dog in front of me, it would be wrong to normatively criticize me for this judgment. I might be wrong that there is a dog in front of me, but it would not be the case that I *ought not* to have judged that.

The question then is, how can the copula be wrong? In other words, how can I be incorrect in synthesizing two concepts  $\langle A \rangle$  and  $\langle B \rangle$  into the judgment  $\langle A \text{ is } B \rangle$ . Well, there are lots of ways I can be wrong! But in this essay we are interested in a very specific way I can be wrong: a way I can be wrong that does not depend upon which particular concepts  $\langle A \rangle$  and  $\langle B \rangle$  are. The principle of contradiction is one such rule. If  $\langle B \rangle = \langle \sim A \rangle$ , then  $\langle A \text{ is } B \rangle$  is a false judgment, regardless of what concept

---

<sup>57</sup> Cf. B332/A266, Ak. 9:104, Ak. 17:344. See however *JL* §25 where Kant claims that in hypothetical judgments “*Consequenz*” takes the place of the copula.

$\langle A \rangle$  is. So, it might be thought, that my judgment that  $\langle A \text{ is } \sim A \rangle$  can now be seen to be subjective to normative criticism: the judgment is an *act of mine* (I synthesized it, rather than passively receiving it from the senses), and the judgment is *false*.

But this *only* issues in an explanation of why I am subject to normative criticism in judging  $\langle A \text{ is } \sim A \rangle$  if we assume that: (i)  $\langle A \text{ is } \sim A \rangle$  is false, regardless of what concept  $A$  is, and (ii) a judgment is subject to normative criticism if it is false. In other words, we need to assume (i) that the principle of contradiction is *true* regardless of the concepts involved, that (ii) that thought as such aims at *non-falsehood* and that thoughts that are false are subject to normative criticism in virtue of failing to be non-false. The philosophical lessons of this are that we can derive from the unity of apperception an explanation of why logical norms apply to all thought as such, but only if we assume the truth of the principle of contradiction. This means that the unity of apperception does not explain *why* the principle of contradiction is true; at most, it explains why the principle of contradiction normatively constrains all thought as such. But if we take seriously Kant's claim that the "entirety of logic" is attached to the unity of apperception and explained by it, this entails that (what Kant calls logic) does not explain why the principle of contradiction is *true*. Logic does not explain why there are no true contradictions, but why *we* are normatively bound not to have any internally contradictory thoughts. The unity of apperception explains the normativity, not the truth, of logic.

Now we can give a unified account of how the unity of apperception explains why logical norms apply to all thought as such. Thought constitutively aims at non-falsehood. Logical laws describe the negative conditions on non-falsehood; no thought can violate these principles and still be non-false. Because judgments and the concepts that compose them stand under the unity of apperception, these judgments are the product of synthesis by the judging subject (they are spontaneously generated, *not* passively received). The synthesizing of concepts into judgments involves uniting them under what Kant calls the 'copula is' in §19 of the B Deduction, which, I argued, is the genus of which different logical functions are the species (the different logical functions are determinate forms of the general copula). Consequently, judgments can be incorrect or correct in the following minimal sense: they can obey or not obey the logical principles (principles of non-falsehood) specified in terms of the various logical functions. But all of this stands under the unity of apperception; so the judging subject is aware of the judgments he or she is making, and can be aware of their agreement or non-agreement with the



logical principles. Furthermore, since these judgments themselves are spontaneous (not passive), the judging subject's consciousness (apperception) of these judgments can be efficacious in modifying or rejecting these judgmental acts. In other words, the subject can modify his or her spontaneous judgmental acts in light of her apperceptive awareness of their agreement or disagreement with logical principles. These logical principles articulate negative conditions on thought's satisfaction of its constitutive end: non-falsehood. So, I conclude, these logical rules obtain the status of norms for all thought as such. The unity of apperception explains why logical principles are normative for thought, why they are *laws*.

To bring this discussion full circle, this also explains why logical norms are formal in Bolzano's sense, namely, why they are substitution-formal. It is because the rules of logic are substitution-formal: they can be specified as abstract schemata from which determinate rules for particular concepts and judgments can be obtained through substitution of determinate concepts for concept-variables. I argued in Section §2 that the most worked-out part of Kant's logic – the theory of inference – is substitution-formal. These logical principles become norms for thought because they specify negative principles that any thought must satisfy in order to be non-false, and non-falsehood is the constitutive aim of thought as such. The unity of apperception then explains how these principles obtain the status of norms for the discursive representational acts of judging subjects, and since the principles are substitution-formal, the resulting norms are substitution-formal as well. Consequently, we can agree with Bolzano that the Kantian thesis of the formality of logic "*scheint noch nicht deutlich zu seyn*", for Kant has left to us much of the work of rendering it clear, but there is at least room to disagree whether it is "*nicht richtig genou entschieden*".<sup>58</sup>

### Abbreviations for works of Kant

A/B *Kritik der reinen Vernunft*, Ak. 3 & 4. Cited by page number in 1st edition of 1781 (A) and 2nd edition of 1787 (B). Translations generally follow Kant (1997), with occasional alterations.

---

<sup>58</sup> I'd like to thank Sandra Lapointe and Clinton Tolley for reading, and commenting on, an earlier draft of this essay. I'd also like to thank Jack Woods, Catharine Diehl, Reed Winegar, James Kreines, Franz Knappik, Tobias Rosefeldt, Peter Yong, Bianca Ancillotti, and the other members of the Colloquium for Classical German Philosophy at Humboldt-Universität zu Berlin for a very helpful discussion.

- Ak. Kant (1905). Cited by volume and page number (e.g. Ak. 29: 1034).
- KrV *Kritik der reinen Vernunft*, Ak. 3 & 4. Cited by page number in 1st edition of 1781 (A) and 2nd edition of 1787 (B). Translations generally follow Kant (1997), with occasional alterations.
- JL *Jäsche Logik*, Ak. 9. Translations generally follow Kant (1992), with occasional alterations.
- Refl. *Kants handschriftlicher Nachlass* ('Reflexionen') in Ak. 14–18. Cited by four digit number.

Where no translation is listed, translations are my own.

### Works of Kant cited

- Kant, I. (1905). *Kants gesammelte Schriften*. Edited by the Berlin-Brandenburg (formerly Royal Prussian) Academy of Sciences. Berlin: Walter de Gruyter.
- Kant, I. (1992). *Lectures on Logic*. Translated and edited by J. Michael Young. Cambridge: Cambridge University Press.
- Kant, I. (1997). *Critique of Pure Reason*. Translated and edited by P. Guyer and A. Wood. Cambridge: Cambridge University Press.

### Works of Bolzano cited

- Bolzano, B. (1837). *Wissenschaftslehre*. 4 vols. Sulzbach. Republished by Wolfgang Schultz, Leipzig, 1929–1931.
- Přihonský, F. (1850). *Neuer Anti-Kant*. Bautzen: A. Weller. Republ. (2003). Edited by E. Morscher. St Augustin: Academia.

### Secondary literature

- Lapointe, S. (2011). *Bolzano's Theoretical Philosophy*. New York: Palgrave Macmillan.
- Lapointe, S. (2012). 'Is Logic Formal? Bolzano, Kant and the Kantian Logicians', *Grazer Philosophische Studien*, 85, 11–32.
- Longuenesse, Béatrice. (1998). *Kant and the Capacity to Judge*. Princeton: Princeton University Press.
- MacFarlane, J. (2002). 'Frege, Kant, and the Logic in Logicism', *The Philosophical Review*, 111, 25–65.
- Rusnock, P. (2011). 'Kant and Bolzano on Logical Form', *Kant-Studien*, 102, 477–491.
- Rusnock, P. and Burke, M. (2010). 'Etchemendy and Bolzano on Logical Consequence', *History and Philosophy of Logic*, 31.1, 3–29.
- Siebel, M. (1996). *Der Begriff der Ableitbarkeit bei Bolzano*. St Augustin: Academia.
- Siebel, M. (2002). 'Bolzano's Concept of Consequence', *The Monist*, 85.4, 580–599.

- Stang, N. (2012). 'Kant on Complete Determination and Infinite Judgment', *British Journal of the History of Philosophy*, 20.6, 1117–1139.
- Tarski, A. (1936). 'The Concept of Logical Consequence', Republished in (1956). *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. Translated by J.H. Woodger. Oxford: Oxford University Press.
- Tolley, C. (2012). 'Bolzano and Kant on the Nature of Logic', *History and Philosophy of Logic*, 33, 307–327.
- Van Cleve, J. (1999). *Problems from Kant*. Oxford: Oxford University Press.
- Woods, J. (forthcoming). 'Logical indefinites', *Logique et Analyse*.