Tyke Nunez

Logical Mistakes, Logical Aliens, and the Laws of Kant's Pure General Logic

Washington University

Philosophy Department

Topic Areas:

Kant, Kant's pure general logic, logical mistakes, logical aliens, normativism about logic, formalism about logic

Abstract

There are two ways interpreters have tended to understand the nature of the laws of Kant's pure general logic. On the first, these laws are unconditional norms for how we ought to think, and will govern anything that counts as thinking. On the second, these laws are formal criteria for being a thought, and violating them makes a putative thought not a thought. These traditions are in tension insofar as the first depends on the possibility of thoughts that violate these laws, and the second makes violation impossible. In this essay I develop an interpretation of Kant's pure general logic that overcomes this tension. It accounts for the possibility of logical mistakes, as the first tradition does, while still establishing the absolute impossibility of logical aliens, as the second tradition does. I then argue that the formalist insight that illogical exercises of the understanding are not alternate ways coherent thoughts could have been, but are mere confusions, is fundamental for achieving a proper understanding of the absolute normativity of the laws of pure general logic.

LOGICAL MISTAKES, LOGICAL ALIENS, AND THE LAWS OF KANT'S PURE GENERAL LOGIC

1 Introduction

Interpreters are divided over Kant's view of the nature of logic's laws. According to normativists, logic's laws are constitutive norms for how we ought to think.[1] According to formalists, logic's laws are indications of the form or nature of thought.[2] Interpreters have taken these two roles of logic's laws to generate a tension over the possibility of illogical thought: normativists seem able to allow it, formalists do not. I aim to show why Kant is unbothered by this tension, and how his conception of pure general logic allows for both.[3] More importantly, I also show that this resolution allows us to explain the absolute normativity of logic's laws.

I begin presenting the normativist and formalist commitments by contrasting them with a third ontological account of logic's laws. I then set up the apparent tension by distinguishing two species of illogical thoughts: logical mistakes and logically alien thought. Although ultimately my concern is interpretive, the tension is broader. It seems that any account of logic that rules out the intelligibility of logical aliens will be unable to account for the possibility of logical mistakes and vice versa. I first show that Kant is able to avoid this tension because he maintains thinking is the successful exercise of the understanding, the fallible faculty for cognition. I then turn to the interpretive matter of reconstructing Kant's account of logical mistakes and presenting how he would reject the apparent possibility of logical aliens. Finally, I return to the relationship between logic's absolute normativity and its peculiar formal necessity, and argue that for Kant the latter grounds the former.

---

[1] For example, John MacFarlane (2002), Béatrice Longuenesse (2005, ch. 4), Lanier Anderson (2005), Robert Hanna (2006), Nick Stang (2014), Huaping Lu-Adler (2016), Jessica Leech (2017) and Matthew Boyle (MS) have defended normativist readings.

[2] For example, Hilary Putnam (1994), James Conant (1992), Clinton Tolley (2006), and Melissa Merritt (2015), all defend formalist readings.

[3] It is noteworthy that there is a discussion in our contemporary philosophy of logic that concerns a similar tension to the one I will be discussing. Unsurprisingly, some participants in it are directly influenced by Kant: for example, Jessica Leech (2015) defends a normativist view, while Manley Thompson (1981) defends a formalist view. Nonetheless, there are also contemporary normativists—e.g., Hartry Field (2015)—and formalists—e.g., Charles Travis (MS)—who are not Kant scholars.

2 Ontological, normative, and formal logic

Between normativist and formalist interpreters there are points of broad agreement. For example, everyone holds that one of Kant's central, abiding innovations is distinguishing pure general logic from both the empirical psychological investigation of the laws governing particularly human thinking, and from metaphysical inquiry into the necessary nature of the soul or reality (Bvii).[4] Further, everyone maintains that Kant does this in part through logic's 'generality' and 'purity.' Pure general logic is 'general' because it studies "the absolutely necessary rules of thinking," and thus contrasts with particular logics that study the rules for correctly thinking about a certain kind of object (e.g., physical ones) (A52/B76). It is 'pure' because it abstracts "from all empirical conditions" under which we think, which contrasts with applied logic: the study of concrete psychological rules governing how human beings actually do think (A53/B77; *JL*, 9:14). So both traditions agree that pure general logic studies the necessary laws of thinking, essential to any thought whatsoever—no matter what it is *about*, or the specific nature of the *thinking subject*.[5]

Because of this agreement, we need to look beyond Kant's characterization of logic as 'pure' and 'general' to get at the differences between the formalist and normativist interpretations. Both formalists and normativists have introduced their interpretations by contrasting the conceptions of logic behind them with an ontological conception of logic. They have tended to do this through examining different strands in Frege's thinking. For example, John MacFarlane contrasts a normativist strand, and James Conant a formalist strand, with an ontological strand in Frege's thought. I think both MacFarlane and Conant do an excellent job of spelling out the philosophical motivations behind normativism and formalism in logic, and so behind their respective interpretations of Kant. Because of this, the way I would like to introduce these interpretations of Kant is by examining these three conceptions of logic, and how MacFarlane and Conant see the contrasts between them.

Logic, on an ontological conception, is a descriptive science of the laws of thought where thoughts are things or complexes of things. These laws say how things are, and are of the form 'such and

---

[4]Kant's *Critique of Pure Reason* is cited using the standard A and B edition numbering; Kant's other works are cited by volume and page number of the Akademie edition. Abbreviations used for individual works are given at the end of the paper.

[5]My focus is on pure general logic. Much of what I say will apply *mutatis mutandis* to transcendental logic, but its receptive and ontological dimensions add significant complexity that make treating it a task for another occasion.

such is the case.' If something violates these laws then it is not a thought, just as if something violates the laws of physics then it is not a physical thing. The laws of logic, on this conception, are maximally general truths that other descriptive sciences, like physics, must presuppose. We might differentiate different versions of the ontological conception through how they conceive of the necessity of logic's laws. For example, perhaps these laws are necessary because they are eternal truths laid down by God, or because they hold in an absolute platonic realm, or because they hold of all possible worlds. Regardless, these laws are *facts* that structure all other facts.[6]

Frege, at times, seems to endorse an ontological conception of logic, and on this strand in his thinking, his logic can also seem both 'general' and 'pure.' The laws of logic, on this view, are maximally general descriptive truths. Consider, for example, Frege's Basic Law IIa: '$\forall x Px \rightarrow Pa$,' where P and a are free variables. Frege renders this in words as "what holds of all objects, holds also for each" (Frege, 1893, §20), and seems to take it to permit, *a priori*, any assertion of this form in *Begriffsschrift*. While as a permission it is normative, presumably the underlying general truth would be '$\forall F \forall y (\forall x F(x) \rightarrow F(y))$.' On the ontological strand in Frege's thinking, this relationship between objects and concepts would be the logical law. This law is 'general' insofar as it holds of all concepts and all objects, in contrast with, say, the laws of physics or geometry, which are true of a particular domain. And this law is 'pure' insofar as appealing to it can grant one's claims an *a priori* justification, because it does not depend on anything psychological or the existence of any specific kind of object in the empirical world.

MacFarlane argues that although Frege holds logic's laws are ontological or descriptive in their content, they imply norms for how we ought to think. To present this point, McFarlane develops one of Frege's analogies. The laws of physics describe what is physically possible. Judgments about the physical world can be assessed as correct or incorrect depending on whether they are right about the physical facts. Because of this, although the laws of physics are descriptive, "they have prescriptive consequences for anyone engaged in the 'game' of thinking about the physical world: such a thinker *ought not* to make judgments that are incompatible with them" (MacFarlane, 2002, p. 36). This is not to say that one's judgments about the physical world will conform to these

---

[6]Both Thompson (1981) and Putnam (1994) introduce and develop this ontological conception of logic, as well as the contrasting formal conception. This is so despite Thompson's main target being an earlier view of Putnam's.

laws—one's judgments about the physical world can be incorrect, after all. Nor is it to say that one must be aware of these laws in order to think about the physical world—one can be assessable by the norms of judging about the physical world while not knowing them. Rather, it is only to claim that if one makes a judgment about the physical world, then this judgment is *evaluable* according to these norms. And it is in this sense that physical laws also "provide constitutive norms for the activity of thinking about the physical world" (MacFarlane, 2002, p. 37).

According to MacFarlane's Frege, "while physical laws provide constitutive norms for thought *about the physical world*, logical laws provide constitutive norms for thought *as such*" (MacFarlane, 2002, p. 37). With both, there is a system of descriptive laws that generate a corresponding system of norms and the only difference is in the size of their domain. For both, one can violate the norms or be unaware of them, while still being assessable by them. And just as one could only evade the force of the norms governing judgments about the physical world by opting out of this activity, one can only evade the force of logic's norms by opting out of the activity of judging altogether.

Now, MacFarlane maintains that Kant's laws of logic would be identical to the norms Frege's logical laws generate. This is because he thinks both are constitutive norms governing all thought. And this is despite acknowledging that Kant's logic has no corresponding ontological component. So along these lines, when Kant claims logic studies "the absolutely necessary rules of thinking, without which no use of the understanding takes place" (A52/B76), MacFarlane understands this necessity normatively and interprets Kant's claim as indicative of how these rules are authoritative for all thinking *as such* (MacFarlane, 2002, p. 35; 43).

The normative conception of logic that MacFarlane attributes to Kant is then the following. Logic's laws say how thoughts ought to be and are of the form 'T ought (not) to ϕ.' If our thinking fails to live up to these norms they still govern it, just as even if we fail to live up to the laws of morality those laws will still govern us. The laws of logic, on this conception, are constitutive norms for thinking *as such*, where to be a 'constitutive norm' is to govern a thing essentially, in virtue of the kind of thing that it is. And this comes together with the claim that the only way to evade the force of logic's laws is to opt out of the activity of thinking and judging altogether. Finally, on this view, these laws are necessary because they are *unconditional*: they apply to all thinking as such (MacFarlane, 2002, p. 43).

4

Turning to Conant, he introduces the formalist conception of logic through a tension between it and the ontological strand in Frege's thinking.[7] In his attempts to communicate the self-evidence of his basic laws, Frege resorts to what he calls "elucidations" (*Erläuterungen*). For example, take the statement 'no concept is an object.' In it 'concept' is in a place that is reserved for predicates true of objects. Because of this, any attempt to render it in *Begriffsschrift* will fail and, on Frege's view, it is nonsense, even though he hopes it will help his readers recognize a real feature of thoughts—the difference between concepts and objects.[8] This example is characteristic of Frege's elucidations in general. Although they seem to be grammatical statements in everyday language, they cannot be formulated in *Begriffsschrift*. And because of this they only give the illusion of sense—of expressing the way things are. Strictly speaking, they are nonsense, even though they are illuminating. Stepping back, there is then a divide between two kinds of facts, those that can be expressed in thoughts and those ineffable facts that can only be indicated through elucidations. For this reason, on the ontological strand even the syntactic laws of logic, which merely govern the construction of well-formed thoughts, are substantive. They circumscribe the limits of thought, and there are some ineffable facts that lie on the other side of this limit which we cannot think, but only elucidate.

Conant, however, finds the germs of an alternative formalist conception of logic in the dissolution of Frege's apparent thought experiment about logical aliens. Frege raises their possibility in the midst of trying to refute the psychologistic philosopher of logic, who takes the laws of logic to be empirically established generalizations. Logical aliens are "beings who could make judgments that contradict our logical laws" (Frege, 1893, p. xvi). For example, they might make judgments that contradict the law of identity, and the psychologistic philosopher maintains that while we must judge in accordance with this law, their laws of logic might allow their judgments to violate it. Now, Conant's Frege maintains that when the psychologistic philosopher of logic is asked the question, 'whose inferences are right, ours or theirs?,' he is faced with a dilemma:

> either 1) he can claim that his account reveals that the judgments of the aliens conflict
> with ours, in which case his idea of one judgment's conflicting with another can be
> shown to tacitly rely upon the idea of their logical incompatibility (that is, upon a non-

[7]In developing this tension, Conant is especially building on an essay by Thomas Ricketts (1985).
[8]Ricketts (2010, § 5) spells out this point in detail.

psychological notion of incompatibility), or 2) he can refrain from telling us anything about the logical relation in which their judgments stand to ours, in which case he can tell us nothing about their thought whatsoever. (Conant, 1992, p. 146-147)

Behind this dilemma, Conant reads Frege as trying to get his opponent to see that no sense can be made of a logical alien. Conant notes that "there isn't any sense to be made of the idea of undertaking to disagree with a principle of logic [because] it is these principles which make both agreement and disagreement possible" (Conant, 1992, p. 147).

Now, at this stage, the proposal about logical aliens looks as though it has the structure of a Fregean elucidation. It looks as though Frege is arguing that because logical aliens are supposed to be reasoning beings, yet the laws of logic are the touchtones of rationality, there cannot be logical aliens. With this, as Conant puts it, it seems we have "grasped the content of the thought experiment—what it would be for beings to be able to think in this remarkable way—and subsequently gone on to reject this possibility" (Conant, 1992, p. 149). But if we take the statement, 'illogical thought is impossible,' to express a sense, then it seems we are equally committed to there being a sense to 'illogical thought is possible,' even though the key move so far was rejecting this. Thus, 'illogical thought is impossible' must be nonsense. But it seems to be illuminating nonsense: it gets at an ineffable fact that can be at best shown, not expressed. Thus, we have a Fregean elucidation.

Formalist philosophers of logic—such as Kant and Early Wittgenstein, according to Conant— would maintain Frege concedes too much to the psychologistic logician, if he concedes that the impossibility of logical aliens is an ineffable fact. These philosophers reject the distinction between illuminating and non-illuminating nonsense. Both maintain the practice of philosophy is rife with illusion, with philosophers who think they have made sense when they haven't. And while these philosophers can be shown to have fallen prey to an illusion, these illusions are not indicative of ineffable facts, but are mere confusions. So when Kant claims that logic studies "the absolutely necessary rules of thinking, without which no use of the understanding takes place" (A52/B76), Conant understands this necessity formally, where these rules are conditions on being thoughts at all and "abstract from all objects of cognition" (Conant, 1992, p. 131).[9]

_____

[9]Conant, of course, is one of the 'resolute' readers of the *Tractatus*.

The formal conception of logic that Conant attributes to Kant is then the following. On it, while thoughts are about facts and objects, they are not themselves facts or complex objects. The laws of logic are "criteria of the form of truth" in that agreement with them is required for thoughts to be about objects, i.e. for the question of their agreement with a fact to even arise (A60/B84). For this reason, logic's laws do not indicate that 'such and such is the case,' and their violations do not indicate something that is of the right kind to have obtained. Rather, on this view, the laws of logic are rules for combining representations or thoughts into new thoughts. They are the rules that these operations of combination must agree with, if the result is not going to be nonsense. In this way, they only articulate the form of coherent thought, the system of coherent relations thoughts stand in with one another, no matter how things are with what they are about. So logic's laws merely articulate thought's nature, but do not tell us anything concerning its material—the objects and facts (or states of affairs) that thoughts represent. And, on this view, logic's laws do not put a substantive limit on thought because they are not a guide to what is. Their violations are mere confusions, not indications of ineffable facts.

3   The tension

MacFarlane and Conant interpret Kant's claim about the necessity of logic's laws in apparently contradictory ways. Nonetheless, as I will elaborate, they are both correct. The obstacle to reconciliation is that normativist interpreters allow illogical thought, while formalist interpreters do not. This tension is related to a distinction between two kinds of cases of such 'thought': logical mistakes and logically alien thought. Logical mistakes are when we, for instance, draw a conclusion invalidly or do not recognize a contradiction in our thinking. They are violations of logical law by the beings that those laws govern. In contrast, logical aliens, as Frege puts it, are beings "whose laws of thought directly contradict our own" (Frege, 1893, p. xvi). So logically alien thought might violate our logical laws, but these laws do not govern them. Thus, if these two species of illogical thoughts can be distinguished, logical mistakes will violate logical laws, but do not threaten these laws; while logically alien thought might not violate its laws, but would threaten the universality or necessity of logical laws.

Normativists hold we bungle our way into making logical mistakes from time to time. They maintain that logic's laws must be violable to be constitutive norms for how we ought to think—that for a normative 'ought,' *ought to ϕ* implies *could fail to ϕ*.[10] So the possibility of logical mistakes is closely tied to the normativist conception of the necessity of logical laws.

Formalists maintain that logically alien thought is not a genuine possibility, but at most the illusion of one. As Conant puts it, "we are simply unable to make *sense* of [the negation of a logical law] in a way which allows the question of [its] truth or falsity to arise in the first place" (Conant, 1992, p. 128).[11] unlike a false judgement I might have made about my misplaced keys, the negation of a logical law does not indicate a way that things could be—a coherent possibility. So the unintelligibility of logical aliens is closely tied to the formalist conception of the necessity of logical laws.

It can seem, however, that while normativists can accommodate the possibility of logical mistakes, they cannot properly reject the intelligibility of logical aliens; and while formalists correctly rule out the intelligibility of logical aliens, they cannot accommodate everyday logical mistakes.[12] This is because if logical mistakes are thoughts, and some of these mistakes violate logical laws, then it seems we can think the negation of logical laws. So if one rejects the intelligibility of beings that think according to laws that contradict ours, then it seems one should reject the possibility of logical mistakes, or if one accepts that such mistakes are thoughts, then it seems one should accept the intelligibility of such beings. And if I am correct that Kant gets the right verdict on both cases, then he will have to be able to resist the pressure both to exclude the possibility of mistakes and to allow the intelligibility of aliens.

---

[10]The normativity of logic should be understood relative to the act type—thinking—not the being whose acts are in question. Thinking is discursive cognition (A131/B170), which involves running through (sensible) material with the intellect. So God intuits and does not think (*VT*, 8:400n). Mistaken thought, according to Kant, results from the interference of a foreign power on the understanding—usually sensibility (A294/B350-A295/B351). And, as Boyle (MS, §2) nicely argues, Kant holds that the laws of logic *describe* how the understanding works, when nothing interferes. Now, what of a perfect thinker: a being with receptive and finite intellectual faculties, but whose understanding cannot deviate from its laws; would logic be normative for it? I know of no place where Kant discusses such a possibility and I do not know whether he would allow it. But supposing he would—because there is nothing inherent in the co-operation of sensibility and understanding that dictates they will in fact fall into error (see Engstrom (2009, §IV.4))—then logic would not be a canon for the *correction* of this being's cognitions. Still, the laws of logic are normative for the *assessment* of its thinking, because this kind of activity can go wrong, even if this being cannot.

[11]See also (Putnam, 1994, p. 257).

[12]I am grateful to an anonymous referee at this journal for pressing me to develop this tradeoff.

4   Cognition and the function of the understanding

Since Frege, it has become orthodox to view logic as the study of entailment relations between sentences or propositions, rather than as the study of the faculty of thinking. Influenced by this, both MacFarlane and Conant do not dwell on Kant's definition of logic as "the science of the rules of the understanding in general" (A52/B76). I think this definition, with its faculty-oriented conception of logic, is the key to overcoming the tension. I will present how this works, through developing Kant's account of thought as cognition and of the understanding as the faculty for this.

According to Kant, there are two sources of our knowledge of objects, "two stems of human cognition, [...] sensibility and understanding" (A15/B30; A835/B863). The function of sensibility is to "give" objects to us, while through the understanding "they are thought." (A15/B30; A51/B75; *JL*, 9:36; A294/B351n).[13] Still, although the representations of sensibility (intuitions) are in a sense cognitions (A320/B376-77), Kant will also often refer to the *understanding* as the faculty for cognition, or the faculty for cognition *properly speaking*, where he is drawing a contrast with sensibility (e.g., A50/B74; A78/B103; B137). Kant will explain cognition (*Erkentnnis*) in this stricter sense as "a whole of compared and connected representations." (A97).[14]

The understanding compares and connects representations through combining or synthesizing (B129-130). We see this in each of its characteristic activities of thinking: conceiving, judging, and inferring. On this picture, intuitions are specific to the time, place, and thing that they represent. And because of this specificity, they are singular. Concepts, however, are general representations, or features, that can be found in many other possible representations, as what is common to all of them (B40).[15] In this way, concepts indicate common features that unify or combine these representations (esp. intuitions). Judgments, in turn, unify or combine concepts. They connect

---

[13]In all of these passages Kant has in mind the the understanding in its "broad designation," which incorporates all of the higher faculties of cognition: "understanding, the power of judgment, and reason" (A130-131/B169; *PölL*, 24:505), and when I use the term this is what I will have in mind.

Relatedly, there are also uses of these faculties and their representations that I will not be addressing here—as, say, in the production of aesthetic pleasure through the free play of the understanding and imagination (*KU*, 5:217). My focus is solely on the role of these faculties and representations in thinking and cognizing—forming new whole concepts, judgments, and inferences—in the strict sense studied in pure general logic.

[14]Which is not to deny that intuitions *can be* run through, compared, and connected, it is just that they need not be. After all, they are the material for thoughts.

[15]Of course, there can be concepts that will only apply to one thing (e.g., the concept of the most real being, *ens realissimum* (A576/B604)). Still, in its *form* such a representation will be general, because it picks out the thing via a rule or feature, and rules or features can be common to many (e.g., *JL*, 9:91).

concepts together, thereby developing the concepts involved. For example, when I judge *bodies are heavy*, I develop my concept *body* by now including *heaviness* as another feature within it. Finally, inferences combine judgments. Paradigmatically, for Kant, these will be syllogisms where two judgments share a middle term, a concept. Through this middle term, one infers a conclusion, a third judgment, which further develops the concepts of the first and second judgments.

In thinking—conceiving, judging, and inferring—Kant takes the understanding to be comparing and connecting representations, combining and ordering them into new wholes. For this reason, in every thought, we can distinguish a matter and a form. The matter is the representations that the thought compares and connects together, and the form is the way that these representations are connected. It is the form of thinking that pure general logic studies. For example, logic will study inference forms, like 'all *A*s are *B*, all *B*s are *C*, therefore all *A*s are *C*.' These abstract away from specific judgments and designate a structural relationship that judgments may instantiate, in order to form an inference. Or again, logic will study judgment forms like 'not all *A*s are *B*,' which abstract away from specific concepts and designate a way concepts can be combined into judgments. Or it will even study the form of concepts, generality (*Allgemeinheit*), with respect to how concepts unify representations, while abstracting away from these specific representations. We can think of these formal rules on analogy with grammar: "Just as grammar is for passing judgment on language as to form," logic is for passing judgment on thoughts as to form (*WL*, 24:792).[16] And just as when we combine words in a way that violates the rules of grammar we arrive at linguistic nonsense, when we compare and connect representations in a way that violates the rules of logic we arrive at cognitions that are not in agreement with themselves (*JL*, 9:51).

Although we will want to develop both the account of logical mistakes and the rejection of logical aliens further, preliminarily, we can now see how this view of thoughts and logic can overcome the tension of the last section. The key is that thoughts are exercises or acts of a faculty. On an ontological conception, thoughts and the laws of logic are of the same kind. So if there are thoughts that negate logical laws, then the negation of a theorem of logic is thinkable, and logically alien thought is intelligible. There seems no space for a division in kind between logical mistakes and logically alien thought. On Kant's account, however, there is space. The understanding is a fallible

---

[16]See also: *WL*, 24:790-792; *DW-L*, 24:693-695; *BusL*, 24:609; *JL*, 9:11-13; *En-F*, 29:31; R1579, 16:19.

faculty. When its exercises violate its laws, these are logical mistakes. These mistakes, however, do not indicate an alternative way coherent thought or its laws could have been. And the mere fact of such mistakes does not suggest that a being with an understanding that operates according to different laws—a logical alien—is intelligible or possible. On this view, logical mistakes and logically alien thoughts are different in kind, and because of this Kant can accept the possibility of mistakes while rejecting logically alien thought as nonsensical.[17]

5 Logical mistakes

According to Kant, the law of non-contradiction is the fundamental law of pure general logic (A151-153/B190-192; A59/B84). Given this, there are many passages that seem to support either the possibility or impossibility of thinking a logical contradiction. For example, he will speak of contradictory cognitions as though they are perfectly intelligible, as when he says "a cognition is false if it contradicts itself" (*WL*, 24:826),[18] while also often denying the possibility of contradictory thoughts or cognitions, as when he says, "I can **think** whatever I like, as long as I do not contradict myself, i.e., as long as my concept is a possible thought" (Bxxvi, note).[19] Kant himself, however, does not seem very exercised by the tension. He sometimes even seems to affirm and deny the

---

[17]This difference is closely tied to another one. Although Kant will speak of thoughts as *products* of the understanding, this is not his fundamental way of thinking about thinking. Rather, fundamentally for Kant, thinking is always an *activity* or *act* of the understanding—an act of combining representations into whole cognitions. The ramifications of this are widespread. For example, because Kant has this act-oriented conception of thinking, the contemporary distinction between 'formation rules,' that govern whether a proposition is well-formed, and 'veridical rules,' that govern whether an inference is valid, does not apply. There are only acts of combination and rules governing whether these are successful or not. For Frege and early Wittgenstein, who distinguish formation from veridical rules, it is important to distinguish between senseless (*sinnlos*) propositions like tautologies and contradictions and nonsense (*unsinn*) that is not syntactically well-formed. In contrast, an important ramification of Kant's act-oriented conception of thinking is that there is no distinction to be drawn between senseless (but syntactic) contradictions and nonsense. There are only exercises of the understanding that fail to combine representations into whole cognitions.

[18]In addition to affirming this about cognition in general, Kant also seems to affirm specific types of illogical cognitions: First, he discusses fallacious inferences that violate logic's laws, where we are subject to a "logical illusion" because these imitate "the form of reason" (A296/B353; A61/B85-86; *WL*, 24:828). Second, in dialectic Kant is concerned with warding off the introduction of contradictions into our system of cognition or knowledge (e.g., Bxxxviii; *Prol*, 4:340-341; *Eberhard*, 8:194), so seems to maintain that we can make contradicting judgments. Third, Kant discusses concepts that contradict themselves, like "bright darkness," whose object is a *nihil negativum* (e.g. A291/B348; *MM*, 29:792).

[19]See also: Bxxvii; A155/B195; *Eberhard*, 8:195. Specifically, he will claim that contradictory concepts and judgments "are nothing," (A150/B189; A291/B348) or that "contradiction entirely annihilates and cancels them" (A151/B190-B191). He will even go on to claim that the principle of contradiction is inviolable (A152/B191), or that if the laws governing a power are essential (as he seems to think the "general and necessary rules of the understanding" that logic studies are (A59/B84)), "then the power cannot deviate from them" (*WL*, 24:824).

possibility of illogical thoughts or cognitions in a single breath.[20]

In this section, I develop an interpretation of Kant on logical mistakes that explains his seeming nonchalance over affirming and denying the possibility of illogical thought. Roughly put, my strategy is to distinguish two levels: exercises of the understanding and cognitions (wholes of compared and connected representations). Logical mistakes are exercises of the understanding that are at odds with logic's laws. Normativists are right that logic is normative for exercises of the understanding, regardless of whether the exercise accords with logic's laws or not. Normativists are wrong, however, insofar as they treat all exercises as cognitions. Formalists are right about logic's relation to cognitions: illogical thoughts are not cognitions at all. Still, formalists are wrong when they say that a putative thought that violates the laws of logic is not even an exercise of the understanding.

Clinton Tolley, a formalist, raises the tension between the normativist and formalist interpretations. Although he recognizes that Kant often speaks of pure general logic in normative terms, he downplays these discussions by claiming that they only appear in the unpublished logical works (Tolley, 2006, p. 392). This overlooks, however, the normative valence of Kant's consistent published references to pure general logic as a canon (e.g., A53/B77). A canon is "the sum total of the *a priori* principles of the **correct** use of certain cognitive faculties in general" (A796/B824, my emphasis), and pure general logic is the canon of the understanding in regards to what is formal in its use. Melissa Merritt, another formalist, recognizes the importance of Kant's discussions of logic as a canon (Merritt, 2015, p. 482), but argues that a canon is normative even on the formalist reading, because just as someone who violates the rules of chess is not making a chess move, someone who violates the rules of logic is not, say, inferring at all. Nonetheless, according to Merritt, such rules regulate practice, and are thus normative "by ruling things out of bounds, as non-thought and non-chess" (Merritt, 2015, p. 483).

I do not think this kind of resolution can work. The problem is that on Merritt's view, if a

---

[20]I have in mind two kinds of passages. First, there are those where he uses the notions of a contradictory concept, judgment, or cognition, to pick out—and thereby affirm the possibility of—those very representations he then claims are impossible. For example, he says contradictory *concepts* and *judgments* "are nothing" (A150/B189; A291/B348; *MM*, 29:792), or agreement of a *cognition* with itself "makes a cognition possible as cognition" (*WL*, 24:823). Second, his claim that the principle of contradiction is "inviolable" seems to suggest contradictory thoughts are impossible, but when he claims this he also says that "we will, to be sure, always be careful not to act contrary to this" (A152/B191), which seems to suggest they are possible.

candidate representation violates the laws of logic, then it is not a thought. If it is not a thought, then it is not governed by thought's laws, and thus these laws cannot dictate how it ought to be. For Merritt, as for Tolley, "that which fails to accord with logical law simply loses its identity as an exercise of the faculty governed by this law" (Tolley, 2006, p. 385). If this were right, however, then the laws of logic do not govern such representations, and so there is no longer a sense in which they *should* accord with these laws. After all, there is no normative valence when I judge, for example, that an associative transition from the thought of Berlin to the thought of lindens is not an inference, or that the juxtaposition of '*Berlin, lindens*' is not a judgment. My associations and juxtapositions are not *worse* because they are not inferences or judgments. They would only be worse if they were failed judgments or inferences—i.e., failed exercises of the understanding—and we can only make sense of the "bungler" who has made a logical mistake or the "sophist" who is trying to trick others into doing so, if logical errors count as such exercises (A298/B254).[21]

Recognizing this, however, suggests a clear path forward. Thoughts are exercises of a fallible capacity or faculty for thinking. With a fallible capacity, Kant will often use the term for its exercises in both a wide sense that includes failed exercises, and a narrow sense that only includes successful ones. For example, a cognitive capacity (*Erkenntnisvermögen*) is a capacity for cognition of objects. Nonetheless, sometimes Kant will speak of 'false cognition' (*falsche Erkenntnis*), which fails as cognition of an object.[22] Similarly, there can be exercises of the understanding, merely as a faculty for thinking, that fail to accord with its laws. These will be 'thoughts' insofar as they are exercises of this faculty. But 'thought' in a stricter sense will only refer to exercises of this capacity that are in agreement with its laws. Distinguishing these senses of 'thought' looks promising because, on the one hand, the laws of logic will be normative, because every exercise of the faculty ought to accord with its laws, although some of them may not. On the other hand, the laws of logic will be formal criteria, because any exercise that does not accord with these laws will be a mere confusion; it will not be a cognition.

Before turning to develop this suggestion, responding to an objection will clarify my disagree-

---

[21]See here the argument Jessica Leech (2015, §2.1) develops against formalist views in general and a related argument by Kathryn Lindeman (2017) against views that make constitutive features both criterial of kind membership and normative for kind members.

[22](E.g., A58/B83; A376; A709/B737; *WL*, 24:826; 24:832; *DW-L*, 719-20, 724).

ment with the formalist. On Merritt's view of chess, presumably the relevant class of non-chess-moves are still chess piece movements that are attempts to move the piece in a way that will be accepted as a chess move. These kinds of piece movements will be evaluable by the rules of chess. So Merritt will be a disjunctivist about chess piece movements: some will count as moves, and some will be failed moves, because they violate the rules. Analogously for thoughts, Merritt might try to appeal to a relevant class of non-thoughts, for which the norms of thoughts are still authoritative, but which fail to be thoughts. Spelling this out, however, is non-trivial, and I suspect that any way of doing this correctly will collapse into my view. The reason is that the formalist must find a genus which will include the relevant non-thoughts and thoughts, such that the same norms are authoritative for both, because they belong to this genus. Kant gets this genus through appealing to exercises of the understanding. These then divide into the logically successful and the logically unsuccessful. So long as the formalist appeals to *exercises of the understanding* in order to explain how the laws of logic are normative for the relevant non-thoughts, their view will be a version of my view. This is because while they might still reserve the use of the word 'thought' for the logically successful species of this wider genus, this linguistic legislation does not make a philosophically substantive difference to the underlying view.[23]

To clarify these two senses of 'thought,' it will help to reflect on how thoughts serve their function of combining representations into whole cognitions. Another way Kant will often get at this function is by speaking of concepts uniting representations into "one consciousness" or a "unity of consciousness," and of judgments as representing the relationship of various concepts to "one consciousness" (i.e., *einem Bewußtsein/Einheit des Bewußtseins*. E.g., B114; A103-104; B133-134n; *PölL*, 24:577u; *JL*, 9:101). This does not happen when the 'concept' or 'judgment' is contradictory. That is, if two cognitions are contradictory, and so thinking them together violates the laws of logic, then my consciousness cannot be unified in thinking both of them. Take two opposing (*entgegengesetzte*) predicates like *brightness* and *darkness*. I can think of *brightness* and I can think of *darkness*, but when I put (*setze*) these together, *bright darkness*, "I think nothing at all" (*MM*, 29:792; R3720, 17:267). *Bright darkness* lacks the form of a concept because it is not general. The predicates cancel out and nothing can fall under it. Further, their opposition

---

[23]My thinking in this paragraph is especially indebted to a conversation I had with Andrew Stephenson.

prevents their combination. They cannot form "one consciousness" or one whole of compared and connected representations. This is because their combination lacks "the logical form of a concept (of thinking) in general," i.e., unity of consciousness (A239/B298).

This is the key to how Kant did and did not think illogical thought was possible. Usually illogical exercises of the understanding arise through a kind of "logical illusion" where, for example, a dialectical inference imitates "the form of reason" (A296/B353), and tricks the inferrer into thinking it follows a legitimate inference rule: as, for example, if one were to believe someone when she says, "All Moors are men. The respondent is a man. Therefore, he is a Moor" (*WL*, 24:828). In this case, "attentiveness to the logical rule" is all that is required to fix the illusion. This suggests that in general, just as all that is required to remedy a grammatical mistake is the recognition that one has violated the relevant rule, all that is required to remedy a logical error is attention to the appropriate rule. So, for example, all that would then be required to fix the mistake of thinking contradictory predicates in the same subject would be attention to the "principle of determinability," according to which only one of two opposing predicates can be added to any concept (A571/B599).

Nonetheless, insofar as I am not conscious of the mistake, there would seem to be a sense in which I could form a contradictory cognition. I could, for instance, judge that a composite substance must consist of simple parts, and also judge that it must not consist of simple parts, so long as I was not simultaneously aware of the contradiction in judging both (A434/B462, A435/B463).[24] In this case, although there was an act that had the phenomenology of judgment, my concept of a composite substance is not a whole of compared and connected representations. I have exercised my understanding in attempting to combine the predicates, but since these have the form *A* and *non-A*, they cannot be united into one consciousness. I am confused. I think I've done something that I haven't.[25] And it is symptomatic of my confusion that as soon as I become conscious of

---

[24]Another kind of example is if I inadvertently combine contradictory concepts, because this contradiction is hidden. Take *squarecircle*. The concepts combined here do not have the surface form *A* and *non-A*. So it may not be evident to me that *squarecircle* is contradictory. But once I analyze these concepts and spell out their content, I will see they will include features that have this form: *round* and *non-round*, or *cornered* and *non-cornered* (*Prol*, 4:341). And so *squarecircle* will fall apart.

[25]My use of 'confused' here is related to Kant's but different. I'm claiming the thinking subject is confused, but confusion in Kant's technical sense applies to concepts (or intuitions), not thinkers. Still, for Kant, a concept is confused when it contains parts that one is not aware of (distinctly), and so its parts are not well ordered (e.g., *JL*, 9:34-35). In my sense, confusion arises when thinkers have one of these confused 'concepts,' but the parts are logically

the contradiction (together with the principle of determinability), my 'concept' of a composite substance falls apart: I recognize that I have tried to combine two predicates that are opposed.[26]

At this point it might sound as though I intend to endorse a solution of the sort that Tolley attempts. As a formalist, Tolley rejects that logical mistakes are thoughts. Still, he thinks logic's laws get their normative grip on these putative thoughts through "second-order" judgments. On these, the failed putative thoughts are judged to be thoughts. And the laws of logic get their normative grip on the putative thoughts through the falsity of the second-order judgments about them (Tolley, 2006, p. 391). For example, on this view, in the above case when I putatively judge that composite substances both must, and must not, consist of simple parts, I do not actually judge. Still, I do make a false second-order judgment that this putative judgment is a judgment. And it is because I think I've made a judgment when I haven't that the laws of logic are normative.

Tolley's view, then, seems to be a disjunctivist view of the sort I discussed above: the relevant genus is 'putative thought that I judge to be a thought,' while the species are 'thoughts' and 'failed putative thoughts.' So does it collapse into the kind of view I am defending? I don't think so, because Tolley does not take failed putative thoughts to be exercises of the understanding. And for this reason, his view misconstrues the normativity of Kant's logic. The laws of Kant's logic are normative for thinking because they are codifications of the rules by which the faculty for thinking (fallibly) functions, not because we might mistakenly take a representation that is not a product of the (infallible) faculty they govern to be one. And although it is true that I might take myself to have formed a judgment or concept when I haven't, the normativity of logic's laws do not require actually making this reflective judgment. Rather, it is essential to exercises of the understanding that they are better or worse insofar as they accord with these laws or fail to. This is so regardless of whatever further reflective judgments I go on to make. Logic's laws are not normative for first-

---

opposed.

[26]N.B. Here the psychological and the logical (in Kant's sense) should be kept carefully distinct. Applied logic, a branch of psychology, will treat the topic of "attention, its hindrance and consequences" (A54/B79). Thus, if I am right, Kant is committed to the psychological claim that attention to a contradiction in one's thinking (together with awareness of the principle of determinability) will cause the recognition that in thinking them together one has not successfully thought anything at all. Kant would explain this psychological fact, however, with the logical point about the opposition of contradictory predicates: that they cannot be combined in one genuine concept, i.e. form a unity of consciousness. To mistakenly think one has so combined them, one must be confused, one must have fallen prey to a kind of illusion. And while the nature of illusion in general—that it involves mistaking subjective grounds of belief for objective ones—is a matter for pure general logic, specific kinds of illusion, and how they have their sources in prejudice, are topics for applied logic (e.g., *JL*, 9:76, *DW-L*, 24:737*ff*).

order putative thoughts via second-order judgments about them, nor (in the first instance) for second-order judgments about first-order putative thoughts. Rather, its laws are simply normative for thoughts, as exercises of the understanding.

I opened this section by contrasting my interpretation with formalist ones, insofar as these maintain that that which fails to accord with logic loses its identity as an exercise of the understanding. Let me close it by contrasting it with normativist ones. In doing so, however, I do not mean to war monger. There is a sense in which I hope both normativists and formalists will count me as one of their own—after all, I hope to preserve core insights from each. Nonetheless, against normativists, I am claiming that when we have a contradictory cognition, there is no genuine thought for Kant—no whole cognition—because this is not a coherent way in which the understanding can compare and connect representations. This is because in such a case there is no unity of consciousness. The various compared representations are opposed, and not merely in what they are about, but in their form (e.g., *mortal* and *non-mortal*). I can only take these predicates to fit together into one cognition if I am confused—if I have fallen prey to a logical illusion. And in this sense, although I have exercised my understanding, I have not done it successfully: I have not coherently thought these predicates together in a single concept.

6 Logical aliens

Having seen how Kant would account for logical mistakes, let us turn to logical aliens. In this section I will develop an argument from Kant's texts that show he would rule out as merely apparent the possibility of logical aliens from the standpoint of reason, the only possible standpoint. This argument will hinge on how the case of beings with other forms of sensibility differs from the case of the logical alien. It begins from an argument for the *unintelligibility* of logical aliens and argues to this *impossibility*, where this impossibility should be understood absolutely: there is not a (intelligible or unintelligible) possibility being ruled out. Logical aliens are merely nonsense.

Now, this concern with logical aliens can seem quite foreign to Kant. He does not, after all, explicitly discuss the possibility of a being with a discursive intellect that functions according to different laws from our own. So why think he must have a position on the issue at all? The answer

has to do with his commitments about the absolute necessity of logic's laws. If Kant cannot rule out logical aliens then, to paraphrase remarks he makes about the categories, the laws of logic "would lack the **necessity** that is essential to" them (B168). Their necessity would be merely subjective, holding for me (and my kind) but extending no further, and I would only be able to say "that I am so constituted that I cannot think" contrary to these laws, which "is precisely what the skeptic wishes most" (B168). For this reason, by his own lights, Kant cannot hold that the laws of logic provide essential, general, and necessary laws for our understanding and thinking, while still allowing the possibility of another kind of being, whose understanding is constituted differently, so that it thinks according to different laws.

After an initial look at Kant's texts, however, rather than rejecting as merely apparent the possibility of logical aliens, one might be forgiven for concluding that he embraces it. Consider the following:

> But for the peculiarity of our understanding, that it is able to bring about the unity of apperception *a priori* only by means of the categories and only through precisely this kind and number of them, a further ground may be offered just as little as one can be offered for why we have precisely these and no other functions for judgment or for why space and time are the sole forms of our possible intuition. (B145-B146)
>
> Even were they possible, we could still not conceive of and make comprehensible other forms of intuition (than space and time) or other forms of understanding (than the discursive form of thinking, or that of cognition through concepts). (A230/B283)

If we examine these and other passages more closely, however, this initial impression is not borne out.[27] While there is a significant sense in which Kant thinks we must remain agnostic about the possibility of beings with forms of sensible intuition, other than space and time, the same does not hold of thinking beings whose intellect is governed by other logical laws. He considers the possibility both of intellectual non-discursive beings that do not think but intuit, like God, and of discursive beings with forms of sensible intuition other than ours (esp. B139), but he never explicitly considers the case of a discursive understanding who thinks according to other logical laws.

It might have been texts and considerations such as these that led Putnam to claim "the whole point of the Kantian line is that logical necessity neither requires nor can intelligibly possess any

---

[27](See also, e.g., A42/B59; B72; B138-B139; B148; A252; A254/B309-A255/B310; A286/B342-A288/B344).

'explanation.'" (Putnam, 1994, p. 248, 255).[28] The unconditional necessity of the understanding's laws is simply the bedrock where our spade turns. Putnam is on to something. Still, there is more to say about why and how Kant thought such an explanation of logical necessity is impossible. For one, while Kant says something like this is true of the functions for judgment, he elaborates these functions from the act of judgment, as the fundamental act of the understanding, and he is convinced that we can be certain this elaboration is complete and entire (A69/B94; A64/B89). Surely there is something to be gleaned for our topic from investigating what *a priori* justification Kant might have for this claim to completeness, aside from the apparent empirical fact that no one since Aristotle had discovered more of them (*JL*, 9:20; *DW-L*, 24:700).

I am not, however, going to dig into that now. Rather, I want to focus on another passage from the end of the outset of the "Paralogisms," where I take Kant to rule out the possibility of logical aliens, through a more general claim he makes about self-consciousness. But before we turn to this, let's look at a more approachable passage that makes a similar point. Kant says:

> it is obvious that if one wants to represent a thinking being, one must put oneself in its place, and thus substitute one's own subject for the object one wants to consider (which is not the case in any other species of investigation). (A353-A354)

Elaborating on this slightly, when I represent a thinking being, I take the being to differ from myself. Nonetheless, there will be other features of this being that I take it to share with me. And I take Kant to maintain it is in virtue of these features that I can represent other thinking beings. Among these shared features there will be those that I take us to both share merely in virtue of being self-conscious thinking subjects. And I will at least substitute these aspects of my own thinking subject, which I take to be necessary in any thinking subject, when I consider the other thinker.

The passage I want to focus on develops a similar thought. It begins as follows:

> Through this I, or He, or It (the thing), which thinks, nothing further is represented than a transcendental subject of thoughts $= x$, which is recognized only through the thoughts that are its predicates, and about which, in abstraction, we can never have even the least concept; because of which we therefore turn in a constant circle, since we must always already avail ourselves of the representation of it at all times in order to

---

[28]See also (Thompson, 1981, esp. §VI)

judge anything about it; we cannot separate ourselves from this inconvenience, because the consciousness in itself is not even a representation distinguishing a particular object, but rather a form of representation in general, insofar as it is to be called a cognition; for of it alone can I say that through it I think anything. (A346/B404)

One of Kant's points here is that in order to make any judgment about the thinking subject, one must deploy that very subject. After all, I do all of my thinking through this subject, and so to think about it, I must rely on it. For this reason, any time I think about this thinking subject, I must always turn in a kind of circle.

A similar point holds of the understanding: whenever I think about it, I must deploy that faculty. That Kant saw this is clear from some of his descriptions of logic. Logic studies the understanding, with the understanding. Kant will thus describe it as "a self-cognition [*Selbsterkenntniß*] of the understanding and of reason, not as to their faculties in regard to objects, however, but merely as to form" (*JL*, 9:14). He will even claim that "in logic the question is only, *how will the understanding cognize itself*" (*JL*, 9:14)? In this sense, Kant was already aware of what Harry Sheffer would dub the 'logocentric' predicament: "*In order to give an account of logic, we must presuppose and employ logic*" (Sheffer, 1926, p. 228).

Similarly, our attempt to think about logical aliens presupposes and employs the laws that govern our thought. The laws of logic are the essential laws of thinking, governing any thinking whatsoever (both as constitutive *norms* for every exercise of the understanding, and as constitutive *criteria* on cognition). Our only means of making sense of the possibility that the laws of logic could be otherwise is thus an activity that is itself bound by the laws of logic. We would have to *think* about a possible coherent thought that violates our laws of coherent thought, while this thinking of ours is itself bound by our laws of coherence. Thus, we cannot make sense of the laws of logic being otherwise than they are.[29] To borrow Frege's vivid description, trying to make out such a possibility is "like trying to jump out of one's own skin" (Frege, 1893, p. xvii).[30] It only leads to confusion.

Still, why hold that because *our* attempt to think about logical aliens ends in confusion, such

---

[29]Kant makes an argument with a parallel structure at the beginning of the *Critique of Practical Reason* (*KpV* 5:12; 5:12-14; *Fort*, 20:275-276; *Sömmering*, 12:35). He argues that any attempt to show (synthetic) *a priori* judgments are impossible would be self undermining, because this claim would itself have to be a (synthetic) *a priori* judgment. For discussion, see Engstrom (1994, p. 374).

[30]As Ricketts (1985) argues, Wittgenstein (1921, esp. 3.03-3.032) sees this more clearly.

beings are *impossible*? Why not retreat to a weaker agnosticism over their possibility, as Kant does with discursive beings who have other forms of intuition? Returning to the main passage of the "Paralogisms," Kant continues:

> But right at the start it must seem strange that the condition under which I think in general, and which is therefore merely a property of my subject, is at the same time valid for everything that thinks, and that on an empirical seeming proposition we can presume to ground an apodictic and universal judgment, namely that everything that thinks is constructed as the claim of self-consciousness asserts of me. But the cause of this lies in the fact that we must necessarily ascribe to things *a priori* all the properties that constitute the conditions under which alone we think them. Now I cannot have the least representation of a thinking being through an external experience, but only through self-consciousness. Thus such objects are nothing further than the transference of this consciousness of mine to other things, which can be represented as thinking beings only in this way. (A346/B404-A347/B405)

The heart of Kant's argument here is, of course, his claim about necessarily ascribing to things *a priori* all the properties that are conditions of our thinking them. Specifically, for our concerns, I am supposed to ascribe to other thinking subjects the property of being governed by the same laws of thought as I am, since this is the only way that I can think of them. I imagine this 'Copernican move'—a kind of move so often at the crux of Kant's philosophy—will be controversial (e.g., Bxvi-Bxviii, A92/B124-B125).

The best way to see its justification is through examining why Kant does not go in for agnosticism about logical aliens in the way that he does for beings with other forms of sensibility. In both cases we cannot form any positive conception of such a being. So, as with the laws of logic, it might seem that our forms of sensibility constitute the conditions under which alone we can conceive of or imagine other sensible beings, and thus that, by parity of reasoning, we must ascribe to them *a priori* all the properties that accompany having space and time as forms of intuition. Yet, Kant does not do this. Rather, he holds we must remain agnostic about the possibility of such sensible beings, while the apparent possibility of logical aliens is nonsensical.

This difference is due to the different relation that the laws of logic and the forms of intuition stand in to our intellectual and receptive faculties. It is not included within the precise concept of a receptive faculty that its form must be spatial or temporal—although it is essential to ours—whereas it is included in the concept of reason or the understanding that it will be governed by

logic's laws. A being's receptive faculty allows it to be affected by objects that it did not create. It is the means by which these objects are given to it to cognize. Objects are given to us spatio-temporally. It is impossible for us to say whether in some other kind of finite being they might be given in some other way. But if they were, we could not imagine it. Thus, although we cannot form a positive conception of a receptive faculty with another form of sensibility than our own—we cannot say anything positive about how such a receptive faculty would be—we can form a negative conception of it: such a receptivity would *not be* spatio-temporal. And because we can form this negative conception, we are not in a position to suppose that the conditions under which we are given objects are also conditions for every thinking being, even though space and time are the essential forms of our receptivity.[31]

We cannot, however, form a corresponding negative concept of a logical alien, because it is essential to any concept of the understanding or reason that it is subject to logic's laws. As we saw, it is the essential function of the understanding to unite representations into whole cognitions—unities of consciousness. The laws of the understanding are the necessary conditions on doing this (whatever these turn out to be). The understanding of a logical alien would unite representations into whole cognitions according to laws that are different from those governing our understanding. But what kind of whole cognitions would these alien 'cognitions' be, exactly? This negative 'concept' is supposed to be the concept of a whole cognition that violates the necessary conditions on uniting representations into whole cognitions. But that is incoherent. So even a merely negative 'concept' of a logical alien is confused. And there is nothing intelligible to remain agnostic over. Thus, only thinking beings subject to logic's laws are really possible, and Kant's "transference" is justified.

At least, it is justified if reason is the measure of real possibility. In the *Prolegomena* and elsewhere, Kant is explicit that he is presuming reason as his foundation and first principle (*Prol*, 4:274).[32] This presumption is not unmotivated: reason (or the understanding) is the intellectual faculty that any thinking presupposes. The standpoint of reason is the only standpoint. To try to

---

[31]Kant also counsels agnosticism over the possibility of beings whose existence is absolutely necessary, despite their inconceivability for us (A792/B820; Kant's discussion of the ontological proof A592/B620*ff* and *KU*, §76-77).

[32]In this respect, Kant is modeling his own project in metaphysics after pure general logic: both undertake an examination of the form of reason, and presuppose reason as their starting point (e.g., *B*, 10:340).

reason about what lies outside of reason would, according to Kant, lead to "a mode of thinking in which reason moves against itself" (*Prol*, 4:274). This is a mode of thinking he thinks we can take up. Antinomies are where both sides of a proposition appear necessarily true because of the nature of reason itself (A421-422/B449-450). After grappling with these for long enough, he seems to think one might surrender to "skeptical hopelessness" and decide to accept both sides of the contradiction as true (A407/B434). This falsely-won peace would be "the death of healthy philosophy," and a "**euthanasia** of pure reason" (A407/B434; *KpV*, 5:120). But even here, where skepticism has reached its highest pitch, what led to this dark place was a "complete despair as regards satisfaction of reason's most important aims" (*Prol*, 4:274). Thus, what motivates this conclusion, according to Kant, itself has its ultimate source in reason and its laws. So while one cannot refute such a skeptic, because she rejects reasoned argumentation, that she is in a position to reject these laws is itself indicative of her being subject to them. And so rather than being a case of logically alien thought, her thought is merely a case of (badly) logically mistaken thought. There is no genuine other side to the laws of logic. All that lies beyond those laws is confusion, nonsense, and failed exercises of the understanding.

7   The necessity of pure general logic

Now that we have seen how the insights at the heart of normativism and formalism explain how Kant could preserve logical mistakes while rejecting logical aliens, what remains is to explain the necessity of Kant's pure general logic. On the account I've developed, for Kant, logic's laws are both constitutive norms for any exercise of the understanding and formal conditions on coherent thought, that is, cognitions. Attributing these two roles to logic's laws might have seemed to saddle Kant with two potentially incompatible explanations of their necessity, but now we are in a position to see how these two roles work together to explain the distinctively absolute normative authority of logic's laws: how the very insight of the formalist—that logic's laws are formal criteria on coherent thought whose violations are mere confusions—explains the normativist insight that laws of logic are absolute constitutive norms for any exercise of the understanding. To see this, it will be useful to contrast the laws of logic with the rules of chess.

MacFarlane develops the analogy between the rules of chess and the rules of logic through the way the laws of physics provide constitutive norms for thought about the physical world. He argues that someone engaged in the game of thinking about the physical world "*ought not* make judgments that are incompatible" with the laws of physics in the same way as someone playing chess ought not violate the rules of chess (MacFarlane, 2002, p. 36). In both cases, the rules provide constitutive norms because both the thinker and the player are assessable in light of the norms, in virtue of the activity they are engaged in. Similarly, MacFarlane thinks that anyone who is thinking (or really judging) is essentially assessable by the norms of pure general logic. And this is because the laws of logic have a similar standing in relation to thought as the rules of chess or the laws of physics have in relation to playing chess or thinking about the physical world. In all three cases, anyone engaged in the activity is subject to the norms.

Still, while both chess and thinking are constitutively governed by their rules, the nature of chess and its rules are contingent, but the nature of thought and its rules are not. In chess we might make our own house rules. It would be perfectly coherent to decide the game is won by capturing the queen instead of the king. We also might investigate the history of chess and the variations in the rules over the centuries. It is an interesting question for philosophers and historians of chess whether in changing our house rules we change chess or begin playing shmess, and whether there has been one game over the years or many variations going under the same name. But regardless of whether our house rules govern chess or shmess, we can actually engage in this activity and the constitutive norms of this activity are authoritative for our games. By changing the rules, we can change the nature of the activity governed by those rules. This is not possible with logic. Attempting to agree on different logical laws cannot change the nature of thinking. Its laws are always the same laws. The nature of the understanding does not change. Although different particular logics govern thought about different kinds of objects, the laws common to all thinking, as thinking, cannot vary. Unlike chess, the nature of thought, its laws, and the understanding are fixed.

An aspect of this unconditional normativity of logic is at the heart of MacFarlane and Leech's readings. According to MacFarlane:

> We can correctly think about what the world would be like if the laws of physics were different, but not about what it would be like if the laws of logic were different. This

> is the sense in which the norms for thought as such are necessary: it is impossible to think at all, even counterfactually, without being constrained by them. (MacFarlane, 2000, p. 55)

We might imagine a world where the gravitational force is half as strong, just as in chess we might play by different house rules. In imagining this case, the new law and its corollaries would govern the truth or falsity of our claims. This is not so with the laws of thought: we cannot coherently posit a world where the laws of thought are different. Jessica Leech puts this point in terms of rational indubitability: the laws of logic are unconditional in the sense that any rational attempt to doubt them is self-undermining. This is because one has to rely on these very principles in attempting to doubt them (Leech, 2015, §3.1).

Though MacFarlane and Leech are correct that the laws of logic inescapably govern our thinking, this falls short of establishing the laws of logic have the "**necessity** that is essential to" them insofar as it leaves open the possibility that these laws hold only for us and not absolutely, for all finite cognizing beings (B168). The reason that they do not go this far is that they do not appreciate the significance of another difference between logic and chess or physics.

The rules of chess prohibit possible moves by designating them as impermissible. Similarly, the laws of physics rule out ways the world could be by designating them as physically impossible. In both of these cases there is something that is ruled out: some way pieces could have moved, or some way that the world could have been. This is not so with logic. The laws of logic do not divide representations of facts (or states of affairs) into logical and illogical ones. There is not some cognition that the laws of logic rule out. There is no way that thought could have worked, other than the way that it does, and there is no intelligible or unintelligible possibility—no effable or ineffable state of affairs—that lies on the other side of logic's laws. It is this aspect of logical laws that leads the formalist to reject the ontological position and it is this that she is after with her claim about the inviolability of logic's laws. When an exercise of the understanding fails to accord with them, it is a confusion, not some alternative coherent way thought could have been.

The absoluteness of the norms of logic—that they hold not merely unconditionally for us, but for all thinking—stems from this feature of exercises of the understanding: that illogical ones could never be whole cognitions. The understanding is the faculty for cognition, for comparing

and connecting representations into new wholes, and the laws of the understanding are formal conditions on this. Any exercise of the understanding that violates these laws is not cognition, but confusion, and such confusions are never elucidatory. There are no wholes of compared and connected representations, which the laws of logic rule out. In this way, the laws of logic are unlike the laws of physics, psychology, or chess: there is not some state of affairs that they claim does or does not obtain, and there is no intelligible or unintelligible possibility that they rule out. This is why the laws of logic are absolute—why they have the necessity essential to them. And so it is because of the way violations of the laws of logic are not like illegitimate chess moves or physical judgments that contravene physical law, but are confusions, that the laws of logic are absolutely necessary norms, while the rules of chess are contingent.

We can see that Kant would endorse this argument if we reflect on his discussions of truth and on how illogical cognition self-destructs. According to Kant, truth "is the agreement of a cognition with its object" (A58/B82). And logic, as the study of the form of cognition or thinking, "must present criteria of truth" (A59/B84). Agreement with these will not be sufficient "to constitute the material (objective) truth of the cognition," but will be a necessary "negative touchstone of truth," the violation of which makes the agreement of a cognition with its object impossible (A60/B84-85). Now, the *reason why* agreement with the laws of logic is a necessary condition on truth, according to Kant, is that when a thought contradicts the rules of logic, it contradicts "the general rules of thinking and thus contradicts," and even "annihilates," itself (A59/B84; A151/B191; *WL*, 24:823). In this respect, it is because cognition (its unity of consciousness) is annihilated by contradicting the rules of logic that no contradictory cognition can be true, no matter how things stand with its object. We even find Kant claiming that logic presents universal and necessary "principles of all logical assessment of our cognition" because it studies "the general and necessary rules of the understanding," agreement with which is required for a cognition's agreement with itself—that is, for not annihilating itself (A59/B84; A60/B84; R1620, 16:40-41; R1628, 16:44-46). Thus Kant takes the absolute normativity of logic to stem from the way exercises of the understanding that violate logic's laws "are nothing" (A150/B189). So for Kant, unlike the laws of physics and chess, the laws of logic are formal conditions whose violations are self-annihilating nothings. They are never elucidatory because there are no illogical facts. And it is because the laws of logic are formal

conditions on the possibility of cognitions that they have absolute normative authority over the understanding, as the faculty for producing such cognitions.

8   Conclusion

In the foregoing we have seen how Kant would accommodate the possibility of logical mistakes and reject as unintelligible the apparent possibility of logical aliens. And we have seen that the laws of pure general logic are unconditionally binding norms because illogical exercises of the understanding are mere confusions. Here I have prescinded from discussing the details of Kant's logical theory, focusing instead on the nature and status of logic's laws, whatever these turn out to be. Of course, this logical theory, its success, and its completeness are immensely important for Kant's philosophy, and in the wake of Frege, then Tarski, this theory has fallen onto hard times. How does this radical shift affect the considerations I have been presenting? Would we, for example, be logical aliens to Kant, accepting as we seem to different logical laws? And are the *a priori* considerations I have been offering thereby refuted by an empirical fact? No. Just as developments in our theorizing about the physical world do not entail corresponding shifts in the laws of physics, developments in logical theorizing do not entail corresponding shifts in the laws of logic. On a Kantian conception of logic, how could they? The laws of logic are the absolutely necessary laws of thought. We can perhaps make our conception of these more precise, and even recognize that what we once would have taken to be coherent is actually confused, but none of this changes the laws themselves.★

Note on Kant's texts

In citing the *Critique of Pure Reason* I use the standard A and B edition numbering. For other works I use the standard "volume:page number" of the Academy edition of Kant's works, as well as the following abbreviations for specific works:

KpV = Kritik der praktischen Vernunft
KU = Kritik der Urteilskraft
Eberhard = Über eine Entdeckung, nach der alle neue Kritik der reinen Vernunft durch eine ältere entbehrlich gemacht werden soll
Fort = Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnitzens und Wolfs Zeiten in Deutschland gemacht hat?
Sömmering = Aus Sömmering: Über das Organ der Seele
GMS = Grundlegung zur Metaphysik der Sitten
B = Briefe
Anth = Anthropologie in pragmatischer Hinsicht
VT = Von einem neuerdings erhobenen vornehmen Ton in der Philosophie
En-F = Vorlesungen über Enzyklopädie (Friedländer)
MM = Metaphysik Mrongovius
JL = Jäsche Logik
DW-L = Dohna-Wundlacken Logik
WL = Wiener Logik
BusL = Busolt Logik
PölL = Pölitz Logik
PhiL = Philipi Logik
BlomL = Blomberg Logik
R = Reflexionen

Of course, we do not have a work on pure general logic that was authored and published by Kant, so we must rely on Kant's notes in the logic textbook he lectured from (Meier's *Auszug Aus der Vernunftlehre*), notes taken by Kant's students during his logic lectures, and the logic text prepared by Kant's student Jäsche (at Kant's request) for Kant's views on the subject. I will generally corroborate claims from the logical works with multiple sources. When available, I have usually stuck to the Cambridge edition translations of Kant's works, although some translations may be my own.

References

Anderson, L. (2005). Neo-kantianism and the roots of anti-psychologism. *British Journal for the History of Philosophy*, 13(2):287–323.

Boyle, M. (MS). Kant on logic and the laws of the understanding. In Travis, C., editor, *Logical Aliens*. Harvard University Press.

Conant, J. (1992). The search for logically alien thought: Descartes, Kant, Frege, and the *Tractatus*. *Philosophical topics*, 20(1):115–180.

Engstrom, S. (2009). *The Form of Practical Knowledge: a Study of the Categorical Imperative*. Harvard University Press.

Engstrom, S. P. (1994). The transcendental deduction and skepticism. *Journal of the History of Philosophy*, 32(3):359–380.

Field, H. (2015). What is logical validity? In Caret, C. R. and Hjortland, O. T., editors, *Foundations of Logical Consequence*, pages 33–70. Oxford University Press, USA.

Frege, G. (1893). *Grundgesetze der Arithmetik: I-II Band*, volume 1. H. Pohle.

Hanna, R. (2006). Rationality and the ethics of logic. *The Journal of philosophy*, 103(2):67–100.

Kant, I. (1979). *Gesammelte Schriften*. 29 vols. Walter de Gruyter, GmbH, Berlin.

Kant, I. (1996). *Practical Philosophy*. The Cambridge Edition of the Works of Immanuel Kant in Translation. Cambridge University Press, Cambridge.

Kant, I. (1999a). *Correspondence*. The Cambridge Edition of the Works of Immanuel Kant in Translation. Cambridge University Press, Cambridge.

Kant, I. (1999b). *Critique of Pure Reason*. The Cambridge Edition of the Works of Immanuel Kant in Translation. Cambridge University Press, Cambridge.

Kant, I. (2001). *Lectures on Metaphysics*. The Cambridge Edition of the Works of Immanuel Kant in Translation. Cambridge University Press, Cambridge.

Kant, I. (2004). *Lectures on Logic*. The Cambridge Edition of the Works of Immanuel Kant in Translation. Cambridge University Press, Cambridge.

Kant, I. (2010). *Theoretical Philosophy After 1781*. The Cambridge Edition of the Works of Immanuel Kant in Translation. Cambridge University Press, Cambridge.

Leech, J. (2015). Logic and the laws of thought. *Philosophers' Imprint*, 15(12):1–27.

Leech, J. (2017). The normativity of Kant's logical laws. *History of Philosophy Quarterly*.

Lindeman, K. (2017). Constitutivism without thresholds. *The Journal of Ethics and Social Philosophy*.

Longuenesse, B. (2005). *Kant on the Human Standpoint*. Modern European Philosophy. Cambridge University Press.

Lu-Adler, H. (2016). Kant and the normativity of logic. *European Journal of Philosophy*.

MacFarlane, J. (2000). *What Does It Mean to Say That Logic is Formal?* PhD thesis, University of Pittsburgh.

MacFarlane, J. (2002). Frege, Kant, and the logic in logicism. *Philosophical Review*, 111(1):25–65.

Merritt, M. M. (2015). Varieties of reflection in Kant's logic. *British Journal for the History of Philosophy*, 23(3):478–501.

Putnam, H. (1994). Rethinking mathematical necessity. In Conant, J., editor, *Words and Life*, pages 245–264. Harvard University Press.

Ricketts, T. (1985). Frege, the *Tractatus*, and the logocentric predicament. *Noûs*, 19(1):3–15.

Ricketts, T. (2010). Concepts, objects, and the context principle. In Potter, R. ., editor, *The Cambridge Companion to Frege*, pages 149–219. Cambridge.

Sheffer, H. (1926). Review of *Principia Mathematica*, volume I, second edition. *Isis: A Journal of the History of Science*, 8:226–231.

Stang, N. F. (2014). Kant, Bolzano, and the formality of logic. In *The New Anti-Kant*, pages 192–234. Palgrave Macmillan.

Thompson, M. (1981). On a priori truth. *The Journal of Philosophy*, 78(8):458–482.

Tolley, C. (2006). Kant on the nature of logical laws. *Philosophical Topics*, 34(1/2):371–407.

Travis, C. (MS). How logic speaks. Unpublished manuscript. King's College, London, UK.

Wittgenstein, L. (1921). *Logisch-philosophische Abhandlung: kritische Edition*. Suhrkamp-Taschenbuch Wissenschaft. Suhrkamp, 1998 edition.