Wittgenstein’s wayward student:
the unauthorized autobiography

Annotated by Curtis Franks

Note: Our hero has been confined until now to the space allotted him by other writers who only want to bounce their own ideas off of him because of their interest in the resulting trajectories. Quite reasonably he feels that he’s never been given a fair shake. His valiant efforts to understand his critics turn up points where their arguments are unconvincing or altogether absent. He delights in these discoveries, but his readers will see them instead as clues to the sort of positions his critics adopt. When he meets another idealized character in Penelope Maddy’s Second Philosopher, he is intrigued by the difference in the treatment he receives. As with his other critics, her arguments appear to him least convincing—when they appear at all—just when the stakes are highest. But this impasse pervades the full range of her positions on everything from elementary logic to the higher infinite, and he doesn’t know what to make of the fact that he cannot seem to get his foot in the door.

1. First things

I think the first thing you ought to know about me, if you’re going to appreciate much of what I have to say later on about my main scientific and philosophical agenda—the nature of logic and mathematics as it happens to be—is that I’m a lot older than I’m often made out to be. You know me as Wittgenstein’s wayward student—not because I ever was a student of the man, and I suppose if I had been I would by now at least be on a ways in years—but because in his *gutes Buch* I play the role of a student who displays, systematically and with impunity, certain “deviant” reactions to instruction, training, and rules, and in particular to the rule “add 2,” no matter how many correctives his teacher issues.

When you hear about such a student, it seems inappropriate just to describe him as poor or lagging. I strike you, no doubt, as fantastically confused, as displaying the sort of behavior that just couldn’t arise in any ordinary course of instruction. And the point Wittgenstein wants

---

1The device of an unauthorized autobiography has been used at least since Larry Rivers’s *What did I do?* and recurs in work by Ray Davies, Tommy Chong, Lemony Snicket, and others—the present work differing from these by actually being unauthorized. Wittgenstein’s wayward student is the persona found in Wittgenstein’s *Philosophical Investigations* and *Remarks on the Foundations of Mathematics*, but the particular label is due to Maddy. Earlier (1985) Goldfarb had used “wayward child,” which the author obviously finds misleading.

2With boundless gratitude to Penelope Maddy, whose vision and imagination are exceeded only by the impact of her patience with her own wayward student.

3*Philosophische Untersuchungen*, cited in text by the author as *Investigations*
you to think about, the point to which we will return later, is why I strike you this way. What you don’t think much about, I trust, what you just take for granted, is that I am a child, and a pretty young one at that. Who but a child, even in such a preposterous scene as Wittgenstein’s, is being taught basic arithmetical operations, much less failing to catch on right away?

The situation reminds me of those Renaissance paintings of Abraham poised to sacrifice his young son. The rabbinical commentators tell us that the “lad” was a full thirty-seven years old, notwithstanding the Bible’s narrative depiction of a passive figure lying unawares atop the altar, naively asking his father if he’d forgotten the sacrificial ram. Similarly, you should know, I found my way into Wittgenstein’s book after more years of writers conjuring me into their work than I care to think about—ordinarily at least, although I find myself now compelled to recount the whole sorry history of these cameo appearances. Suffice it to say, for now, that my role in the *Investigations* was hardly my first, and, far from being a child, I was by that time already quite old.

The first thing you need to know about Wittgenstein, meanwhile, is that he talks to himself. I mean throughout the *Investigations* one finds extended passages comprised of records of voices debating some delicate topic or another—one voice defending a philosophical position about meaning, another trying to debunk it, sometimes even what seems to be a third voice butting in to clear up a misunderstanding between the first two or to tell the debunker to cool it off—and it’s tempting, even natural, to think that somewhere in that stew is the voice of Wittgenstein himself, pitted against his foes.

But then one comes to a remark like, “How does the philosophical problem about mental processes and states and about behaviorism arise?” (§308) and it becomes clear that the whole chorus of voices that preceded it was a display of how philosophical thinking creeps up on us and leads us to think that we have to fit ourselves into one side of a debate: “Our thinking plays us a queer trick. We want . . . to say ‘Either such an image is in his mind, or it is not; there is no third possibility!’” (§352). Well, Wittgenstein proposes, there may be nowhere to run in this arena other than into one or the other corner, but don’t mistake the arena for reality. And pay attention to how you found your way in!

“For this,” Wittgenstein declares, “is what disputes between [for example] Idealists, Solipsists, and Realists look like” (§402). So arises a second temptation: to take most of the voices in the *Investigations* as caricatures of the various philosophical schools that Wittgenstein is mocking. But this can’t be right, either. Because, you see, this litany of voices just won’t stop. Just when you think he’s smothered them for good, Wittgenstein turns his head, asks another innocent question, and before you know it gets drawn into another arena, or sometimes, through a different entrance, back into the same one he just left, or thought he’d left.

And the point of all this repetitiveness and lack of resolution is supposed to be that
Wittgenstein can’t quite close the deal. The voices are his own. Or perhaps they are the personification of the spirit of the times in which we live. And he wants to show us how insidious they are, how tricky it is to escape from them, and how there is no argumentative strategy that allows you to shut them down and move on. Instead, he says, “the results” he is aiming for are “the dents that the understanding has got by running its head up against the limits of language” (§119). He records himself talking to himself in order to show us how easily, perhaps inevitably, and unwittingly we slip into philosophical messes, and in order to bang up his own habits of thinking in his attempts to find his way out of them. Again, these dents aren’t the cost of what Wittgenstein is up to. They are the point.

Anyway, it’s within that project that you find me, summoned from the annals of history, seated in some surreal classroom, unable to take to the instruction that some exasperated teacher is offering. And I tell you, it’s more than a little uncomfortable to star in someone else’s identity crisis.

2. No laughing matter

Now, one obvious consequence of the way I am situated in Wittgenstein’s discussion is that I don’t get much of a say about anything. And I don’t just mean that I don’t get many turns to speak, though in fact I do get but a few—I get no chance at all to carry on with anyone other than this teacher. Some other voices opine about my lot. One suggests that the instruction I’ve been issued is ambiguous, equally well interpreted as I have done and as the teacher intended—that some other more explicit instruction is needed. Another disagrees, claiming that no amount of instruction will do, that no statement or formula or even belabored training session could ever capture the teacher’s full intention, which has infinite applications. And so on and so on.

But what you can’t find anywhere is Wittgenstein just saying outright what I’m doing wrong, or where the teacher has gone wrong, or whatever. And this, it seems, is because he’s more interested in showcasing the assembly of voices and exposing the frustrating cacophony that ensues than in putting forward his own theory of meaning or account of what it is to understand a rule. He is, in a phrase, more interested in playing around at my expense than in any serious philosophical work.

It’s a point I can’t overemphasize. In Lewis Carroll’s treatment of the issue, he casts me as a tortoise—*a tortoise!_*—and, instead of speaking directly to me, has me in conversation with a particularly feeble Achilles. I’m sure it’s a rollicking good time for the right audience, but you’ll excuse me if I can’t appreciate such joking at my own expense. Fair to say that if

---

*I do not know why the author renders “Beulen” here as “dents” rather than “bumps.” Elsewhere he follows Anscombe’s translation.*
there is a serious topic being discussed at all, it is being treated lightly, rather than with the
care and sobriety that real philosophical work calls for.

To give you a full sense of the predicament, consider Norman Malcom’s report that
“Wittgenstein once said that a serious and good philosophical work could be written that
would consist entirely of jokes (without being facetious)” (2001, pp. 27–28). Now try to
imagine making headway on, say, the problem of reference with a joke. I have my suspicions
about this, but let us even grant that a comedic genius like Shakespeare could pull it off. Is
there any evidence that Wittgenstein could?

Speaking of Shakespeare, let me take this opportunity to weigh in on the great authorship
controversy (you will recall that I have been around for quite a long time) and report first-
hand that the man did not, after all, write all those wonderful plays. In fact, someone else also
named William Shakespeare did.

But to return to the issue at hand, think about the elaborate layers separating Wittgenstein,
or Carroll, or any of my other detainers from me, to say nothing of the distance between me and
their readers. Apparently direct confrontation would spoil the joke, so these writers keep their
distance—but at what cost to philosophy? Ordinarily one thinks that the task of philosophers
is to interpret such phenomenon as they are writing about, whereas what these writers do is
better described as passing the buck. If anything, they complicate the matter, so that their
readers are left to debate just what point they themselves are trying to make.

For this reason, it seems to me, one cannot assume without further ado from the similarity
of my role in Carroll’s and in Wittgenstein’s writing, or in anyone else’s for that matter, that
one and the same idea is being put forward time and time again. It’s often unclear to me
whether any idea about anything at all is on offer.

3. Naturalism

When I appear in the Investigations, I’ve already been taught various things about basic
arithmetic. To my teacher’s satisfaction, I’ve demonstrated reasonable mastery of the material.
Then I’m asked to perform some skip-counting with big numbers, 882, 884, 886, 888, 890,
and this all goes pretty well. When I get to 998, I continue in the same way, 1000, 1004, but
the teacher objects that I should instead have followed 1000 up with 1002.

Naturally, I ask why. He replies that I was meant to add 2. I insist that I had done just
that. The teacher says that I’d added 4, whereas I was supposed to continue on in the same
way. But so far as I have been able to understand, I had added 2, just as I had up to that point.

Here’s where those voices chime in. One remarks that I’ve somehow mistaken the in-
struction “add 2” as “add 2 up until the number 1000 and add 4 thereafter”—something that
all the training and testing I’d been through couldn’t reveal until this point. To clear this up, another suggests, I ought to be told that adding 2 is no different among 4-digit numbers as it is among 1, 2, and 3-digit numbers. Someone else, rightly, observes that I don’t take myself to have done anything different. I didn’t think that “add 2” meant “add 2 up until the number 1000 and add 4 thereafter,” but “add 2 no matter what number you’re starting with.” And that’s what I thought I was doing.

This is when my favorite voice calls out: “But isn’t the same at least the same?” (§215).

And I suppose there are a number of different things you could ask about this exchange, but the thing I think we ought to focus on is what exactly a proper understanding of a rule like “add 2” amounts to. For apparently I have misunderstood that rule. My teacher expected me to know, based on my ability to perform in accordance with it as expected for so long, that when starting with 1000 I was supposed to proceed in a way that strikes me as entirely different. So even though it is expected that this amount of training would suffice for me to see that 1002, and not 1004, is the right way to continue, it didn’t. What does?

Wittgenstein’s readers have suggested three different types of answer to this question. The first type of response seeks the thing—an algebraic expression that describes an entire infinite sequence, a computing machine, a mental image, the teacher’s intention—that in addition to the training I received would have prevented any misunderstanding. The Investigations are pretty widely read as emphatic about the inadequacy of this sort of answer.

So a second type of answer is to grant that no amount of training, no mental image or intention, and no formula or machine of any sort, could ever suffice to guarantee understanding. Because there are always multiple ways to move consistently beyond any given data, there’s also always a question about what makes one of those ways correct. And—this is the rub—nothing does. According to this answer, we have to get by with community agreement of some sort—i.e., grounds for taking our responses to be right and others’ wrong are lacking, but this is of no practical concern in ordinary circumstances because beings like us have other ways of reaching consensus. This “skeptical solution,” as it has been called by Saul Kripke and others, following Hume, sees me in a soft light. I’m not mistaken, just different.

Then there’s the naturalistic answer. According to the naturalist, the ambitious seeker, chasing down the elemental thing that understanding consists of, and the skeptic, denying that there is any such thing, share in a common confusion. Wittgenstein called it a bewitchment. These philosophers both start out looking for an answer to the question, “What makes 1002 the correct response to the instruction ‘add 2’ when one is starting at 1000?” One of them is sure that something, perhaps a self-interpreting fundamental rule or state of mind, will be found that fills this role. The other denies that any such thing could, because 1002 is in fact no more correct than 1004. These attitudes seem to be poles apart. What could their commonality be?
Our seeker and our skeptic both demand that something outside of our ordinary practice of issuing and following instructions somehow ground that practice in order for there to be any sense in saying that some applications of a rule are correct. Well, if you think this demand is a fool’s errand, the person who keeps searching doesn’t look an awful lot sillier than the person who throws up his hands and says “Oh well, I guess we’ll have to get by without grounds, somehow or another.”

And that’s the naturalist’s read of the terrain. The seeker’s various theories about what makes 1002 the correct response all fail, but this doesn’t mean that we lack grounds for how we respond. “The . . . naturalist’s point,” in P. F. Strawson’s words, “is that there can only be a lack where there is a need” (1985, p. 41). What makes 1002 the correct response to the instruction “add 2” when one is starting at 1000 is just that you are starting at 1000 and asked to add 2, and “there is, philosophically speaking, nothing behind all this, and no good need for anything beyond or behind it all to constitute a philosophical explanation of it” (pp. 77–78). To look beyond is to lose your grip on what following a rule is.

Warren Goldfarb memorialized this attitude. The skeptic, he wrote, concludes that “nothing determines how to go on with the series” from our inability to point to any ground for the correctness of an application of a rule. “Yet the scenario doesn’t show anything like that,” he replied. “[H]ow to go on is certainly determined, in the sense that if [the wayward student] is to continue the series we asked him to continue, then he must go on to write 1002, 1004, 1006.” Instead, Goldfarb elaborated, “the conclusion . . . is that he is incapable of understanding what we mean” (1985, p. 485). Recall my description of the skeptic’s line of thought above. I said that there is always a question about which among multiple ways to go on is correct. According to Goldfarb, that very idea is a step away from reality into the philosophical arena, “for in ordinary life there isn’t always a question” (p. 486).

Well, where does that leave me? If you tell me there are grounds for the teacher’s expectations, I’ll ask what they are. If you tell me there are no grounds, that I’m just taking to training differently than other students, because I’m not part of their community or something, I can’t be too surprised. I didn’t exactly wind up in this classroom in the most ordinary way to begin with. But how am I supposed to take the news that I’m just wrong and that there’s nothing more to say about why I’m wrong beyond just repeating the same instruction? Isn’t that just the description of the difficulty that we wanted to resolve? “OK, now, you’re up to 1000. Add 2 again. Oops, you’ve got 1004 there, are you sure?’”—“Yes, that’s what I meant to write.”—“Well, you’re wrong. The answer is 1002.”—“1002? Why is that correct?” “Because you were to add 2, and you were up to 1000.” . . .

4. Turtles all the way down
I believe that Peter Winch was the first person to identify me as the tortoise in Lewis Carroll’s (1895) essay about logical inference. I mean, I’d already made a big splash in this role, and as Winch himself noted (1958, p. 57), the “moral” of the parable that I’d been drawn into there, “that the actual process of drawing an inference, which is after all at the heart of logic, is something which cannot be represented as a logical formula,” is a little hard to miss. Winch apologized for “boring” his readers by pointing it out. But I’m not aware of anyone before Winch seeing through my disguise enough to connect this essay with Wittgenstein’s thought.

So what is the connection?

In Carroll’s essay, I’m pondering with Achilles a few propositions labeled A, B, and Z, allegedly so related that Z follows logically from A and B.\textsuperscript{5} Achilles and I agree that even granting the truth of A and B, it would be rash to infer Z without further ado, that a perfectly rational person could assent to A and B and, for all that, deny Z. So I inform Achilles that I’ll play the part of a perfectly rational person for a change, and I challenge him to get me to see my way to the truth of Z by logic alone.

Achilles begins by asking me to consider the conditional expression, “If A and B, then Z.” This strikes me as obviously true, quite independently of my commitment to A and B, so I happily assent. Achilles expects this to prepare me to see my way to the truth of Z, by inferring it from A, B, and “If A and B, then Z.” But this, too, strikes me as too hasty. Pressed about my reluctance, I point out that Z doesn’t seem to follow straightaway from A, B, and “If A and B, then Z” any more than it does from A and B alone. Surely, though, Achilles insists, I would agree that if A, B, and “If A and B, then Z” are all true, so too must Z be. And I assure him that, now that he has drawn my attention to it, I do indeed. Somehow Achilles expects me at that point to be prepared to infer Z form A, B, and “If A and B, then Z,” but I remind him that we had just agreed that the inference isn’t licensed without the extra premise, “If A, B, and ‘If A and B, then Z,’ then Z.”

The narrator of the story—the person who lets this discussion play out without ever weighing in himself, and who I never get to talk to—departs at this point and leaves the reader to sort out on her own what to think.

Now, the first thing to think is that the regress that arises in Carroll’s story is very similar to one that can arise in Wittgenstein’s. Achilles’s search for the missing item that will carry the inference home only leads to a situation like the one he began in. Barry Stroud saw a perfect analogy in the search for the thing that will secure proper understanding. He focuses on the proposal that mental representations will fill the role:

\textsuperscript{5}Carroll’s sentences are geometric and related according to the inference pattern \textit{modus ponens}, but as the author does not dwell on their contents and later considers in this vein the inference pattern known as disjunctive syllogism, it is appropriate to think of A as “It’s either red or it’s green,” B as “It’s not red,” and Z as “It’s green.”
Even if “representations” were little plastic flash-cards and “the mind” was a special little pocket they fit into, and [understanding] was defined as a card’s being in that pocket, we could still not explain someone’s thinking that \( p \) as just a matter of there being a card with “\( p \)” on it in that person’s special pocket. . . . “Representations” alone are not enough because, at the very least, the person also has to understand what is on his card, he has to “grasp” the “content” that is represented there. . . . But then those graspings or understandings are themselves presumably psychological phenomena [. . . and] to say that they too consist of nothing more than the presence before the mind of “objects” or “representations” would lead to a regress. (1991, p. 131)

But what interest me more are the disanalogies between my roles as the tortoise and as the student. Just consider the available reactions to the tortoise regress. Is a skeptical response even possible?

As a wayward student I’m supposed to elicit the thought that, if there’s nothing that can be said or given to me to ensure that I will fasten onto the right understanding of a rule, then there’s nothing you or my teacher or anyone else can point to as a reason to know that they’ve understood it correctly, either. The skeptic concludes that there just is no fact of the matter about whether I or my teacher or anyone else understand correctly—even denies that there is something to understand correctly—and, in the accounts by Dummett, Kripke, Bloor, and others, follows this up with a theory about how one (i.e., everyone but me) gets by in the absence of such facts of understanding. The naturalistic response is supposed, then, to be hard-won against the pull of this sort of skepticism.

Does this line of thought make sense in the case of logical inference? It doesn’t seem at all right to conclude from my conversation with Achilles that there just is no fact of the matter about whether \( Z \) follows from \( A \) and \( B \), moreover to suggest that it is a matter of contingent social practice that people infer \( Z \) from \( A \) and \( B \), that they could just as well reason otherwise. And in fact, no one suggests anything like this.

Instead, the naturalistic response just seems immediate. Winch spent no time considering whether \( A \) and \( B \) might not fully justify \( Z \) in the absence of some conditional statement or another, because he thought that Carroll’s joke is more about me raising the question than with the defect of any of Achilles’s particular attempts to answer it: “To insist on any further justification is not to be extra cautious; it is to display a misunderstanding of what inference is” (p. 57). Similarly, Stroud observed that “it is difficult to see how the tortoise could understand and accept \( A \) and \( B \) without accepting \( Z \)” (1979, p. 44). One is reminded of Goldfarb’s idea that “the correct conclusion is” not that no one means anything at all determinate, but simply “that he is incapable of understanding what we mean.”

Why does one of these scenarios invite a skeptical response and the other not do so?
Interestingly, commentators like Winch and Stroud who see the connection between them do not explore this disanalogy. In “Inference, Belief, and Understanding,” for example, Stroud summed up his account with comments that would apply quite broadly to the general point about understanding and applying a rule, with no mention of any extra wrinkle added by the fact that his route to this conclusion began with considerations specifically about a rule of logical inference: “it cannot be a condition of explaining how a person comes to understand a word or sentence in one way rather than another that there should always be something he goes by or relies on, in the sense of something he can be said to know or believe, to determine that the expression is to be applied this way rather than that” (1979, p. 33).

Similarly, in “Wittgenstein and Logical Necessity,” despite the title’s suggestion to the contrary, Stroud’s analysis does not appear to rely on anything particular about logic. He began just by observing that

Part of human behaviour consists of calculating sums, distances, quantities, of making inferences, drawing conclusions, and so forth. . . . There are various facts which make [such behavior] possible . . . . For example, our memories are generally good enough for us not to take numbers twice in counting up to 12, and not to leave any out . . . and so on. The inhabitants of the earth might have lacked these and other simple abilities, and if so, there would be no such thing as calculating at all. In that way the possibility of calculating depends on such contingent facts. (1965, p. 12)

This set up his naturalistic strategy to dehumanize me: throw the human tendency to “take ‘1002, 1004, . . . ’ to be going on in the same way as putting down ‘996, 998, 1000, . . . ’” into the same pile of contingent facts responsible for how they calculate. He elaborated:

It is a fact that we naturally go on in this way, but people might not have done so. Since they might naturally have followed the rule in a different way, our rules alone do not . . . guarantee that that they will not be taken or understood in deviant ways [. . . because] a rule is not something which stands apart from our understanding of it . . . . How we naturally understand and follow the rule determines which applications of it are correct, and the way the rule is followed will depend in part on what we take to be “going on in the same way.” (p. 12)

That’s the naturalistic response, all right. But doesn’t it seem extraordinary to treat “making inferences” and “drawing conclusions” as just more cases of applying a rule? Stroud said that a being such as me “would not be fully intelligible” to humans, that I “would ultimately be unfathomable” (1965, p. 13). My inhumanity as a student boils down to the fact that I can’t follow the rule correctly. If humans were different in just the right way in how they took the instruction “add 2,” the idea goes, then I could understand them, and they me. But isn’t my error as a tortoise somehow deeper than this? Could humans have naturally followed the
rules of logic in a different way and still be reasoning correctly? If not, then it is plain why as a tortoise I do not elicit a skeptical response, hence why the naturalistic response is not hard-won.

Penelope Maddy seems to be the first person to observe this elementary distinction. She located a discussion in Wittgenstein’s *Remarks on the Foundations of Mathematics* where, she says “we find the interlocutor lodging a line of protest that . . . doesn’t appear in the *Investigations*” (2014, p. 77). The discussion there involves someone drawing inferences other than the ones humans customarily draw rather than someone just finding himself, like me, lacking sufficient grounds to draw any inference at all, but the point is the same: “But isn’t there a truth corresponding to logical inference? Isn’t it true that this follows from that?” Wittgenstein asked. “What should happen if we made a different inference—how should we get in conflict with the truth?” (1933–44, §5).

To illustrate, Maddy considers someone adopting the inference patterns associated with “tonk,” a connective introduced just like the disjunction and eliminated just like the conjunction. This person and the wayward student both “come into conflict with the truth,” Maddy says, but in different ways. The student continues skip counting to 1004 “in the context of” the human practice of following the rule, “where it’s incorrect.” The deviant logician, on the other hand, is “following the tonk rule correctly; it’s just a bad rule” (p. 78).

We’ve seen how the naturalist tries to avoid dealing with me. When I demand an explanation for what I’ve done wrong by writing down 1004, this figure just points back to the fact that I was at 1000 and instructed to add 2, refusing to acknowledge any deeper reason why 1002 rather than 1004 is correct (insisting, in fact, that any such acknowledgement is a step into the philosophical arena leading to regress or skepticism). My demand for an explanation for what I’ve done wrong by not finding my way to ‘It’s green’ from “Either it’s red or it’s green, and it’s not red” could be met with a like-minded response: just repeat the premises, maybe with some emphasis or exasperation. This appears to be both Stroud’s and Winch’s attitude, but it is even less satisfying in this case than in the case of arbitrary rule-following.

Maddy offers a more direct engagement. Calling some rules good and others bad acknowledges something making them so. Couldn’t an explanation of my alleged error point to this something?

5. Science

Decades before I appeared in Carroll’s essay, I snuck my way into the margins of Bernard Bolzano’s *Wissenschaftslehre*, “tempting” him to count the rule according to which a valid inference is drawn among the logical grounds of that inference’s conclusion. I have always marveled at Bolzano’s attitude about this. Rather than claim that my suggestion is inhuman
or that it somehow reveals that I harbor a deep misunderstanding of what logical inference is, he wrote that the ensuing regress (together with what I can only imagine is a theological assumption that the grounds for any truth have to be finite in number) “forces” him to accept that A and B (returning to our example) are the complete grounds of the truth of Z (§199).6

I think the reason behind Bolzano’s reluctance to acknowledge that nothing can be said in favor of Z over and above A and B is that doing so seems to leave unanswered any question about what makes this particular inference correct. That’s why it was a surprise to him to learn that A and B are the entire ground for Z. Of course, I didn’t mean to tip Bolzano off to any such discovery. I meant for him to realize that even if “It’s green” follows from “It’s either red or it’s green” and “It’s not red,” the reason it does must have something to do with the fact that in our world, when a disjunction is true and one of its disjuncts is false, the other disjunct must be true. That’s what Bolzano called the “temptation” to think that “It’s either red or it’s green” and “It’s not red” cannot be the full grounds of “It’s green.” It’s what I call plain sense.

How can a card-carrying naturalist deal with this question?

After rehearsing what he called “perhaps the best skepticism-rebutting argument in favor of the existence of body,” namely, that the hypothesis of an external world is part of the most plausible available explanation of our memory and experience, Strawson wrote:

We accept or believe the scientific theories (when we do) just because we believe they supply the best available explanations of the phenomena they deal with. That is our reason for accepting them. But no one accepts the existence of the physical world because it supplies the best available explanation [of anything]. That is no one’s reason for accepting it. Anyone who claimed it was his reason would be pretending. . . . Similarly, the best argument against other-minds skepticism is, probably, that, given the non-uniqueness of one’s physical constitution . . . it is in the highest degree improbable that . . . . But, again, this is no one’s reason for believing in the existence of other minds, of other people . . . . We simply react to others as to other people. (pp. 20–21)

A naturalistic defense of ordinary logical inference in this vein could go like this: Perhaps we infer “It’s green” from “It’s either red or it’s green” and “It’s not red” because our habits of reasoning have been shaped by features of the world in which our species evolved and in which we as individuals matured. So the reason it’s green, whenever it’s either red or green and it’s not red, is that the world we navigate is one with certain regularities and the like. Those features are, perhaps, both the reason the inference is correct and the reason why it’s

---

6Neither Carroll’s original geometrical example nor our substitute instance of disjunctive syllogism accurately illustrate Bolzano’s program. According to Bolzano, the ground/consequent relation is anti-symmetric and analytic, so the grounds of compound truths are always simpler than those truths. The author is apparently unaware of this central feature of Bolzano’s logical theory.
the inference we draw. But it is no one’s reason for inferring as they do. The reason each of us infer that it’s green is just that it’s either red or green and it’s not red.

Bolzano, though, was not after agents’ reasons for inferring as they do, but ontological grounds for those inferences’ correctness. Of course an agent need not be consciously aware of and appeal to such ontological grounds to be justified in her inferences. But Bolzano wasn’t interested in an agent’s justification. He was interested in the actual reasons for a claim’s truth. His retreat to what looks like the naturalist’s position—the conviction that “It’s either red or it’s green” and “It’s not red” are all the reasons there are that “It’s green”—is based on a realization that what seems like an additional essential fact, the validity of the conditional “If A and B, then Z,” cannot play any role.

Maddy’s arrival on this scene could hardly be more eventful. But to make sense of her position, I first need to return to the ditch she dug between the general problem of rule-following and the special problem of the correctness of logical rules. Her first unexpected move is to fill it back in.

Taking up the question, “what makes the tonk rule bad?” that she so carefully teased apart from the question, “what makes ‘1004’ the wrong answer?” Maddy says, “I think Wittgenstein’s answer to this question marshals the same considerations as his answer to the [other] question—which is why the two tend to blend together.” We already noted Maddy’s observation that, whereas it makes sense to say that 1004 is the wrong answer to the instruction “add 2” when you’re starting at 1000, it makes no sense to say that “add 2” is the correct rule. Still, she says, what makes 1004 the wrong response to the instruction is that the instruction “is embedded in a practice that requires 1002,” and it is perfectly legitimate to ask what makes “add 2” a good rule in that practice:

The practice of elementary arithmetic is as it is for a range of reasons, among them that we have certain interests and goals in which counting and the arithmetical operations are important and that we find “+2” than a rule that changes to what we call “+4” after 1000. . . . There might be nothing particularly disadvantageous about employing an arithmetical rule that changes from out “+2” to our “+4” after 1000, but a rule that calls from random choices or coin flips at each point would not serve the purposes of the practice of arithmetic. (2014, p. 78)

“Add 2” is not the correct rule in the sense that it is the only good one, but it is a good one all the same, one of the rules that facilitates one of our important human activities.

The same sort of considerations are what make the tonk rule bad: “There would be nothing particularly disadvantageous about using an exclusive ‘or’ in place of our inclusive

---

7This distinction between agents’ justification and actual reasons underlies Bolzano’s emphatic distinction between the grounding relationship under discussion here and the relationship of mere derivability.
‘or,’ but using ‘tonk’ defeats the purpose of inferring.” That purpose: “to know about the world around us” and to have “a reliable method of moving from truths to truths.” Summing up: “Both correct application of existing rules and the choice between possible rules rest on . . . our interests, our nature, and the world’s regularities” (2014, p. 78).

How far we’ve gone while standing in one place! Maddy’s inquiry began with her observation that logical inference presents a set of considerations absent from the general case of rule-following. Commentators like Winch and Stroud who missed this distinction suggested that giving in to the temptation to offer an explanation of the correctness of Z that cuts deeper than referral to A and B reveals a misunderstanding of the nature of inference. Spotting it led Maddy, like Bolzano, to find sensible an investigation of what makes some inferential practices good. But that investigation revealed that rules of inference do not in fact differ substantially from other rules. “Our interests, our nature, and the world’s regularities” together license some inferences and not others, just as they, in different proportion, explain what sorts of rules comprise our counting and calculating practices. The key distinction prompting Maddy’s investigation is exposed by that very investigation as an illusion.

What has not proved illusory, according to Maddy, is the fact that inquiry into the nature of logical truth is both coherent and possible:

In the logic case, if someone asks why I’m justified in concluding that it’s green, given that it’s either red or green and it’s not red, there are things I can say—“look here, there are only two possibilities, don’t you see? and one of them is ruled out”—but in the end, if my conversation partner doesn’t get the point, I have nothing more to offer. (2014, p. 102)

So far Maddy is perfectly in line with the naturalist we’ve become familiar with. She agrees with Strawson that providing reasons outside these ordinary ones would be both ineffective and insincere. It would be ineffective because no one who balks at the inference in the first place can be expected to come around after hearing an account of how or why we came to adopt it. It would be insincere because they’re not, after all, our reasons for concluding that it’s green. “But the point at issue now,” she continues, “is what’s open to the philosopher, to the observer of, the commentator on, these various behaviors,” and here she feels herself parting ways with the naturalist we’ve depicted:

Questions like “if it’s red or green, and it’s not red, why must it be green?” or more generally “what is the nature of logical truth?”], traditionally understood as “philosophical,” would remain open after Wittgenstein’s therapeutic efforts were complete, and . . . attempts to answer them would remain legitimate. (2014, p. 110)

But it’s wrong to construe this parting of ways as a flight from naturalism. Following
Strawson, it seems instead appropriate to describe Maddy as “an exponent of a more thoroughgoing naturalism,” someone who “could accept the question, What causes induce us to believe in the existence of body?” or What causes induce us to infer according to the pattern known as disjunctive syllogism? “as one we may well ask, as one that can be referred to empirical psychology, to the study of infantile development” (1985, p. 12).

Here is a naturalist who keeps old philosophical questions alive by transforming them, sometimes, it seems to me, mutilating them nearly beyond the point of recognition. Someone asks, “Why do you suppose bread will nourish you?” You answer, “Because it has always done so in the past,” prompting the obvious follow-up, “But why suppose what has worked in the past will do so again?” The skeptical response is to concede that you have no grounds for this supposition. A Strawsonian response, by contrast, maintains that “the correct way with the professional skeptical doubt” of this sort “is not to attempt to rebut it with argument, but to point out that it is idle, unreal, a pretense; and then the rebutting arguments will appear as equally idle; the reasons produced in those arguments to justify induction . . . are not, and do not become, our reasons for these beliefs; there is no such thing as the reasons for which we hold these beliefs” (pp. 19–20).

Our “thoroughgoing naturalist,” though, gleefully and deliberately misunderstands the question: “Why do I suppose that the future will be like the past? Why, because my species evolved in and I matured in an environment exhibiting key regularities, so my cognitive apparatus is attuned both to pick up those features and to anticipate like regularity. There is some question still at what stage of my development this sort of anticipation became fully formed, how much and in what ways it is determined by my neurological structure as opposed to experience . . . .” It won’t help to point out to these types that they aren’t answering the question you’re asking. I know. I’ve tried.

6. Second Philosophy

Let’s not get too down about Maddy, though. Not yet. Here finally is someone who finds it perfectly legitimate to understand the question “Why must it be green?” so that the coy “Because it’s either red or green, and it’s not red” won’t do for an answer.

Maddy takes her lead from Robert Fogelin, who balked at Wittgenstein’s suggestion that “the desire to find physiological explanations of psychological phenomena” was some sort of prejudice of the age of science. Fogelin said such phenomena “just seem to be of the wrong order to be simply brute or inexplicable.” He pointed out how absurd it would be to adopt a similar attitude about rain: “Imagine someone saying ‘it just rains, that’s all; explanation has to stop somewhere’” (1987, p. 207). Similarly, according to Maddy, someone might ask for a justification of Z—be prepared to take A and B as sufficient grounds or not—but in any
case to see them as the right sort of candidate grounds. Someone else, though, can ask for a justification of $Z$ knowing that $A$ and $B$ are its alleged grounds (perhaps even convinced that they are), and instead be soliciting an explanation of what makes them so, of the validity of the inference from $A$ and $B$ to $Z$. This is the sort of insistence that Carroll mocked but Bolzano took seriously. Maddy doesn’t just side with Bolzano. She’s even more resilient in her inquiry.

So how does she do it? Not directly, of course, but with a “device,” a “character,” a “sort of idealized inquirer” she dubs the Second Philosopher (2007, p. 1). Such is my fate. When I’m not cast in some such role myself, I’m represented in negative by someone who is. David Pears, incidentally, has attributed common misreadings of Wittgenstein to a tendency to “mistake the play within the play for the main drama” (1983, p. 12). Never having been welcome in the main drama, it’s hard for me to see these as misreadings. Nor can I much appreciate the hand-wringing over questions like whether Wittgenstein’s disdain for science is really directed at science per-se or merely at the idea that science will fill, rather than supplant, the role that philosophy used to play. The *Investigations* are chock full of appeals to science (“What is the relationship between name and thing named?”—“Well, what *is* it? Look at [this example] or at another one. There you will see ….” §37) in deliberate, even comical misunderstandings, clinging to the idea that empirical answers are the right sort to offer to philosophical questions. If the play-bill instructs you about when to take these comments seriously, I suppose part of the joke is my naivety.

Still, I can’t help thinking that my very distance from the audience, and from Maddy for that matter, draws me especially close to the Second Philosopher. Even with no opinion about whether what she has to offer ought really be called philosophy, or whether Wittgenstein or anyone else would accept it as such, I think I can appreciate what this character has to say about logic.

Which brings us to Maddy’s second big surprise: “Why must it be green?” Second Philosopher: “Actually, it needn’t be.”

Now, this is just what I’ve been saying all along! But before we get ahead of ourselves, we need to understand what sort of reply this is. The Second Philosopher still faults me for failing to infer that it’s green from it being either red or green and not red. She just observes that on her account, $Z$ follows from $A$ and $B$ only contingently, so that there’s no sense in which it “must” so follow: “the force of the logical must is compromised” (2014, p. 31).

Maddy describes several routes to this view of logic, but her principal reference point is the framework of Kant’s transcendental idealism. On Kant’s view, she says, “the truths of logic can be understood empirically,” as depending on very general structural features of the world, “but the force of the logical must [can be] explained only transcendentally” (2014, p. 13). That explanation: Logic is true of the world because of its most general structural features. Humans believe logical truths because no other sort of discursive, conceptual thinking is possible. But
the world we experience is a projection of such constraints on our thinking. No wonder, then, that the world would be structured so as to support logical inference (2007, p. 223).

The Second Philosopher wants to preserve a kernel of this Kantian picture, the ideas that (1) the truths of logic depend on general structural features of the world and that (2) the human belief in logic can be accounted for in terms of “primitive cognitive mechanisms.” But she’s not at all satisfied with Kant’s “transcendental” move of explaining (1) with (2). There is, she claims, just no evidence that the world is structured as it is because we think the way we do, compared with considerable evidence of a converse explanation: (3) “human beings are so configured cognitively because they live in a world that is so structured physically” (p. 226).

The casualties to logic of this view are legion. To begin, as already noted, the “force of the logical must” gives way immediately. Kant’s transcendental argument supplied us with an explanation of logical truths’ necessity. It’s second-philosophical converse tells us why “we tend to overlook the contingency of rudimentary logic” (p. 273). In place of a picture of why logic is necessary, we get a picture of why we erroneously think it is. Z follows from A and B, though it needn’t do so—despite the fact that humans cannot readily conceive of how it could fail.

Worse, a little second-philosophical poking around uncovers more cracks in the Kantian scheme. The weirdness of the quantum world, for example, makes it plain that necessity gives way, not to “always, but only contingently so,” but to “in many but not all cases,” for the truth of logical principles depends on the world exhibiting certain structures, which it does “in many but not all respects” (p. 244). The link between (1) and (2) has to be modified to (3’): Humans’ primitive cognitive mechanisms are the way they are because they live in a world by and large appropriately structured for that way of thinking to come out right and they interact almost exclusively with its structures of that sort (p. 245). Z needn’t follow from A and B, and in fact doesn’t when “it” is a photon and terms like “red” and “green” are replaced with “horizontally polarized” and “vertically polarized.”

Finally, some logical principles that the Second Philosopher urges us to accept fail fairly consistently, even in the ordinary world that we habitually navigate. Maddy cites the principle of excluded middle and the material conditional analysis of if-then claims. These are “ideal-

8Maddy does not point out that adding the law of excluded middle to the framework of basic inferences presumably licensed by her rudimentary logic makes “if A, then B” and “either B or not A” interderivable, so that really only one “idealization” separates classical logic from rudimentary logic. Instead she wages a war on two fronts, contending with issues of vagueness in her defense of LEM (pp. 285–91) and navigating the “paradoxes of the material conditional” separately (pp. 291–96). Possibly this is due to one direction of the interderivability depending on the ex falso quodlibet rule, which figures prominently in her discussion of the material conditional. It is unclear whether ex falso quodlibet is part of rudimentary logic, whose description Maddy deliberately leaves vague (p. 290). On the other hand, as she points out herself in 2012 (p. 484), Maddy’s flagship example of rudimentary inference, disjunctive syllogism, is not derivable (though it is admissible) in the bare inferentialist framework of the the minimal calculus—its validity rests on ex falso quodlibet.
izations” no different than the friction-less planes evoked by physicists. They are “admittedly false,” but still effective “in a range of applications” and worth adopting for that reason alone (p. 289). “Just as in the rest of science,” the Second Philosopher maintains, “we must be sensitive to the benefits and dangers of idealization, and satisfy ourselves that the idealization is appropriate to the case at hand” (p. 287).

The putative benefits recommending these idealizations are that they lend simplicity to the overall framework of reasoning (via truth-functionality at the propositional level) (pp. 285, 292), that they are familiar to us (p. 288), that they appreciably strengthen our inferential resources (p. 288), and that they usher in meta-theoretic niceties like the Deduction Theorem (p. 287). It’s possible to contest these individually: Much less than classical logic is needed for the Deduction Theorem to hold; the material conditional strikes very few people as familiar from their days of untutored reasoning; etc. But let that not distract us from the view itself. The propriety of laws of logic lies not in their necessity, nor even in their universal applicability. Few, if any, are applicable in all contexts, and some are more widely apt than others. Even when a logical law is applicable, though, this need not be due to it being literally true. Maddy cites approvingly H. G. Wells’s observation of the mismatch between logical validity and objective truth: “In my way of thinking, relentless logic is only another name for stupidity—for a sort of intellectual pigheadedness. If you push . . . enquiry through a series of valid syllogisms—never committing any generally recognized fallacy—you nevertheless leave behind you at each step a . . . marginal loss of objective truth and you get deflections that are difficult to trace, at each phase in the process.”

The combined force of these points suggests that even as the Second Philosopher distinguishes “rudimentary logic” from classical logic, they differ from one another only in the degree to which, and the frequency in which, their principles are true and, what’s not the same thing, in the range of their application. So what exactly is my error if I hesitate to infer Z from A and B? Other than the unhelpful reminder that I’ve accepted there being only two possibilities and ruled one of them out, the Second Philosopher has no all-purpose answer. She will look at the details of the case. Some worldly structures may be in place to underwrite the inference. Should they be absent, the inference may yet be effective. In any case, everyone draws it as a matter of course. Usually, that’s a good sign, quantum weirdness aside. Just don’t get carried away!

I think it’s time to learn a little more about the Second Philosopher and how she arrived at such iconoclasm. Her name suggests that she’s a philosopher who has caught a second wind. Perhaps after some initial exhaustion from the repeated fizzling out of her ambitions at Cartesian styled first-philosophy, she has, rather than simply thrown in the towel, re-tooled with methods continuous with science.

Maddy emphasizes that this is not the case—and that, although her position resists “out-
right definition,” we should get to know her by her ways (2007, p. 1). So far we’ve seen her reverse a key argument of Kant’s in light of the lack of evidence for his “transcendental idealism” and the extensive evidence for its converse. It turns out that several features of this one move are indicative of the Second Philosopher’s style: Although she has no working definition of what counts as science or active bias towards science over other topics or pursuits, she’s looking for evidence for claims of all sorts—the same sort of evidence, in fact—for what would commonly be recognized as a scientific claim as for what is thought of as philosophical or logical. Moreover, she finds obvious attractions in any number of schemes, as exemplified by the “considerable appeal” she sees in the central facet of Kant’s account of logic. So she willingly sets out to see how much of those schemes can be salvaged under scrutiny, even if their origins are, by her lights, arcane. The Second Philosopher takes up questions of meaning, truth, and all the rest “under the low and leaden sky,” not as a consolation prize, but because nothing has ever made “transcendental,” “ultimate,” or a priori perspectives appealing to her.

What really brings out her true colors, though, are her reactions to attempts to lure her out of these tried and true ways. Recall Strawson’s description of the thoroughgoing naturalist, the inquirer who “could accept the question” What causes induce us to infer according to the pattern known as disjunctive syllogism? “as one we may well ask.” In the remainder of that passage, paraphrased to fit our logical example, Strawson said this inquirer would remand that question to psychology and the study of child development, “but would do so in the justified expectation that answers to it would in fact take for granted” the validity of disjunctive syllogism (1985, p. 12).

This is precisely the Second Philosopher’s tack: “She acknowledges . . . that she can’t justify anything without using the methods of inquiry she’s developed for that purpose; the very suggestion that she attempt this seems wrong-headed to her” (2007, p. 32). But given her provisional attitude towards the validity of such logical principles, one might expect a distinctive response to allegations that her inquiry is fueled by the very principles that it’s meant to account for. After all, the outcome of her inquiry is that these principles are, despite all appearances, only contingently true and even fallible.

But the Second Philosopher is unmoved. “Our Second Philosopher freely acknowledges one poignant aspect of the human condition: we can’t step outside our system of beliefs and methods and justify them from an external perspective; the only perspective we can occupy is our own” (2007, p. 35). Never mind that her appraisal of logic falls short of what ordinarily is thought of as justification! “Her study of various methods of reasoning . . . are all [conducted] relentlessly ‘from the inside’” (2007, p. 32).

Maddy lets us see up close how this relentlessness colors the second philosophy of logic. The first vignette occurs when the Second Philosopher “rests with her conclusion that the world has, at least, a distinctive structure of individual objects” of the sort that rudimentary
logic depends on:

Here some will object on grounds of circularity—we use our beliefs about bounded, cohesive, spatiotemporally continuous objects to confirm that the world is populated with such objects . . . . We noted [before] that this objection often betrays the objector’s underlying wish for a “higher,” extra-scientific justification of science, so that science can be preferred to pseudo-science . . . on some neutral grounds. But the Second Philosopher finds no footing outside her own methods; she simply argues that astrology and creationism are wrong, on her own terms, unimpressed by the reply that they could run a parallel argument against her. She knows what’s wrong with that argument, too! (2007, p. 235)

But, one wants to object, what else could you expect to discover, having taken for granted the very thing you’ve verified! “The Second Philosopher’s justifications are not circular in this sense,” Maddy replies. The discoveries about the quantum world are enough to illustrate “that their success is not a foregone conclusion” (2007, p. 235).

Strawson, again, anticipated the ensuing dynamic:

Our inescapable natural commitment is to a general frame of belief and to a general style . . . of belief-formation. But within that frame and style, the requirement . . . that our beliefs should form a consistent and coherent system, may be given full play. Thus, for example, though Hume did not think that a rational justification of induction in general was either necessary or possible, he could quite consistently proceed to frame ‘rules for judging cause and effect.’ Though it is Nature which commits us to inductive belief-formation in general, it is Reason which leads us to refine and elaborate our inductive canons and procedures and, in their light, to criticize, and sometimes to reject, what in detail we find ourselves naturally inclined to believe. (1985, p. 14)

An obvious concern, though, in the case of logic, is that in our attempts to distinguish cases where idealization is innocent from cases where it isn’t we will be susceptible to the very deceptions of idealization that we are trying to harness. This sets the stage for a second vignette. After reminding us that “we must come to terms with the fact that even this rudimentary logic cannot be assumed to apply effectively in every situation,” Maddy presents the Second Philosopher’s project of validating disjunctive syllogism in a specific situation: “The last card in her hand is either a club or a heart; it can’t be a heart (because Joe played the last one); so she’s holding a club.” How do we know that this inference is valid, that nothing akin to quantum phenomena are at play? “The cards and bridge players here are objective individuals; each card has a rank and suit; each card either has been played or remains in one

\[9\]See 2007, p. 108.
and only one player’s hand; and so on. The world of the bridge game is [the sort in which] disjunctive syllogism is valid” (2007, p. 272).

But doesn’t this “defense” of disjunctive syllogism make use of disjunctive syllogism? In reply to this question, Maddy just suggests we observe that “the seemingly transparent soundness proof can be carried out only because the necessary objects, properties, relations, dependencies, are available in this context, . . . there to be exploited.” But isn’t it an assumption that they are there?—Not at all, it’s a discovery!—But isn’t that discovery aided by disjunctive syllogism as well, to say nothing of other fallible principles and idealizations that might or might not mislead in this task?—Indeed, and the kitchen sink!—Then how will you ever know that you haven’t been misled?—“All this can be disheartening to the philosopher in search of certainty, but [I] never undertook to philosophize from a point of view more secure than that of science” (2007, p. 299).

Maddy sizes the Second Philosopher up pretty well: “Her equanimity in the face of this admission may frustrate us” (2007, p. 35).

7. Mathematics

Wittgenstein had abandoned philosophy after writing the Logisch-Philosophische Abhandlung and for years resisted others’ efforts to rekindle his interest when Friedrich Waismann and Herbert Feigl “managed to coax Wittgenstein, after much resistance, to join [them] in attending the lecture” of the Dutch topologist Luitzen Egbertus Jan Brouwer. The lecture, “Mathematics, Science, and Language,” was delivered in Vienna in 1928. According the Feigl, “a great event took place” just afterwards:

Suddenly and very volubly Wittgenstein began talking philosophy—at great length. Perhaps this was the turning point, for ever since that time, 1929, when he moved to Cambridge University Wittgenstein was a philosopher again. (Feigl 1981, p. 64)

Feigl later wrote in a letter to George Pitcher that the evening “marked the return of Wittgenstein to strong philosophical interests and activities” as he immediately “began sketching ideas that were the beginnings of his later writings” (Pitcher 1964, p. 8n).

Also attending Brouwer’s lecture were such luminaries as Kurt Gödel and Rudolf Carnap. And although there is little doubt about Lewis Carroll’s considerable direct influence on Wittgenstein’s late thought,10 I’m convinced that my appearance that evening is what led to my role in the Investigations.

10I am unsure just what the author has in mind, but Pitcher presents some evidence of influence in 1965.
Brouwer described language as a natural development in the social history of the human species, organized in order to facilitate increasingly complex tasks. To set the frame for a genetic analysis of language, he contemplated a simple social order organized with the help of the first forms of linguistic behavior: “At primitive levels of culture and in primitive relationships, the transmission of the will is achieved by a simple gesture; and here the cry is particularly effective.” This scene, of course, would later be echoed in the opening sections of the *Investigations*. So too would Brouwer’s continuation: “But in matters belonging to the organization of a higher human community the tasks to be imposed are too various and too complicated to be brought about by simple cries,” leading to more elaborate forms of communication (§I.3).

As an outgrowth of human social behavior, language by its nature does not simply or straightforwardly encode concepts. Uprooted from its context, an utterance or inscription has no meaning. It’s only in its social role, among humans who share a natural history and set of purposes, that language functions.

Naturally, this leaves me in a predicament of being unable to interpret occasional instructions thrown my way, when these instructions’ natural homes are contexts foreign to my own experience and training. “For transmission of the will mediated by speech,” Brouwer declared, “there exists neither exactness nor certainty” (§II). Therefore such bits of language will be open to multiple interpretations, with nothing inherent in the language favoring one among them. And crucially, “for pure mathematics as well there exists no certain language, i.e. no language that excludes misunderstandings in conversation and, when it is being used to prop up the memory, protects against errors . . . .” So even simple instructions like “add 2” will be subject to misunderstanding.

By now we are familiar with the wide range of potential correctives one might propose for this situation: an algebraic formula, a mental image, etc. Brouwer focussed on one, which he associated with a prominent research program in the foundations of mathematics: “This circumstance cannot be remedied, as the formalist school attempts, by submitting mathematical language itself to a mathematical contemplation, revising it . . . and then coming to an understanding about it in a language of the second order, a metalanguage.” After all, if the new-fangled metalanguage is to function at all, it must to that extent figure into the social practice of a human community, so though it may “indeed guard against misunderstandings and errors with great probability . . . , in keeping with the essence of language it cannot do so with absolute certainty” (§II).

Far from leading Brouwer to skeptical conclusions or despair, this realization prompted a familiar naturalistic pronouncement: “On the other hand stability and exactness of language is not necessary in practice because people are drilled by a common will to an automatism
of understanding incomplete sentences.”11 There can only be a lack where there is a need, and humans, sharing as they do an evolutionary and social history together with like training, habituation, and community, need no such exactness, or “completion,” in their language. Their every utterance is made possible by, and therefore is informed by, this elaborate backdrop.

The right conclusion to draw from my inability to catch on, therefore, is not that something is defective with language, but, alas, that something is defective with me—specifically, that I am not part of the same world as those for whom the instruction is meaningful. The idea that something could be supplied that would guide me to the right interpretation is the same as the idea that there is something guiding ordinary humans to the right interpretation. Brouwer called it “the false belief in a magic reach of language . . . , a reach that exceeds its character as a means for transmission of the will” (§II).

If the image coming into focus here evokes Wittgenstein’s dismissal of his early view of language—as a “picture linked with, reaching up to reality” (1921, §2.1511), as one in which “naming appears as a queer connection of a word with an object” and “as an occult process” (§34)—in favor of considering language as “part of the natural history of human beings” (§415) and “seeing that nothing out of the ordinary is involved” (§94), consider that Brouwer even fielded an objection based on the possibility of a “private language”:

Language is therefore absolutely a function of the activity of social man. Even if the individual man in total isolation uses language to prop up his memory, this is only because he must take the sciences and the organization of the community into consideration. (§I.3)

Pressing further, Brouwer diagnosed the misunderstanding of language just described as “the natural consequence of a much older, more primal, more consequential and more deeply rooted error, namely, the reckless trust in classical logic” (§II). To explain both what he meant by trusting logic and in what sense doing so is careless, Brouwer turned to etiology:

Already in antiquity man possessed a very perfect language (i.e. one that practically excluded misunderstandings) for the mathematical contemplation of finite groups of things of the objective spatio-temporal world. For this language there are certain forms of transition from correct statements to other correct statements; these forms of transition . . . were gathered together under the name of logical principles. (§II)

These “principles proved themselves,” Brouwer said, as the statements derived from them turned up time and time again to be independently verified. “And then one came to trust statements that had been derived by means of the logical principles even when they were not susceptible to direct check” (§II).

---

11This sentence appears only in the 1933 Dutch version of the lecture, translated by Heyting in Brouwer 1975.
The suggestion here is that our “trust” is out of line because such statements might in fact be false. Brouwer did caution against this, but in fact it is just a consequence of a more general, and insidious, error. Our “trust” in logical principles leads us to see logic as something sublime when in fact it is mundane. “The practical reliability of the logical principles,” at least when they are reliable, “rests on the fact that much of the world of perception, in its finite organization” exhibits a certain “loyalty and contentedness” (§II).

Of course we do not tend to think that its being green because it’s either red or green and not red depends in any way on contingent features of the world, much less the fact that we are considering it only in finite swaths. We think the inference is inevitable, conceptual, necessarily secure in a way that empirical principles cannot be. Brouwer has an explanation for this fact about us, one that hooks into his more general account of language: “Man has for ages been blind to this sober interpretation because he did not recognize that words are nothing but a means of transmission of the will, and regarded them [instead] as a means of indicating fetish-like concepts.” As they are “supposed to represent a priori laws governing the concepts and their linkages,” logical principles do not strike us as devices whose reliability could depend on features of the world, however ubiquitous (§II).

Now, the Second Philosopher disagrees with some aspects of this account. Humans feel the force of the logical must, not because they are basically prone to superstition, but because evolutionary pressure has attuned them to those features of the world where logic applies and because their cognitive development occurred almost exclusively in this environment. Finitude is not obviously the relevant feature for the applicability of rudimentary or even classical logical principles. And it’s not at all clear that general misconceptions about language figure into the illusion of logic as somehow special, even if on this point Wittgenstein might agree with Brouwer.

But Brouwer’s core account, his “sober interpretation,” is the Second Philosopher’s own. The truths of logic depend on the world exhibiting certain very general regularities, which by and large it does, though not always. Our belief in logical truths, even our mistaken view that they are true universally and independently of any facts about the world, is due to our habits of thinking, themselves contingent facts about us. And these habits of thinking are ours because in our natural history of navigating the world we interacted nearly exclusively with those structures that allow our logical principles to “prove themselves.”

The Second Philosopher agrees with Brouwer, too, that the characteristic laws of classical logic fare worse in this scheme than a set of more rudimentary principles. Sometimes Brouwer is mistaken for an advocate of a traditional view about “intuitionistic” logical principles—as holding that just these are a priori truths, constitutive of reasoning, whereas the validity of LEM is a contingent matter.12 We’ve seen that he actually held that no principles of logic

---

12 This misunderstanding seems to lie behind the popular, but historically strained, opinion that Wittgenstein’s
are true \textit{a priori}, that their appearance as necessities of thought is an illusion, and that at most such principles can be “practically reliable.” Still, an unmistakable thread in Brouwer’s polemic is the insistence that \textsc{LEM}’s practical reliability is extremely limited compared to that of the intuitionistic laws—precisely the Second Philosopher’s assessment.

Then these two part ways. Faced with the realization that \textsc{LEM} isn’t literally true beyond a limited range of applications, Brouwer wondered whether it still functions, “in the mathematics of the infinite, as a practically reliable means of transition between constructions.” Famously, he concluded that it cannot. Because logic, as language, has no inherent meaning and serves only to transmit a mathematician’s will, the logical derivation of a statement is a promisory note that a direct mental construction can be performed. \textsc{LEM}, “although it is not \textit{correct}, nevertheless, if one presupposes it exclusively for finite species of properties, is \textit{consistent}.” Statements derivable with \textsc{LEM} from verifiable truths about infinite assemblages, on the other hand, are among those that are not susceptible to direct verification of this sort. As a consequence, Brouwer advocated for the “collapse” of “considerable portions of the previous mathematical edifice” (§III).

Faced with the same realization, the Second Philosopher goes to the other extreme. After all, \textsc{LEM} is an idealization. “The debate isn’t over [its] truth or falsity . . . but over its effectiveness . . . in a range of applications” (2007, p. 289). Brouwer offers one proposal for what it would be for a logical principle to be effective in mathematical reasoning. But his focus on constructive verification and mathematical contemplation strikes the Second Philosopher as dogmatic. As every application has its own purpose, no one criterion could determine whether a tool is effective in carrying it out. We need to find out what the study of infinite aggregation is good for—why the Second Philosopher should take interest in it in the first place—and determine from there whether \textsc{LEM} or other principles facilitate or hinder that goal. She concludes that classical logic and much more, even if misleading and inappropriate in other applications, are particularly appropriate in set theory.

To understand how she reaches that conclusion, let’s consider first how the Second Philosopher comes to take interest in mathematics. All human activities, as such, are “apt subjects” for her “sociological or anthropological study of the role they play in our culture, her psychological study of their role in the individual psyche, [and] perhaps biological . . . studies of their basis and origins.” But mathematics stands apart from such activities as astrology and theology because of its “contribution to our scientific description of the world”:

In the course of her examination of her own best methods, she will need an account of how and why pure mathematics plays the role it does. Indeed, given its importance, she comes to see that one way of pursuing our understanding of the

\footnote{return to philosophy was prompted, not by an affinity with anything Brouwer was saying, but by an aversion to it. Details about this and a related opinion can be found in Marion 2003.}
The Second Philosopher won’t be content as a distant observer of mathematics. Because it’s not just part of the world she wants to understand but also informs and impacts her own understanding of that world, she wants to be involved in assessing its methods as part of her project of refining her own tools of inquiry.

Turning specifically to set theory, her next step is to determine what its purpose is.

No doubt a mathematical practice as rich and varied as contemporary set theory functions in the service of a range of different goals and subgoals, but for present purposes, the Second Philosopher isolates one fairly uncontroversial motivation: set theory hopes to provide a foundation for classical mathematics, that is, to provide a dependable and perspicuous mathematical theory that is ample enough to include all the objects of classical mathematics and strong enough to imply all the classical theorems about them. (2007, p. 354)

This motivation, we then find out, recommends a strong methodological agenda in the face of the incompleteness of axiomatic frameworks for set theory: “Given that set theory is (at least partly) designed to provide a foundation for classical mathematics, to provide a single arena for mathematical existence and proof, it does make sense to try to make our theory of sets as decisive as possible, to try to choose between alternative axioms, [and] to try to rule out models that do this foundational job less well than others” (2007, p. 355).

These scruples suggest, first, that we ought not rest content with an axiomatic framework like ZFC, in which questions like the Continuum Hypothesis (CH) are unsettled (“as decisive as possible”). But recognizing that there are multiple ways to move beyond ZFC, settling CH one way or the other, this is not enough. Some of these ways will advance the foundational program while others will not (“models that do less well than others”). So the Second Philosopher has to ask what it is for set theory to serve its foundational goals well. Must it in some sense provide the right answers? Must it always provide some answer rather than none, or can success at the foundational program sometimes conflict with the ambition to be decisive? Perhaps, even, there’s nothing to say in advance, as we learn about good set theory only by watching it evolve. These are complicated matters that lead us away from our examination of logic.

Setting aside the question of how CH might get settled, what interests me is the Second Philosopher’s conviction that we know in advance that it has some determinate truth value or another. Other writers—John Steele, Hartry Field, Solomon Feferman—have supposed no such thing. So what are the Second Philosopher’s grounds for this opinion?

First of all, the Second Philosopher observes that statements like “CH or not-CH,” “CH is
either true or false,” and “Either CH is true or not-CH is true” are “straightforward assertions of set theory.” Now, she concedes, the difficulty in coming up with new axioms with enough intuitive appeal raises a serious question about whether we will decide the question one way or the other, but by her lights this doubt “does not change the fact that CH is either true or false” (2007, p. 370). That alone provides some motivation to make the attempt.

Here we see LEM at work in the philosophy of set theory, assuring us that CH cannot be inherently indeterminate, setting an agenda for its solution. This might strike one—and would surely strike Brouwer—as incredible given how tenuous the Second Philosopher’s account of the propriety of classical logic is. Addressing these concerns, Maddy points out that “the idealizations of classical logic present none of the potential headaches here that they do in ordinary physical cases” (2007, p. 382).

This is a familiar mantra. Students the world over who find the truth-functional analysis of conditional statements unconvincing are directed to mathematics as evidence that the material conditional is not completely worthless. But given that Field and others, who propose that CH might not have a determine truth value, specifically doubt whether the set concept has sharp boundaries, extending this defense to set theory seems, well, circular. If the universe of sets exhibits vagueness, then the idealizations of classical logic will mislead in that arena. Appeal to LEM cannot convince anyone that CH has a determinate answer. I detect the Second Philosopher’s true colors showing once more.

But even granting the validity of statements like “Either CH is true or not-CH is true,” all that would seem to follow is that in each model of set theory either CH or its negation will hold. The appeal to classical logic fails to secure the determinate truth of CH apart from a further conviction that set theory eventually will be sufficiently developed so that CH has the same truth value in each of its models.

The Second Philosopher has just this conviction. She observes that when mathematicians confront two alternative set theories, the end-directed aims of providing a decisive foundation for mathematics should compel them to adopt whichever of these can subsume the other as a definable sub-structure. Her preference for the theory of large cardinals over Gödel’s axiom of constructibility follows this scheme. Building on this, she claims that if set theorists “were faced with two attractive theories . . . neither of which is capable of recapturing the other,” the same goals “would recommend a concerted search for a more powerful theory that could encompass both” (2007, p. 387).

In fact “as soon as those universes are posited, we find ourselves wanting to talk about the relations between them—about isomorphisms and embeddings and so on—which requires a larger arena containing” them all (2007, p. 387n). Maddy considers the possibility that efforts along these lines are frustrated, that the goods of unification and maximization come to conflict with one another. In this case, the Second Philosopher would be surprised to learn
that “there are different kinds of sets, that truth and existence in mathematics turn out to be relative to a particular universe of sets.” Maddy calls this imagined scenario “dire” (2007, p. 389).

Considering my difficulties skip counting past 1000, I shouldn’t speak for set theorists. But as a natively curious inquirer, the “dire” scenario Maddy considers would only pique my further interest in the field, much as the undecidability of CH in standard axiomatic frameworks has done. All speculation aside, it seems to me that Maddy’s own observation about mathematicians’ reflexive interest in pursuing relationships among different universes of sets, once they’ve been isolated, can be turned on its head.

According to the Second Philosopher, the primary motivation for the study of set theory—the foundational aspiration—harnesses this reflex in the form of the articulation of increasingly general background theories that will, for example, decide more questions about real numbers. But I wonder if the reflex isn’t itself the primary motivation of set theory. Just scanning the most often cited articles in the Journal of Mathematical Logic, I discover that every countable model of ZFC embeds into its own constructible universe, that it is possible to preserve CH even when adding, with finite conditions on a forcing poset, a club subset of $\omega^2$, and that some problems of compelling interest are about which cardinal numbers remain inaccessible under certain extensions of the universe of sets, assuming that CH holds. In other words, I see unbridled interest in the playground of ideas: “What do sets like those look like when I put on these glasses?” I do not find a background motivation of trying to narrow in on the one true theory of sets, or on foundational goals, as ordinarily conceived, of any sort.

By the same token, debates about whether principles of logic like LEM or axioms like CHOICE are true in the universe of sets seem to me to be a holdover from the era of foundational preoccupations. Shedding these hangups, set theorists delight in observing that the Cantor-Schroeder-Bernstein Theorem—a trivial corollary of the well-ordering principle—is provable without CHOICE, assuming LEM. Precisely: The theorem holds in every Boolean topos, but fails in, for example, the arrow category $\textbf{Set}^{\rightarrow}$, and so is not intuitionistically valid. The whole intrigue of this observation depends on a nonchalance about what principles are factually true. One might even say that the theorem only makes sense when you suppose that CHOICE is false, an attitude that does not naturally lend urgency to the question whether it really is.

But I think I can anticipate the Second Philosopher’s response. She isn’t working with set theorists’ motivations, but with her own. Were it not for set theory’s foundational prospects, Cantor’s Paradise might strike her as it did Wittgenstein, as one worth leaving after a quick look around (Wittgenstein 1939, §103). The Second Philosopher presents her special interest in mathematics as an honorific, but it seems to me that she doesn’t extend the same courtesy to set theorists that she does to astrologers and theologians, leaving them to follow their own
complicated, competing, even contradictory motivations instead of selecting from among them one that makes their inquiry worth the time.

8. The last word

I fear I’ve now reached the time at which we are most gifted with prophecy, and what I have collected over all these years are questions, not answers. They will be recognizable from my most recent encounter, but each carries the weight of history, which is why I’ve prolonged my tale. If logic isn’t necessarily true, why pursue the development of its theory? Why even call it logic, or, for that matter, anything at all, if it differs from ordinary physical laws only in its breadth of applicability and generality? Why assume, unlike Brouwer, that $LEM$ applies to infinite sets, if you agree with him that it’s a mere idealization?

For some time I suspected that the Second Philosopher could never satisfactorily answer these and other questions. Her recorded attempts all seem circular and her response to being called out for that inscrutable.

But it’s hard to sustain suspicion about what someone just owns up to. Here’s Maddy wrapping up her introduction of the Second Philosopher: “Finally, of course, none of this amounts to an argument that we should all strive to conduct ourselves as Second Philosophers. My hope is that the appeal of the approach will be obvious to the susceptible” (2007, p. 3).

“Where does that leave the unsusceptible?” I often wondered. Now I’m beginning to understand. The Talmud (Pesachim 110b) says that demons prey on whoever takes measures to guard themselves from them, making those provisions the more urgent. People who don’t give them a thought are left alone. And so it is with me. I torment just those who take precautions against me. For everyone else, I simply don’t exist.

I can’t enroll in your class on my own. But how easily, until now, I’ve always obtained instructors’ permission! Don’t you see that your provisional account can never explain how logic, “viewed all at once,” hooks up with the world, “as it is in itself?” That your defense of disjunctive syllogism rests on that very principle? That your use of $LEM$ to settle questions about sets already presumes their answers? That your explanation of this feature of logic and your justification of that logical inference are rough-and-ready and for that reason not getting to the heart of the matter?

—“Granted, it isn’t the sort of explanation/justification that would persuade in all possible situations . . . and thus not the sort of thing to satisfy the traditional philosopher, but that’s as it should be” (Maddy 2014, p. 119).
Works Cited


