Inescapable concepts

Thomas Hofweber
University of North Carolina at Chapel Hill

July 13, 2022

Abstract

It seems to be impossible to draw metaphysical conclusions about the world merely from our concepts or our language alone. After all, our concepts alone only concern how we aim to represent the world, not how the world in fact is. In this paper I argue that this is mistaken. We can sometimes draw substantial metaphysical conclusions simply from thinking about how we represent the world. But by themselves such conclusions can be flawed if the concepts from which they are drawn are themselves flawed. I propose that we can overcome these limitations by focusing on a special class of concepts: inescapable concepts. Combining arguments about what the world is like from considerations about our concepts alone, together with an argument that the relevant concepts are inescapable, leads to a novel method for metaphysics, which is broadly neo-Kantian.

Keywords: inescapable concepts, Kant, language-metaphysics gap, rational traps, immanent metaphysics, immanent stance

1 The language-metaphysics gap

It is widely accepted that one cannot draw substantial metaphysical conclusions about the world from reflecting on our language or our concepts alone. Thinking about our language or our concepts only tells us how we represent the world, not how the world is. Metaphysics concerns reality, and thus how the world is, not how we represent it to be. Our representations of the world are only revealing about us, and since we are small part of reality, they are revealing about a small part of reality. But metaphysics concerns reality in general, not just the small part of it that is us. Thus focusing our our own representations of the world is not a proper way to achieve results in metaphysics. At least not unless one thinks that results in metaphysics are analytic or conceptual truths. We can grant that one can establish analytic truths by reflecting on ones own representations, but few would hold that this would lead to substantial metaphysical results.

This simple argument has serious implications for how metaphysics should proceed, what it can hope for, and how it relates to other philosophical disciplines. Although we certainly will need to clarify both the argument as well as the conclusion it hopes to reach, which I will do momentarily, we can already state why this is significant. If thinking about our own concepts were a source of metaphysical insight, then at least it would make sense how we can achieve such insights. But if that is ruled out, then how should metaphysics make progress, in particular when it is conceived as being a part of philosophy? Maybe metaphysics has to move closer to the sciences, and read off metaphysical consequences from the largely empirical results of the sciences. Or maybe metaphysics has to make decisions on the basis of theoretical virtues alone, considering simply which metaphysical theories are simplest or most parsimonious. Or maybe metaphysics does not answer question of fact at all, but instead merely concerns the construction of models of what is believed to be the facts, or with the selection of useful concepts, or the repair of flawed concepts, and so on.

All this affects the status of the philosophy of language within metaphysics. If considerations about language can't lead to positive metaphysical insight, understood as answering questions of fact, then thinking about language can still prevent us from making mistakes and from accepting the wrong answers, but it won't help us otherwise in finding the right answers. The philosopher of language can point out that this argument here overlooks a scope ambiguity, and that argument there relies on the wrong semantics of certain expressions, and so on. This helps to correct errors, but not to make a positive, constructive contribution towards how things are in metaphysically relevant ways. We can say that so understood the philosophy of language only as a *corrective*, but not a *constructive*, role in metaphysics. This is just like how we normally think of the significance of epistemology in metaphysics, something that will become relevant later on:

it, too, only has a corrective, but not a constructive, role in metaphysics. The epistemologist can point out that these conclusions are badly justified, and those theories have little evidence as their support, but this by itself does not lead to a positive, constructive proposal about how things are in metaphysically relevant ways. Consequently, the standard view to hold is that epistemology can at best tell us how well supported our metaphysical theories are, but not more directly what these theories should say.

In this paper I would like to argue that all this is a mistake. It is indeed possible to draw substantial, synthetic metaphysical conclusions directly from reflecting merely on how we represent the world. I will give an example of this just below, and explain why such an example is indeed possible. Nonetheless, I will argue that by themselves such conclusions are insufficient to properly resolve the metaphysical issues at hand. By themselves these arguments are not enough, but they can be improved upon. How that can be done is the main topic of this paper. I will argue that such conclusions can properly be drawn by thinking about a special class of concepts: inescapable concepts. Combining largely empirical considerations about our own representations together with that these representations are inescapable, in a sense to be explained below, allows us to draw conclusions about what reality is like which properly settle metaphysical questions. This gives rise to a general strategy for finding out about reality by reflecting on our concepts, which naturally leads to an approach to metaphysics that is broadly Kantian in spirit. It holds that a distinguished class of our concepts can be a guide to reality, and that we can find out about general features of reality simply by reflecting on this special class of our representations. I will argue that on this approach both the philosophy of language as well as epistemology will have a constructive, not just a corrective, role in metaphysics.

All this must surely sound rather programatic, and it is. I will in the following outline the main parts that need to come together to defend this approach to making progress in metaphysics, without developing any of these parts in detail. The goal of this paper is to put the general approach on the table and to make clear what needs to be done to carry it out for a particular case. I hope to make clear that this is a fruitful and promising strategy for progress, but how much progress can actually be achieved this way is beyond the scope of this paper. This paper presents the key ideas of the general approach and how they come together, other work will have to fill in the details.[1] To start, let's see how we can straightforwardly answer metaphysical questions about reality by thinking about our own representations alone, even though the answers are not analytic, and why this, by itself, is not enough.

[1]For more on this general approach, with a defense of a version of idealism as a special application, see [Hofweber, 2023].

To put a label on it, let us call *the language-metaphysics gap* the thesis that one cannot draw substantial metaphysical conclusions about the world from considerations about our language alone.[2] The key here is "language alone". One can, of course, draw all kinds of conclusions from our language together with which sentences are true. But that is not about language or representations alone. It instead is about language and how language relates to reality. Language *and* how it relates to reality concerns not just us and our representations, but also reality. Uncontroversially, language plus truth allows for substantial conclusions about reality. But it is also fairly uncontroversial that from language alone one cannot draw such conclusions. At best one can draw analytically true conclusions, or derive conceptual truths, or draw conclusions about a small part of reality: human beings. But these are never substantial metaphysical conclusion in the intended sense, since metaphysics concerns synthetic hypotheses, and metaphysics does not just concern the small part of reality that is us. Apparently no real and substantial metaphysical question can thus be reached from considerations about language alone.

Plausible as this might be, I would nonetheless like to argue now that it is false. One can indeed derive substantial metaphysical conclusions simply from reflecting on our representations alone, without assuming anything about whether these representations are accurate or true. I will first give an example of an argument that answers a substantial metaphysical question from considerations about language alone, and then explain how such an argument is possible in light of what we have seen so far.

Consider the question whether natural numbers exist. This is a question about reality, what it contains and what things exists, not just a question about us. Nonetheless, this question can be answered from considerations about language alone. The key to this is to reflect on the semantics of number words: 'one', 'two' 'three', etc.. There is a real question about the semantics of number words, which is connected to an old puzzle about the occurrence of number words in natural language, one going back to at least Frege in [Frege, 1884], and has been actively discussed ever since.[3] On the one hand number words sometimes appear syntactically like adjectives, as in

(1) Jupiter has four moons.

On the other hand, they sometimes appear in a position usually occupied by a proper name, as in

[2]Heather Dyke calls drawing metaphysical conclusions from our representations "the representational fallacy" in [Dyke, 2008]. See also [Hofweber, 2019].

[3]See, for example, [Dummett, 1991], [Wright, 1983], [Hodes, 1984], [Hofweber, 2005], [Moltmann, 2013] and many others.

(2) The number of moons of Jupiter is four.

How can one and the same expression appear in both of those syntactic positions? There are a number of options that wold explain this, and all of them are simply options about our language alone. One is that number words are ambiguous: there is one that is a name, and another one spelled the same way that is an adjective. Another option is that number words are essentially the same in both kinds of occurrences, but that there is some explanation of why they appear syntactically in a position that is contrary to what it would normally be expected to be in. So, maybe number words really are adjectives, but for some syntactic reason or other they appear to be like names in (2).[4] Or maybe number words really are names, but for some syntactic reason or other they appear like adjectives in (1).

Suppose the first uniform account is correct: number words are like adjectives, but they can appear syntactically like names for syntactic reasons that such an account would spell out. In particular, they have the semantic function of adjectives even when they falsely appear to be names. Adjectives are not referring expressions, they do not have the semantic function of picking out some entity. Adjectives modify a noun, but they don't refer. All this so far are simply considerations about language alone. The reason why the adjective appears like a name in (2) is one about our language, not the world. Furthermore, suppose that this picture extends to number terms more generally. Then even terms like 'the number four' are not referring or denoting expressions.[5] They can occupy the same syntactic place as names or other referring expressions, but their semantic function is different. All this is still only about language alone, what we do when we use number words or number terms.

Number words and number terms more generally are non-referential on this picture. It is important here to distinguish two kinds of non-referential expressions. One can be seen to be non-referential by considerations about language alone, while the other one can't. A paradigmatic case of being non-referential in the second sense is an empty name. An empty name aims to refer, but fails, since the world does not cooperate: it does not contain the object the name aims to refer to. In this sense, being non-referential is not about language alone. It is about the relationship between language and the world. But a second kind of being non-

[4]To be clear, this classification of number words as adjectives is just to illustrate the issue at hand. A more realistic syntactic proposal along those lines will treat them as determiners or modifiers, but the details of this classification is not really relevant for us now.

[5]How could this be? One option is that 'the number four' is not a description, but 'the number' is aposited to the non-referring number word 'four', similar to how 'the philosopher' is aposited to 'Aristotle' in 'the philosopher Aristotle'. If so, then assuming 'four' is not referential to begin with, it is arguable that 'the number four' isn't referential either. More details on all this can be found, for example, in [Hofweber, 2016], chapters 2, 5 and 6.

referential is about language alone. This is the sense in which an adjective is non-referential. It does something completely different semantically than referring. It is non-referential, since its semantic function is something other than reference. If the above option about how to understand the different occurrences of number words and number terms were correct, then number words would be non-referential in this second sense. They are not even in the business of referring as they do something different altogether semantically.

Suppose then that by considerations about language alone we have determined that number words and number terms are non-referential in the second sense. Then we are in a position to answer the question whether natural numbers exist. The argument for this is simply the following: consider all the existing objects and pick one of them at random. Could it be the number four? I am now assuming for the sake of this argument that number terms in English are non-referential in the second sense. This, in particular, applies to my use of number terms just now. Thus when I ask whether this randomly chosen object might be the number four, then I use 'the number four' non-referentially. It occupies the syntactic position that can also be occupied by referring expressions, but it itself isn't a referring expression, by assumption. It merely appears syntactically like one, for some syntactic reason. But then the sentence 'this thing is the number four' is perfectly grammatical, but it can't be true. Here 'this' refers to the object we picked at random, but 'the number four' does not refer at all. So, this thing we picked out is not the number four. And since that thing was picked at random our conclusion generalizes: nothing in the domain of existing things is the number four. The number four thus is not among the existing things and so it does not exist. And the same holds for all the other natural numbers. Thus natural numbers do not exist. We have answered the metaphysical question about the existence of natural numbers simply from considerations about our own representations alone. And this bridges the language-metaphysics gap.

But how can an argument like this possibly work in light of what we have seen above, and in light of the motivation for the language-metaphysics gap? What explains that we could bridge the gap after all? The key to seeing how this can be is to focus on the question we asked and the concept we used to articulate it. When we tried to find out about reality we asked a certain question: do natural numbers exist? This question is one about what reality is like, not one about our representations of reality. Nonetheless, we used our representations to ask this question. We used our concept of a natural number to ask whether there are natural numbers. Focusing on the question we ask and the concepts we employ in that question gives rise to the possibility that thinking about our concepts allows us to see what the answer to that question

must be. And just this seems to be the case in this example. If our representations are as we assumed them to be in this example, then we can answer the question we asked simply by investigating the representations we used in asking it. If number terms are non-referring and natural numbers just are the number one, the number two, and so on, then the question whether there are natural numbers can be answered in the negative by thinking about our representations alone. That's how it was possible to bridge the gap. But it should be noted that even though we answered the question we asked by reflecting on our concepts, the answer is not a conceptual truth. That there are no natural numbers is not a conceptual truth, at least not in the sense in which this notion is often understood. Possessing the concept of number by itself does not put one into a position to determine in a fairly immediate way that there are no numbers. Instead, this requires substantial empirical investigation in linguistics and related fields, and is thus a broadly scientific discovery, based on empirical investigations into our own language.

Now, even if all this is correct, it might seem like a shallow victory for those who hope to find out about reality from considerations about our representations of reality. What we were able to conclude was that natural numbers do not exist, whereby natural numbers are simply the number one, the number two, and so on. If number terms are indeed non-referring, then maybe this conclusion can be drawn, but it might seem not much of an insight into what reality is like. Sure, it does not contain natural numbers, under those assumptions, but that tells us little about what it does contain. It might still contain number-like things, say, von Neumann ordinals, positions in structures, abstract objects forming an ω-sequence, and so on. These things wouldn't be numbers, but they would be a lot like what numbers were thought to be by many. Whether there are those number-like things is completely left open even if there are no numbers. And questions about the existence of various number-like things are perfectly legitimate questions which are so far unaddressed. And maybe they are better questions to ask than our original question about numbers. Maybe the real issue, the one we should be pursuing, is not about numbers, but about those number-like things. And furthermore, one might think that if number words are indeed ultimately like adjectives that appear for some syntactic reason like names, then this can be taken to show that our talk of numbers is flawed and defective in a way, and thus should not be taken too seriously for the purposes of science and metaphysics. The proper question to ask is thus not one about numbers, but about something else, something articulated with different concepts.

The significance of the answer to the question we asked is thus tied to the quality of the concepts employed in that question. If the concept is defective or otherwise unsuitable for metaphysics, then maybe a question asked with it can be answered this way, but the result, although

correct, will reveal little about what reality is like. To have a result of proper metaphysical significance more needs to be done. We also need some reasons to think that the concepts we employed in the question are high quality concepts for our purpose. Simply answering a question of fact alone does not speak to that issue.

Nonetheless, we answered the question that we ask at the outset: do natural numbers exist? That counts for something, but the issue remains whether we asked the right question in the first place. Did we probe reality in the right way when we asked about the existence of numbers, or should we better have asked about something else instead? Maybe we should not have used our concept of a natural number in the question, but some other, more metaphysically suitable concept. It is one thing to answer the question we asked, and quite another thing to ask and answer the right question or at least a good question. The issue thus gets pushed to this: what is the question we should have asked in the first place? What concepts should we have used in the questions we should ask, if not the concept of a natural number?

This leads to a second aspect of reflecting on our concepts and what metaphysical conclusions we can draw from them. It concerns not what we can conclude from the concepts we do in fact employ, but what concepts we should employ instead. To properly answer a metaphysical question we should not only find the answer to the question we did ask, but the answer to the questions we should have asked. And to determine that, we need to find out what concepts we should employ in the questions we ask, and then answer those questions. Thus simply having answered the question we asked is not quite enough to declare victory. We need to find out in addition that we asked a good question to begin with. It does not have to be the very best question, nor does there have to be a uniquely best question. But we should at least ask a proper, metaphysically suitable question, and not just any old defective question. If our original question was bad, then we need to do better, and ask and answer a different, better question instead. But not just any better question, a better question in the neighborhood of the original question. Maybe the best question to ask overall is why God is hidden, and if so, then there is a sense that we should have asked that question instead of the question whether natural numbers exist. But our concern here is not what the best question overall is, but what good questions there are which are similar to the one we asked. Are there better questions that are nearby alternatives, or better questions in the neighborhood of the original question? Should we ask about natural numbers or about some number-like things instead? Or alternatively: should we employ the concept of a natural number, or the concept of some number-like things instead when we ask a question about what exists? To properly make progress in metaphysics by reflecting on our concepts, we need to answer these questions as well, and not just be content with having

answered the question we happen to ask.

Let us call a *(metaphysically) deep result* an answer to a question that not only is the correct answer to that question, but also is the answer to a proper and good question to ask relative to nearby alternatives. Let us call a *(metaphysically) shallow result* an answer to a question which simply answers the question as stated, with no evaluation of the question itself, in particular whether it was a good question to ask. What we need are deep results, not just shallow ones. If a result is deep, then it can't just come from aspects of a flawed and defective concept, since such a concept could be improved upon, leading to a better question to ask. A deep result does not have to be the answer to the uniquely best question we can ask, but it must be the answer to a question that can't be improved upon, and that can't be replaced by an even better question.

To get deep results it is not enough to bridge the language-metaphysics gap; we also need to find out what questions we should ask and which concepts we should employ in these questions. But how could we possibly determine what the proper concepts are without first finding out what reality is like and thus what concepts are well suited to describe it? It surely seems that reflecting on our concepts alone won't give us any guidance on which concepts we should use and which questions we should ask. But, nonetheless, there is a way to make progress here, and how to do this is the main topic of this paper.

2 Inescapable concepts

It is hard to know what concepts we should employ in inquiry without knowing what reality is like. Our concepts should be perfectly matched to reality, and so we face a circle: we need to find out what reality is like to determine what concepts we should use, but to find out what reality is like we already need to represent it with matching concepts. It thus seems that we can't find out what reality is like unless we first find out what concepts we should employ, and also that we can't find out what concepts we should employ without first finding out what reality is like. We can call this the *conceptual circle* and it is this circle that we must break free of. In ordinary cases of inquiry the hope is that we can slowly break free of the conceptual circle in a holistic way, by making simultaneous progress on both parts: what reality is like and what concepts we should use. In a step by step way, we might get slightly better concepts allowing us to be slightly better at representing reality, which in turn should lead to even better concepts, and so on. Whether this answer for the case of inquiry is sufficient is one thing, but it certainly would be insufficient for the project of finding out what reality is like by reflecting on our concepts alone. Here we need to somehow find assurances that our concepts are the proper ones to employ by reflecting on these concepts alone, without first finding out what reality is

like. It is hard to see how that could be so, and with it it is hard to see how the project of finding the proper questions to ask, and answering them, could possibly succeed by reflecting on our concepts alone.

But maybe there is another way out of the conceptual circle. Maybe we can conclude for certain special concepts that they are the right ones to apply to reality, even though we otherwise do not yet know what reality is like. Maybe there is a reason to think that at least a certain concept C is the right one to employ in describing reality and that we can appreciate this reason without knowing much else about reality. But how could we possibly know that such a concept C is the right one? Here is one idea about how this might be. Maybe we can see that any attempt to replace that concept C with a different one, C^*, would always make things worse. Maybe we can recognize that replacing C with another concept would make things worse for us representationally, and thus that we should never do it. And if this were so for a particular concept C, then we might have good reason to think that C is just the right concept to employ in inquiry without first having to find out what reality in general is like.

When I consider whether a particular concept C should be replaced with an alternative C^*, then I consider the reasons for and against such a replacement. Should I employ concept C^* in situations where so far I employed concept C? Should I use C^* instead of C on the occasions where so far I used C? When I consider such questions I try to assess whether the switch would be a good one to make. After all, I have to decide on whether to make the switch, and so I have to consider and evaluate the reasons for and against the switch. I must assess, by my own lights, whether the switch should be made, since this decision has to be made by me, in my present situation. And in such a case it could in principle be that I can recognize that no matter what concept C^* might be, it would never be rational by my own lights to switch from C to C^*. In my own present situation I could then never rationally move from C to C^*, since it would not be rational by my own lights to do so, no matter what C^* might be. And thus I will have to conclude that there is no acceptable improvement on C, and thus that C is already perfect when it comes to describing reality: no other concept can do better.

Let us call a concept *inescapable* if one cannot rationally, by one's own lights, replace it with another one for the purpose of inquiry. That is to say, when I reflect on whether I should give up employing a particular concept C in inquiry, and rather employ a different one, C^*, instead, then I can see that this would always be irrational for me to do, by my own lights. If this were so, then I would be rationally required to reject such a replacement and thus to stick with the

original concept C. This notion should be clarified in at least the following ways.[6]

First, the issue is not whether a concept can be replaced with another one somehow, but only whether such a replacement would be rational. It might well be that every concept can be replaced through brain manipulation, hypnosis, drugs, or what have you. We are only concerned with whether a proposed replacement would be rational by our own lights. We thus should distinguish inescapable concepts from what we can call *hardwired concepts*. The latter are ones that are simply fixed in our minds, and can't be replaced at all, no matter what methods we might employ. But even if every concept can be replaced somehow, the question remains whether we ourselves could always find such a replacement rational by our own lights. This latter question is the question of concern for us here.

Second, the issue is whether a replacement would be rational *for the purpose of inquiry*. It might make a lot of sense to replace a concept in exchange for a good amount of money, but that would be giving up on a concept for the wrong kind of reason. The right kind of reason for replacing a concept is instead a reason of the kind that this concept itself is somehow better, on its own, not because it comes with cash attached. How to make this distinction between the right and wrong kind of reason more precise is notoriously difficult, and I use "for the purpose of inquiry" to make clear that it is supposed to be rational weighing the right kinds of reasons only, which concern how well the concept does in capturing what reality is like, not necessarily how well it does in benefiting us in other ways.

Third, the issue is whether the replacement would be rational *by my own lights*: whether I can accept, reflecting on and evaluating the reasons for and against, that C^* is better than C to employ for the purpose of inquiry. The issue is whether we ourselves should replace one of our concepts with another one, and to assess that properly we need to be able to be in a position to appreciate the reasons for and against doing this, and so we need to evaluate those reasons by our own lights. It is not enough for such a replacement to be rational to rely on a purely external conception of rationality, where how reasons are to be weighed is determined by something external, which is not necessarily accessible to us. So, if there is a sense of rational where it is rational to have the concepts God prefers us to have, without it being accessible to us why those are the right concepts to have, then it would not be rational by our own lights to switch to those concepts. For it to be rational by our own lights involves that the reasons in favor outweigh the ones against according to our own best standards for evaluating those

[6]The notion of an inescapable concept is related to, but different from, those of a 'conceptual fixed point' in the sense of [Eklund, 2015] or of a 'bedrock concept' in the sense of [Chalmers, 2011]. I will, in particular, discuss Eklund's views of the significance of what he calls 'conceptual fixed points' below.

reasons. That does not mean that everything we in fact accept is rational by our own lights, since it can be that we misapply the standards we in general endorse as the right ones. But it does mean that the reasons must appreciable by us in our present situation.

Fourth, the issue is *replacement in application*, not replacement in repertoire. It is arguable that it would always be a mistake to get rid of a concept and replace it with a different one instead. Getting rid of a concept from one's repertoire of concepts would always limit one's representational capacities, and thus is arguably always irrational. What matters for us instead is whether we should stop applying one concept in certain situations, and apply a different one in those situations instead. Even if we have both concepts in our repertoire, it makes sense to wonder which one we should apply in certain situations, in particular situations where we need to make a choice between one or the other. And this is our issue here: should we employ C^* in situations where so far we have employed C? It might be useful to have both concepts in our repertoire in general, but the issue remains which one we should use in particular situations.

I would furthermore like to contrast both inescapable concepts and hardwired concepts from what we can call *inevitable* concepts. A concept is inevitable if every thinker must have it. An inevitable concept thus is not merely a concept hardwired in this or that mind, but one that must be present in every thinker in order for them to have thoughts at all. If there are any inevitable concepts, then every thinker must have every one of them. But different thinkers can have different concepts that are inescapable for them: While thinker A has concept C_A inescapable for them, thinker B might not have that concept at all, but instead has a different concept C_B, which is inescapable for them. We will see this possibility illustrated shortly. Thus strictly speaking being inescapable is not a feature of concepts, but a relation between a concept and a thinker: a concept is only inescapable for a particular thinker who has that concept. Although it could seem that no concept will be inescapable, since one could always have strong reasons in favor of replacing that concept, there are some prima facie good examples of such concepts.

A first example of a candidate for being inescapable concepts are logical concepts. The question whether logical concepts are inescapable is closely tied to the question whether logic is rationally revisable. Can it ever be rational for us to give up our own logic in favor of a different one? Suppose I presently reason classically and endorse such reasoning as the right one, but someone suggests that this is a mistake and that I should reason with a different logic instead, say intuitionistic logic. Since intuitionistic logic differs from classical logic by whether the inference from $\neg\neg p$ to p is valid, the question comes down to whether I should replace my concept of classical negation, for which this inference is valid, with that of intuitionistic negation, for which it is not valid. For the sake of the example, I am here assuming a picture according

to which such inferences are constitutively tied to the relevant logical concepts, a view that is widely, but, of course, not universally shared. On this picture a proper change in logic comes with a change in logical concepts, in particular which concept of negation I should employ in reasoning in general.

Should I continue to employ classical negation and accept double negation elimination, or should I give up on classical negation in favor of its intuitionistic alternative? To answer this question I will have to assess the two options and determine which one wins out on the balance of the reasons for and against making the switch. And to do this I will use my best methods for evaluating options in general, including reasoning and inference. But these methods for evaluating anything include deductive reasoning, in particular reasoning with classical logic, which at present is still the logic I employ in my deductive reasoning. So, using this method I can easily conclude that I should continue to rely on double negation elimination, and that adopting intuitionistic logic would be a mistake. I can reason simply as follows: suppose $\neg\neg p$ is true. Then it must be that $\neg\neg p$. But then, reasoning classically, p must be the case, and thus p is true. So, the truth of $\neg\neg p$ guarantees the truth of p, and thus double negation elimination is truth preserving, and therefore valid. Thus I should not switch to intuitionistic negation. If I were to switch I would lose a valid form of reasoning, which would be a clear mistake.

This argument is a purely deductive argument from no premises, and it would thus seem to outweigh any reason in favor of switching that I might encounter. Thus it seems that not only ought I not to switch given the information I have so far, but I have an argument that I ought never to switch, no matter what. It would therefore be irrational by my own lights to revise my own logic, or so the argument in a nutshell.

Whether this argument is in the end correct is a tricky question. One worry is that there seems to be something question begging in continuing to reason classically to show that classical logic is best. But whether this is question begging in a sense that would invalidate these considerations is not at all clear. At most classical reasoning in favor of classical logic is *rule-circular*, but not *premise circular*. After all, I do not rely on a premise like "classical logic is the right logic", I merely use the classical rules, which I presently accept, to draw a conclusion about classical logic.[7] Second, the easiest way to show that classical logic is to be preferred is to show that the classical rules are all valid, something we can do using the classical rules together with the truth rules. But as is well known, the classical rules together with the "naive" truth rules are inconsistent, as is illustrated by various paradoxes like Curry's paradox or the liar paradox. It might be natural to take this to show that something is wrong with the truth rules,

[7]See [Dummett, 1978] and [Boghossian, 2000] for more on the status of rule-circular reasoning.

but the same issue arises with them as does with the classical rules for the standard connectives: we can show that the truth rules are truth preserving, and thus are to be preferred over any alternative, in the same rule-circular way. The truth rules can be shown to be valid just as the rest of the classical rules can be, with the same resources. But how can the argument that the classical rules and the truth rules are valid be any good when we can show that these rules allow us to derive anything? These issues quickly lead to thorny questions about our rational response to the paradoxes, whether it is rationally required to revise our own rules in light of paradoxes, if so, which ones, and if not, what other option should be favored. This is not the place to resolve these questions. I simply want to use the argument against the rational revision of logic as an example of an argument that would show that logical concepts are inescapable in our sense.Whether this is in the end correct in obviously a substantial further question, but the point remains that if logic is not rationally revisable then in effect logical concepts are inescapable: they cannot rationally be replaced with different ones by our own lights.

If this is correct and logical concepts are indeed not rationally replaceable, then this would illustrate a key feature of inescapable concept: it is not required that everyone has the same ones. It would be irrational for me to switch from classical logic to intuitionistic logic, but it would also be irrational for an intuitionist to switch to classical logic. Thus classical negation would be inescapable for me, while intuitionistic negation would be inescapable for an intuitionist. This shows that inescapable concepts do not need to be inevitable. One can be a thinker while not having a concept that is inescapable for someone else.

Arguments for inescapability are not restricted to more narrowly understood logical concepts. They can also be given for normative concepts like ought, which I take here in the sense of 'all things considered' ought, as opposed to something more narrow, like moral or prudential ought. Suppose I consider whether I should replace my concept of ought with an alternative ought*. Suppose further that I can recognize that this new concept would be extensionally different from my old one, in the sense that what I ought to do is different than what I ought* to do for certain cases. Then I seem to be in a position to reject the proposed replacement out of hand, and conclude that it would be irrational for me to give up my concept of ought in favor of ought*. I can simply reason as follow: if I were to switch, then I would think about what I ought* to do, and if my thinking would be effective, then I would do that. But what I ought* to do is not always what I ought to do, and so I might do things that I ought not to do even when I reason perfectly with the concepts that I would then have. Thus switching would lead me astray, and

therefore I should not switch.[8]

A similar, although more complicated, case can be made for the inescapability of the concept of agency. Could it be rational for me to replace the concept of agency with a different one, call it 'shmagency'? Here I am reflecting on what I should do: replace the concept or keep it? But doing something is engaging in agency; a conceptual connection between the concepts of doing and of agency make this vivid. I thus can't just replace the concept of agency and leave the rest alone. I must also replace doing with schmo-ing, whereby shmo-ing relates to doing as shmagency relates to agency. But then I face a conundrum: should I do something that will result in my not thinking about what I should do in the future? After all, I will employ a different concept than doing something, and so I won't think of myself as doing something any longer, and I won't reflect on what I should do after making the switch. I will consider what I should shmo, of course, but can I rationally conclude that I should make the switch and accept that I subsequently won't reflect any longer on what to do? Could it be rational for me to do something that would effectively eliminate future rational reflection what I should do? This question deserves serious consideration, but I hope even so far one can get a sense that an argument that this would not be rational might be forthcoming here. At least it would establish that there is a constraint on rationality that would prohibit such a concept replacement,and if so, then this would generate an argument that agency and doing are inescapable concepts: they can't rationally be given up.

Considering rational replacement of one concept with another is often not an isolated affair, as the case of agency illustrates: a replacement would also likely bring with it a replacement of other concepts that are conceptually closely related to the original concept in question. The concept of agency is connected to the concept of doing, and with it also to various other action concepts: reading, singing, etc.. A replacement of the concept of agency would thus engender a replacement of a family connected concepts. Concepts whose replacement would give rise to large-scale follow-up replacements are likely harder to replace rationally, but that by itself does not rule out replacement. But sometimes these conceptual connections go deep, so deep that it is arguable that a wholesale replacement of the whole group of connected concepts would be irrational. The concept of a fact is a candidate for being inescapable for this reason. Suppose we consider whether we should replace our concept of a fact with some alternative: fact*. Maybe something is suspicious about our concept of a fact, and a proposed alternative concept of a fact* is suggested as an improvement. Could it be rational for us to make the switch? In particular, could it be rational to make the switch for the purpose of inquiry? There is an argument tied to

[8]See [Eklund, 2017] and [Eklund, 2015] for a detailed discussion of thinking about alternative normative concepts.

conceptual connections that this could never be rational. Inquiry has a constitutive aim: to find the truth. But the truth is conceptually connected to the facts. It is a conceptual connection that if it is true that p, then it is a fact that p, and the other way round. Thus to replace the concept of a fact with that of a fact* would also suggest a replacement of the concept of truth with that of a corresponding truth*. But that would not be enough. We can't coherently think of ourselves as being engaged in inquiry with the goal of trying to find out what is true* and which facts* obtain. This would not respect the constitutive aim of inquiry, which is truth, not truth*. We would need to replace the concept of inquiry itself with its starred alternative: inquiry*. But then, could it ever be rational, while being engaged in inquiry, to replace the concept of inquiry itself with that of inquiry*, which does not have truth, but truth*, as its constitutive goal, and which does not aim at the facts, but the facts*? That any such replacement would bring with it a different constitutive aim for the activity associated with the starred replacement concepts might suggest that on reflection one could never rationally make this switch, since it would lead one badly astray in what one hopes to achieve in inquiry, the activity one presently is engaged in. Thus, arguably, the concept of a fact is inescapable for us.

All this now puts us in a position to argue, at least in rough outline, that the concept of number is also inescapable. The concept of number has tight conceptual connections to other, rather basic concepts like quantifiers. To illustrate, there being an F is conceptually connected, even conceptually equivalent, to the number of Fs being larger than 0. So, when I wonder whether there is an F, then this is conceptually equivalent to wondering whether the number of Fs is larger than 0. To replace the concept of number with a different one, number*, would thus suggest a larger replacement of a family of connected quantificational concepts. But if I were to make such a replacement, then this would affect what I would wonder about in the future. In particular, I would put myself in a position of no longer wondering whether there is a concept C^* which would be an improvement over one of my own concepts C. Considering this question, after all, is conceptually equivalent to considering whether the number of such concepts C^* is larger than 0. I can thus see that if I switched out the concept of number and with it the concepts that are conceptually tied to it, then I would not longer be in a position to consider just the kinds of questions that I am considering now, when I wonder whether I should replace a particular concept. Thus such a choice of concept replacement would limit my future reflection on what concepts I should replace, as well as my investigation in many other areas. I would thus constrain myself and limit my future rational reflection in a central area if I were to make that switch. And this is, arguably, irrational to do. It would be irrational to do something now that would knowingly limit my future rational reflection in a significant way,

in particular with regard to whether there are better concepts. Since the concept of number has close conceptual connections to quantificational concepts that are crucial for such rational reflection, it would be irrational for me to replace it with an alternative. Or so the outline of an argument that the concept of number is inescapable for us.

Obviously, all of these arguments and examples of candidates for being inescapable concepts deserve much further discussion.[9] I should also note that the examples given are not the only ones worth considering. Next in line for candidacy for being an inescapable concept would be concepts like belief, meaning, content, reason, and similar notions. The above examples were simply raised to motivate the idea that some concepts might be inescapable, which concepts they might be, and what kinds of arguments one might be able to give that they are indeed inescapable. But these examples also are not completely unrelated; there is a certain suggestive similarity between them. The concepts that were candidates for being inescapable where in some way or other related to the question that we raise when we ask whether we should replace these concepts with other ones. The question is one about what I ought rationally to do, and that involves the concepts of ought and of doing, and thus of agency. And if deductive, logical reasoning is tied to rationality, then indirectly the question also involves logical concepts, as well as that of number. Whether this similarity is a guide to other candidates for being inescapable concepts has to be left open for now, suggestive as it might be. Instead, I would in the following like to focus on the importance of inescapable concepts, if there are any, for metaphysics and for inquiry.

3 The significance of inescapable concepts

Suppose we have found a concept where we have good reason to think that it is inescapable. That would mean that we have found an example of a concept C and a good argument that C cannot be rationally replaced by a different one C^* by one's own lights for the purpose of inquiry. What of philosophical significance would follow from this?

There are essentially two ways to approach this: one more positive and one more negative. On the positive way, we could take a concept being inescapable as a reason to think that we have found the right concept to apply to the world, and thus we have found something out about what the world is like: we found a concept that is a perfect match for the world. This would be a real insight. Even though it leaves open what else we can conclude from this fact, at least

[9] I discuss the inescapability of logical concepts in [Hofweber, 2022b], that of agency in [Hofweber, 2022a] and that of facts in [Hofweber, 2023]. Eklund discusses thin normative concepts like ought in [Eklund, 2017] and [Eklund, 2015].

we made some good progress in finding out about the world. But there is also a more negative outlook on this situation. On this negative way of looking at it, nothing has been shown about how our concepts relate to the world. Instead, something has been shown about our rational situation: we cannot rationally move on from where we are; we are rationally stuck and forced to remain in place. But that might be so whether or not our concepts perfectly match the world. We simply cannot escape from having this concept while being rational. Rationality traps us into sticking with this concepts, for better or for worse.

Thus inescapable concepts can be seen in two rather different ways: either they trap us to remain in place, on pain of irrationality, even if it is a bad place to be in, or else the requirements of rationality tied to these concepts lead to an insight into the world. We can consequently distinguish between thinking of inescapable concepts as *traps* or as *insights*. Only if they are insights can we hope to draw any significant conclusions from them about the world, and only so understood can inescapable concepts help with the project of drawing metaphysical conclusions from thinking about our representations alone. If they are traps, then we can conclude something about our own situation and what rationality does not permit in it, but little can be concluded about how our minds relate to the world. The question thus is whether inescapable concepts, assuming there are any, are insights or traps. This is the first of two related issues that we need to make progress on.

The second issue concerns how reflecting on our concepts relates to the methodology of metaphysics and philosophy more generally. Here, too, there are two main ways of thinking of it. One is neo-Carnapian, inspired by view of Rudolf Carnap in [Carnap, 1956]. On this view, philosophy properly done does not ask questions of fact; that is the job of the sciences. Instead philosophy reflects on the concepts that are used by the sciences when they ask questions of fact, and it contributes to inquiry by making a proposal about what concepts might be improvements to the ones presently employed. Philosophy is thus closely tied to reflecting on which concepts we should have. The proper method of philosophy concerns conceptual engineering: finding improvements to our concepts and figuring out which concepts we should have. And this 'should' here is, of course, tied to rationality: it is tied to which concepts it would be rational for us to employ in inquiry. If we were to encounter any inescapable concepts, then this project would come to a roadblock. Since it would not be rational for us to employ a different concept instead, this would mean that no further philosophical work on this concept is legitimate. And this, again, can be either a good thing or a bad thing: the work has reached its goal and the concept is perfect, or somehow we are hindered from doing what we are supposed to do. Either way, the work of philosophy ends when inescapable concepts are encountered. There is nothing left

to do for philosophy here, assuming the Carnapian outlook on the subject.

This, in essence, is how Eklund sees the significance of what he calls 'conceptual fixed points': concepts that we are unable to engineer away, for some reason or other. Conceptual fixed points, in Eklund's sense, include hardwired concepts, inevitable concepts, and presumably also inescapable concepts. And as such they show the limits of what Eklund calls an attractive conception of the methodology of philosophy: improving our concepts (see [Eklund, 2015, 16]). Several other philosophers have recently defended a similar Carnapian conception of the role of reflecting on our concepts and the methodology of philosophy, including [Scharp, 2018] and [Thomasson, 2016]. On such a conception of philosophy, it is less ambitious than many have thought. It does not aim to find general facts about the world, or stand next to the sciences in the project of inquiry. Instead, it focuses on our concepts, leaving questions of fact to others. But it is also possible to closely associate the more ambitious, fact-finding conception of philosophy with inescapable concepts, and that is what we should turn to next.

A very different approach to the significance of inescapable concepts is a neo-Kantian one, inspired by the work of Immanuel Kant.[10] Kant's theoretical philosophy is a paradigm case of an attempt to derive metaphysical conclusions from thinking about our concepts alone. There is a distinguished class of concepts, the categories, to which the empirical world must conform. That this is so is found out in philosophy, by thinking about our concepts, and with no input on what the world otherwise is like. Crucially, for Kant the categories are arguably inevitable concepts in our above sense, and thus he can proceed via a transcendental argument: experience would not be possible at all unless the categories relate to the empirical world in a certain way. And that the categories relate to the empirical world in this way allows us to derive various metaphysical conclusions: that every change has a cause, that objects persist through changes, and so on. The neo-Kantian approach to inescapable concepts takes inspiration from this general strategy, but modifies it in crucial places. First, the neo-Kantian takes inescapable concepts, not the categories, to be the special concepts reflection on which leads to insights into what reality is like. Since inescapable concepts are not necessarily inevitable, the neo-Kantian will not proceed via transcendental arguments, but in a different way, one outlined below. Finally, the neo-Kantian does not have to hold that there is a division between the empirical, phenomenal world and the noumenal world, as Kant does. The metaphysical results that might be achieved can simply hold for reality full-stop.

[10]See [Kant, 1781]. This approach is neo-Kantian simply in the sense that it is inspired by Kant, not in the sense that it is part of the neo-Kantian movement of the late 19th century, involving such philosophers as Hermann Cohen and Ernst Cassirer, amongst many others. For a survey of the historical neo-Kantians, see [Heis, 2018].

How this neo-Kantian approach could possibly work as a strategy for metaphysical and philosophical progress is, of course, left open. But this is exactly the kind of project that would get us what we were looking for above: Not only would it show that certain questions of fact can be answered by reflecting on our ways of representing the world, it would show that the questions we asked are the right questions to ask. Putting these things together would give us a deep result in the sense spelled out above on page 8: an answer to a proper question to ask. If this could be pulled off, then we would answer a proper metaphysical question simply by thinking about our own representations alone, together with that they are inescapable concepts. The case of thinking about numbers illustrates this possibility. If the argument outlined above against the existence of numbers from thinking about our representations alone is correct, and if the concept of number is inescapable, then we answered a proper question. No other question in this neighborhood can rationally be accepted as a better question, although other questions might be good questions as well. The question whether von Neumann ordinals exist can also be a good question, but it isn't a better question. And so the question about numbers is not simply some defective and flawed question, but one that is worthy of our metaphysical attention. And thus answering it is a deep result, in our above sense. And it is a result about reality which we have gotten from thinking about our own representations alone, as the neo-Kantian approach had hoped.

4 The immanent stance

Even if all of the above is correct, the worry remains that this might not be as profound of an insight as it might seem. To illustrate the issue, consider the example of the inescapability of the concept of a fact. It might well be that our concept of a fact is inescapable for us, and that it would never be rational by our own lights to replace it with an alternative, fact*. Thus any results we can achieve by thinking about our representations of facts will be deep results for us. But other creatures could have started differently. They might have started out with the concept of a fact*, and with it with the related concepts of truth*, reality* as the totality of facts*, and inquiry*. For them, those concepts might be inescapable, and thus for them different results will be deep results. And this possibility shows that achieving a deep result might not be as important as it might have seemed, or so the worry. Thus different starting points can rationally force you to never give up certain concepts, and they can rationally force you to accept certain questions as the right ones. But since different people with different starting points can be required to accept different concepts as inescapable for them, not much can follow from this about the world. At best it shows that the notion of a deep result must be seen as one that

needs to be relativized to starting points. Different starting points can lead to different deep results, but since there is only one world, the whole issue reflects more on the starting points than on the world. Rationality here is like a trap: we are forced to have to think that these concepts are best, on pain of irrationality. The forces of rationality keep us stuck in place, just as causal forces keep a mouse in place in a mouse trap.

Inescapable concepts, according to this objection, are thus traps, not insights, and deep results only seem special to those trapped by their respective inescapable concepts. That it is rationally compelling that they are special results is part of what traps us. Showing that inescapable concepts lead to insight, not traps, is the final challenge that needs to be met to properly defend the neo-Kantian project outlined in this paper.

The solution to this problem is closely tied do having one of two general takes, or stances, on what we do in inquiry. There are two ways to conceive of the target of inquiry, which are connected to how seriously we should take our conceptual starting point and how seriously we should take the questions we ask when we initiate such inquiry. One of these stances takes what needs to be achieved in inquiry as being determined by an initial question. That question determines what needs to be done: it sets the *goalpost* for this particular part of inquiry. It could be something general like "what is reality like in general ways?" or "what should I do?", or something more specific like "how does communication work?" With such a question we set the goalpost that we need to reach. We set the goalpost with our initial question, formulated using our own concepts, the ones we happen to have, with all their benefits and problems. Once the goalpost is set it can, of course, be adjusted or moved as things develop. We might accept new questions or concepts as being better, and then use those to set a different goalpost instead, having realized that the old question was flawed in some way. The goal of our project of inquiry has then shifted from trying to answer the old question to trying to answer the new, better one. But even if so, we still set the new goalpost with a question, possibly formulated with new concepts. I will call this approach to philosophical inquiry the *immanent stance*. It contrasts with the *transcendent stance*, which takes the goalpost for inquiry not being set by us at all, in particular not by a question we ask. Instead, it takes the goalpost to be already there, waiting for us to meet it. The goalpost simply is reality, or some aspect of reality, something already there before we got started with inquiry. The goalpost so understood is thus not set by us with a question, formulated with our concepts. It is there already, calling upon us to meet it.[11]

[11] Good example of philosophers taking the transcendent stance are Matti Eklund in [Eklund, 2017] and Ted Sider in [Sider, 2011].

Let us take the immanent stance seriously for a moment. On it we start out with some concepts or other, and using those concepts we ask questions about the world, thereby setting the goalposts for various parts of inquiry. We then aim to improve our concepts, in part so that we can ask even better questions. Such improvements need to be carried out by us, and they need to be assessed by our own lights: they need to make sense as improvements to us, as evaluated according to the best of our present abilities. But at first we have the concepts we have, and we have the standards for evaluating the reasons for changing those concepts that we have. We have to start somewhere, and we need to improve from there. That we could have started somewhere else to begin with, with different concepts, does not by itself undermine our present use of our present concepts. We have to consider and evaluate alternatives from our present position. We don't need to first justify our starting point and prove it to be better than alternatives before we can legitimately employ the concepts we have. That would seem to be impossible, since we would have to provide such a justification without using our concepts. Since no justification of our conceptual starting point seems possible, it seems right to take ourselves to be at first defeasibly entitled to employ the concepts we have. That entitlement is only defeasible in the sense that we might well find out quickly that the concepts we have are flawed and other ones are better even by our own lights.[12] Still, at first we are entitled to use the concepts we have. The question remains whether the mere existence of alternative concepts or the mere existence of alternative conceptual starting points defeats our entitlement to our own concepts. My own judgment here is that he answer is 'no'. By analogy, the mere existence of other belief-forming mechanisms in perception does not defeat our own perceptual beliefs. If we also had reason to think that alternative ways of forming perceptual beliefs were better, then this would be different. But there merely being alternative belief-forming mechanisms is not enough for defeat. And the same is true for alternative concepts and alternative conceptual starting points: there simply being alternatives is not enough to rob us of our entitlement to use the ones we have. What would rob us of this entitlement would be some good reason that our own concepts are inadequate or inferior to the alternatives.

These two aspects of the outlined position are congenial with each other, but strictly speaking independent: First, there is the immanent stance, which concerns how the goalpost of inquiry is set by our questions, and not antecedently there. And, second, there is our defeasible entitlement to employ the concepts we find ourselves having at the start. One could hold one without the other, but it is natural, although not required, that one augments the immanent stance with a

[12]There are some clear similarities here to a dogmatist position about our perceptual entitlements in the sense of [Pryor, 2000].

view about our defeasible entitlement to our own concepts, as I am inclined to accept, although this, just as the immanent stance itself, is certainly debatable. Whether we should take the immanent stance, and in addition accept a defeasible entitlement to our own concepts at the start are issues that are not easy to resolve. But there is much to be said for the immanent stance. It doesn't take the world to demand a description with certain concepts, but it takes all such demands to ultimately come from us. Not, of course, the correctness of the description, that is settled by the world, but only the demand for a description itself. That there is such a demand at all on us comes from our engaging in inquiry and our project of finding out what reality is like. Without us engaging in inquiry, it would seem to be fine to remain silent, and not describe the world at all. The world by itself doesn't legitimately demand anything of us, including that we use these or those concepts. But once we engage in inquiry, we are bound to live up to the standards of success of inquiry, as they were set by our placing the goalpost with our initial question. We set the goalpost with our questions using our concepts, and with this we generate the demand that we meet it. The world determines whether we met it, but it does not determine what we need to do and what activities we need to engage in. And on this way of understanding our own situation, the goalpost for our activity is properly stated as representing reality, not reality*. Taking our own starting point and our own questions seriously breaks the tie.

The immanent stance augmented by a defeasible entitlement to our own concepts makes clear that inescapable concepts are insights, not traps: We start with some concepts or other which we are defeasibly entitled to employ. Using these concepts we then ask questions that set the goalpost for what we hope to achieve in inquiry, while at the same time we also try to improve our concepts to ask better questions. We pose those questions, and evaluate alternative concepts, to the best of our abilities by our own lights, using the standards for weighing reasons that we at present accept as the best ones. These standards in turn are not written in stone, but improving them must again make sense by our own lights. If it turns out that there is a concept which is inescapable for us, then we can never rationally replace it with a different one, by our own lights, and thus we must conclude that this is already the best concept to have in this neighborhood. When we find out that other creatures can have different inescapable concepts, then we should also conclude that they are the best concepts for them to have by their own lights. For them, fact* might be inescapable, while for us fact is. But when we try to assess whether their concept is just as good as ours, and whether their questions about reality* are just as good as ours about reality, then it is us asking that question, using our concepts, which we are defeasibly entitled to use, and engaging in our activity in finding out what is the

case and what reality is like. So understood the answer must be that their questions and their concepts are not as good as ours. When we evaluate alternative concepts, we wonder which ones are better, and better here means better for the purpose of inquiry. That is the standard for being better that is operative when we ask the question about which concepts are better. True enough, their concepts are better for inquiry*, which aims at reality*, not at reality. But that is not the issue at hand when we ask which concepts are better. The goalpost for that question is which ones are better for inquiry, and better for representing reality. For the relevant question, the one that sets the goalpost that is presently operative, the answer favors our concepts. And unless our entitlement to these concepts is defeated, we are entitled to take this to be the right goalpost articulated with the right concepts. Simply because there are alternative conceptual starting points does not defeat our entitlement to our own concepts, and so we are entitled to hold that we employ the right concepts.

All this seems to now invite a meta-question: sure, our concepts are better for inquiry, and theirs for inquiry*, but what makes inquiry so much better than inquiry*? But here, too, it is us who ask the meta-question, and so understood the answer is again obvious: inquiry aims at the truth and at reality, while inquiry* does not. It instead aims at something that we have little reason to pursue. There might be good reasons* to pursue it, but those have no pull on us, although they do have a pull on them. The question was what we should pursue and what we should believe. And the answer to that questions is what we have most reason to do and what is true.

This means that we are not trapped by our inescapable concepts. The issue is just the same as above: it is me who is asking whether I am trapped by an inescapable concept, i.e. whether the forces of rationality constrain me to never give up that concepts, even though in a sense I should be moving on. But that goalpost, set by that question, is easily reached by answering the question in the negative. I can conclude that an inescapable concept gets me to exactly where I should be. The fact that other creatures can have other concepts that are inescapable for them does not undermine this. It does not take away my entitlement to applying my own concepts, and as long as I remain entitled to apply my own inescapable concepts, I can reason conclusively that they are the proper ones to apply. Those who start out with fact* as an inescapable concepts for them are indeed trapped by that concept. They cannot see that they should be aiming at truth and reality and that they should engage in inquiry. They are trapped into sticking with inquiry* aiming at the truth*, and so much the worse for them. But my own inescapable concepts cannot be traps, and the immanent stance brings out why not.

When I myself ask whether I am trapped by my inescapable concepts, then the answer to

that question must be that I am not trapped. It is once again a kind of harmony between the question that we ask with the concepts that we have and the answers we give with just those concepts. Focusing on the question that we initially asked not only highlights how the language-metaphysics gap can be bridged, as discussed above, it also makes clear why inescapable concepts lead to insights, not traps.

5 Conclusion

Our initial problem was whether one could ever reach any substantial metaphysical conclusions merely from reflecting our our own concepts. Although the language-metaphysics gap seems to suggest that this is never possible, I argued above that it is possible after all. Since the questions we ask involve our own concepts, it can in principle be that reflecting on the concepts used in the question guarantees that this question must have a certain answer, even if that answer is not an analytic truth. The example of the non-existence of natural numbers illustrated this possibility. Nonetheless, the real issue remained, since so far we were only able to answer the question we happen to ask, not necessarily the question we should have asked in the first place, using the concepts we should use. Although it seems impossible to know which concepts one should use without knowing what reality is like, I argued that there are possible exceptions to this with the case of inescapable concepts. If a concept is inescapable for us, then we cannot rationally accept by our own lights that there is a better one with which it should be replaced. And if so, then any answer to a question asked with such inescapable concepts in a central place would lead to a deep metaphysical result. If the argument briefly outlined for the inescapability of the concept of number is correct, then the non-existence of natural numbers would be such a deep result, established by reflecting on our concepts alone. Furthermore, appreciating the significance of the immanent stance in philosophical inquiry shows that such concepts are not traps, and thus that deep results are real insights into what reality is like. These three parts together are the key components of the neo-Kantian vision for making progress in metaphysics: bridging the language-metaphysics gap, inescapable concepts, and the immanent stance. I would like to call this approach to answering questions in metaphysics *immanent metaphysics*, and at the same time stress that it is just one of several ways in which one can make progress in metaphysics. No claim is being advanced that all metaphysics should proceed this way, only that it is a fruitful strategy for achieving results. On this approach considerations from the philosophy of language as well as from epistemology have a constructive role in metaphysics, not just a corrective one, in the sense of page 1. The philosophy of language will be central to uncover certain features of our representations that allow us to bridge the language-metaphysics gap, while epistemology

will determine which concepts are inescapable. Metaphysics, the philosophy of language, and epistemology need to work closely together to allow for progress in immanent metaphysics. Naturally it remains to be seen how fruitful this approach can be in the end, but I hope to have at least made the case that this approach is more promising than it might appear.[13]

References

[Boghossian, 2000] Boghossian, Paul (2000). Knowledge of logic. In Boghossian, Paul and Peacocke, Christopher, editors, *New Essays on the A Priori*, pages 229–254. Oxford University Press.

[Carnap, 1956] Carnap, Rudolf (1956). Empiricism, semantics, and ontology. In *Meaning and Necessity*, pages 205–221. University of Chicago Press, 2nd edition.

[Chalmers, 2011] Chalmers, David (2011). Verbal disputes. *The Philosophical Review*, 120(4):515–566.

[Dummett, 1978] Dummett, Michael (1978). The justification of deduction. In *Truth and other Enigmas*, pages 290–318. Harvard University Press.

[Dummett, 1991] Dummett, Michael (1991). *Frege: philosophy of mathematics*. Harvard University Press.

[Dyke, 2008] Dyke, Heather (2008). *Metaphysics and the Representational Fallacy*. Routledge.

[Eklund, 2015] Eklund, Matti (2015). Intuitions, conceptual engineering, and conceptual fixed points. In Daly, Chris, editor, *The Palgrave Handbook of Philosophical Methods*, pages 363–85. Palgrave Macmillan.

[Eklund, 2017] Eklund, Matti (2017). *Choosing Normative Concepts*. Oxford University Press.

[Frege, 1884] Frege, Gottlob (1884). *Die Grundlagen der Arithmetik: eine logisch mathematische Untersuchung über den Begriff der Zahl*. W. Koebner.

[Heis, 2018] Heis, Jeremy (2018). Neo-Kantianism. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2018 edition.

[Hodes, 1984] Hodes, Harold (1984). Logicism and the ontological commitments of arithmetic. *Journal of Philosophy*, 81:123–49.

[13]For very helpful discussions and/or feedback on an earlier draft I am indebted to Dave Chalmers, Matti Eklund, Dave Estlund, Markus Kohl, Matt Kotzen, Ram Neta, Jim Pryor, Kevin Scharp, Alex Worsnip, two anonymous referees, as well as audiences at Brown, Cambridge, Düsseldorf, St. Andrews, Syracuse, Tübingen, and UNC.

[Hofweber, 2005] Hofweber, Thomas (2005). Number determiners, numbers, and arithmetic. *The Philosophical Review*, 114(2):179–225.

[Hofweber, 2016] Hofweber, Thomas (2016). *Ontology and the Ambitions of Metaphysics*. Oxford University Press, Oxford, UK.

[Hofweber, 2019] Hofweber, Thomas (2019). Idealism and the harmony of thought and reality. *Mind*, 128(511):699–734.

[Hofweber, 2022a] Hofweber, Thomas (2022a). The inescapability of the concept of agency. Unpublished manuscript.

[Hofweber, 2022b] Hofweber, Thomas (2022b). The unrevisability of logic. In Hawthorne, John and Turner, Jason, editors, *Philosophical Perspectives*, volume 35. Wiley-Blackwell.

[Hofweber, 2023] Hofweber, Thomas (2023). *Idealism and the Harmony of Thought and Reality*. Oxford University Press, Oxford, UK. Forthcoming.

[Kant, 1781] Kant, Immanuel (1781). *Kritik der reinen Vernunft*. Johan Friedrich Hartnoch, Riga.

[Moltmann, 2013] Moltmann, Friederike (2013). Reference to numbers in natural language. *Philosophical Studies*, 162(3):499–536.

[Pryor, 2000] Pryor, James (2000). The skeptic and the dogmatist. *Noûs*, 34(4):517–549.

[Scharp, 2018] Scharp, Kevin (2018). Philosophy as the study of defective concepts. In Burgess, Alexis, Cappelen, Herman, and Plunkett, David, editors, *Conceptual Engineering and Conceptual Ethics*, pages 396–416. Oxford University Press.

[Sider, 2011] Sider, Theodore (2011). *Writing the Book of the World*. Oxford University Press.

[Thomasson, 2016] Thomasson, Amie L. (2016). What can we do when we do metaphysics? In D'Oro, Giuseppina Guiseppa and Overgaard, Soren, editors, *The Cambridge Companion to Philosophical Methodology*, pages 435–458. Cambridge University Press.

[Wright, 1983] Wright, Crispin (1983). *Frege's Conception of Numbers as Objects*. Aberdeen University Press.