Reliable deduction

Dedução confiável

Luis Rosa[1]

Abstract: In this paper I address the question of what makes deduction reliable (when it is reliable). A successful answer to that question would help us understand how deduction can expand and generate knowledge. I explore two answers to the main question. The first one makes use of the notion of formal-logical entailment, whereas the second one makes use of the notion of metaphysical entailment. The latter is found to be superior to the former, in that it allows us to explain the reliability of a wider class of reliable types of deduction.

Keywords: Reasoning, inferential knowledge, reliability.

Resumo: Neste artigo trato da questão sobre o que torna uma dedução confiável (quando ela é confiável). Uma resposta satisfatória a tal questão nos ajudaria a entender como dedução pode expandir ou gerar conhecimento. Eu exploro duas respostas a tal questão. A primeira faz uso da noção de acarretamento lógico-formal, enquanto que a segunda faz uso da noção de acarretamento metafísico. A última é superior à primeira, pois nos permite explicar a confiabilidade de uma classe mais ampla de deduções.

Palavras-chave: Raciocínio, conhecimento inferencial, confiabilidade

1.

Here are some examples of reliable deductive inferences:

1. I know that *Ana was born in Durban*, and also that *Mosi was born in Durban* (and I know that *Durban is a city*). I infer on that basis that *Ana and Mosi were born in the same city*.

2. I hypothesize that *Judy is a feminist philosopher*; under the scope of that assumption, I conclude that *Some feminists are philosophers*. On the basis of that piece of suppositional reasoning, I infer that *If Judy is a feminist philosopher then some feminists are philosophers*.

3. I believe that *Natural science is grounded on empirical evidence*, and also that *Mathematics is not grounded on empirical evidence*. On that basis I infer that *Mathematics is not a natural science*.

---

[1] Assistant professor University of Cologne <fsopho@gmail.com>

There is more than one way to explain what is 'deductive' about these inferences.[2] It is possible to draw a distinction between deductive and non-deductive reasoning even if we agree with Gilbert Harman (1986) that the existence of deductive *arguments* by itself does not show that there is such a thing as deductive *reasoning*, understood as reasoned change in view––if only because the rules of deductive argumentation are not rules about how one should revise one's beliefs.

Consider my inference that *Left-wing political agendas are bad*, as made on the basis of my belief that *The adoption of left-wing political agendas e.g. in North Korea and Serbia brought about bad consequences*. I may change my mind about left-wing agendas later without changing my mind about what happened in North Korea, Serbia, etc.: I still take the results of their political programs to be bad ones, but now I think that some left-wing agendas may turn out to be good. My rejection of the original conclusion does not lead me to change my mind about the claims I have used as grounds for it in the first place (and I was not disposed to do so at the time of the original inference either). Furthermore, I am not liable to epistemic criticism for changing my mind in this way: I can coherently believe now that *Some left-wing political agendas are good* even though I still maintain my original opinion about the bad things that happened in North Korea, Serbia, etc.

Compare that to my inference in 1 above. If I learn now that actually *Ana and Mosi were not born in the same city*, that will prompt me to revise one of the beliefs I have used as grounds for my original conclusion. I was disposed to do so at the time of the original inference. E.g. perhaps I will revise my belief that *Ana was born in Durban*, or maybe my belief that *Mosi was born in Durban* (that depends on what more exactly motivated me to give up on my belief that *Ana and Mosi were born in the same city*). Furthermore, if I still maintain my belief that they were both born in Durban *despite* having learned that they were not born in the same city, I am liable to criticism. These future/dispositional doxastic developments and their respective epistemic properties indicate that the ties between my grounding beliefs and my conclusions in these two cases

---

[2] As I am using the notions of 'inference' and 'reasoning' here, inference is a *species* of the *genus* reasoning. Inference is a type of reasoning in which the reasoner ends up forming a new belief/high-credence attitude, or rather she ends up giving additional support to a belief/high-credence attitude that she already had even before making the inference.

are of a different kind. To distinguish them, we might say that my original inference about Ana and Mosi was a *deductive* inference, whereas my inference about left-wing political agendas was not.

Another way of drawing the distinction between deductive and non-deductive reasoning is by making reference to the reasoner's disposition to form (new) beliefs of a specific kind. E.g. one might suggest that the piece of suppositional reasoning I perform in example 2 is a process of deduction because I am disposed to believe (upon consideration) that *It is impossible for 'Judy is a feminist philosopher' to be true and 'Some feminists are philosophers' false* (or something along these lines). In contrast to that, when I conclude that *Amanda will celebrate* under the scope of my assumption that *Amanda will get the job*, I am nevertheless disposed to believe that it is possible––even though it is unlikely––that *Amanda will not celebrate despite getting the job* (say, because she could get sick in the meantime). Accordingly, the latter piece of reasoning would not be a process of deduction.

As an alternative to that classificatory strategy (or in addition to it), one might also make reference to *actual* states or attitudes of the reasoner (instead of dispositions to hold doxastic attitudes) that suitably connect the contents involved in her reasoning process. Boghossian (2014), for whom inferring necessarily involves the reasoner *taking* her premises to give support to her conclusion, seems to adopt this position. The way in which the reasoner takes her premises to be related to her conclusion when she performs a piece of *deductive* reasoning differs from the way in which she takes her premises to be related to her conclusion when she performs a piece of *non-deductive* reasoning (Boghossian 2014, p. 5). Whatever the taking-attitude might turn out to be here, that is yet another way of drawing the relevant distinction.

Perhaps there are yet other ways of specifying what is distinctive about deductive reasoning. Notice that we are indeed drawing a distinction between different types of *psychological processes* here, instead of different types of *arguments* (relatedly, the ways of drawing the relevant distinction mentioned above make no use of the actual alethic properties of arguments, e.g. logical entailment as opposed to probabilistic support––only to the reasoner's doxastic dispositions or her take on the relationship between 'premises' and 'conclusion').

Some types of deductive reasoning are *reliable*, some are not. But what makes it the case that some types of deduction––understood as types of cognitive processes––are reliable or truth-conducive? Or: what are the *properties* of reliable types of deduction that are responsible for their reliability, thus distinguishing them from the unreliable ones? That is the question I am going to address in this paper.

2.

Before addressing that question, however, let me first describe some important *epistemic roles* of deduction. Deductive inferences can only play these roles if they are reliable, which in turn motivates us to pursue the matter at hand.

First, deduction allows us *to expand our knowledge*, or to ground new knowledge on the basis of knowledge we already have. It allows us to put two-and-two together and discover truths that were hidden in the pool of information previously available to us––e.g. as when a detective discovers that *The butler did it* by deducing it from the available evidence, or when a mathematician proves rather surprising results on the basis of apparently trivial axioms, etc.

In order to make sense of the claim that we gather *new* knowledge through deduction we have to take the contents of doxastic attitudes to be relatively *fine-grained*. All necessary truths, or so standard epistemic logics would have it, are known by us all the time (see Meyer 2003). You know that ϕ when ϕ is true in every epistemically accessible world––e.g. every world in which your evidence is identical to the evidence you actually have. Since necessary truths are true in all worlds, you know them no matter what world you are in. In general, you always know everything that follows from what you know (that is, your knowledge is closed under entailment).

Standard modal epistemic logics provide a systematic and rigorous framework for studying epistemic notions, and they can be fruitfully used to model cognitive processing or information exchange phenomena. But these logics involve just as many idealizations as any other interesting mathematical model. It turns out that one of these idealizations––closure of knowledge under entailment––does not quite allow us to capture the sense in which we make discoveries and expand our knowledge through deduction. Granted, one can modify the classical

modal logic framework in such a way as to better represent this aspect of rational cognition (e.g. see Duc 1997 and Jago 2014 for a more recent approach), but that only acknowledges the need to take more fine-grained doxastic attitudes into account.

So maybe one should take the contents of doxastic attitudes to be Russellian propositions, or Fregean senses, or perhaps interpreted sentences in the language-of-thought (see Richard 1997 for different accounts of so-called 'propositional attitudes' in general). I do not have space to assess these different approaches here, however. Suffice it to say that they are acceptable at least insofar as they allow us to make sense of the claim that we can gather *new* knowledge through deduction.

The claim that reliable deduction is used to increase our knowledge is consistent with the claim that *sometimes it fails* to transmit epistemic status. Perhaps deduction fails to generate knowledge of the falsity of fancy skeptical hypotheses (e.g. that *I am a deluded brain in a vat*) on the basis of one's knowledge of ordinary empirical claims (e.g. that *I have hands*), where the latter ones are also known to be incompatible with the former ones (e.g. *If I have hands I am not a brain in a vat*). Transmission failure here may be due to the fact that the subject's actual cognitive state is indistinguishable from that in which the skeptical hypothesis is true (see Wright 2002 for a defense of this idea and Brown 2004 for a critical discussion). Or perhaps knowledge-transmission fails in these cases because possessing conclusive reasons is required for knowledge: I do have conclusive reasons to believe that *I have hands*, and yet I fail to have conclusive reasons to believe that *I am not a brain in a vat* (see Dretske 1971).

A less controversial claim is that the risk of error may accumulate through deduction. As an example, consider a subject who deduces a big conjunction each conjunct of which is believed by her with less than maximal warrant or justification, but enough so as to constitute knowledge (as illustrated in the *preface paradox* situation). E.g. I have knowledge about the properties of physical objects, about the past and about other people's mental states. Each of these beliefs is risky in its own way, in the sense that it is still possible for them to be false given my evidence/cognitive situation. But when I put all these slightly risky claims together and infer their conjunction, I end up forming a highly risky belief––if only because the big conjunction inherits the risk of each of its

conjuncts. (See Lasonen-Aarnio 2008 for a defense of the view that single-premise deductive inference can be risky as well).

Finally, reasoners may also have *rational self-doubt* about their inferential performances (I borrow the term 'rational self-doubt' from Schechter 2013). E.g. I may perform a competent deduction on the basis of things I know with certainty, but fail to know the conclusion because I am aware that I was very tired and sleepy when I performed the inference. Despite the fact that my deduction was competent and I still maintain my belief in the conclusion, that belief arguably fails to constitute knowledge.

All these cases pose interesting challenges to the epistemology of reasoning. But, again, they do not conflict with the claim that deduction expands or increases our knowledge in many other cases. What they show (if they do) is only that *deduction sometimes does not* increase our knowledge.

3.

I have just described one of the epistemic roles of deduction: it gives us new knowledge on the basis of knowledge we already have. Call that property of deduction 'knowledge-expansion'. Moreover, deduction can arguably also *generate* new knowledge––knowledge that is not based on any previous (propositional) knowledge. This phenomenon is nicely illustrated through example (2) presented above: I believe that *If Judy is a feminist philosopher then some feminists are philosophers* on the basis of a piece of suppositional reasoning. My conclusion in this case is not grounded on any set of premise-beliefs, but rather on a previous process of reasoning with assumptions. This is not at all uncommon: we frequently engage in 'what if' activities and reach new conclusions as a result of exploring the consequences of our hypotheses.

The process of drawing conclusions on the basis of pieces of suppositional reasoning and its property of generating new knowledge or justified belief has already been explored elsewhere in the literature––e.g. see Schechter (2010), Balcerak Jackson & Balcerak Jackson (2013) and Dogramaci (2016). Some have characterized the fact that we can form beliefs through processes of this kind as some sort of *byproduct* of our more down-to-earth inferential processing of doxastic input (see Williamson 2007, p. 162, and also Rumfitt 2008, p. 63). That I am equipped to infer that *Some feminists are philosophers* on the basis of my

belief that *Judy is a feminist philosopher* has the consequence that I am also equipped to infer that *If Judy is a feminist philosopher then some feminists are philosophers* from my armchair.

It is not just beliefs toward *conditionals* that we can form in this way, even though the examples involving beliefs in conditional claims are perhaps the most intuitive ones. As an alternative example, consider the following case: I hypothesize that *1023 is a prime*; under the scope of that assumption I conclude that *It is not the case that 2013 is not a prime*; on the basis of that piece of suppositional reasoning, I finally infer that *It is not the case that 1023 is a prime and not a prime* (after all, the assumption that *1023 is a prime* turned out to be incompatible with the assumption that *1023 is not a prime*).[3]

I have introduced the term 'knowledge-expansion' above to refer to the property of deduction to give us new knowledge on the basis of other things we know. Now let me introduce the term 'knowledge-generation' to refer to the property of deduction to give us new knowledge on the basis of pieces of suppositional reasoning. These properties correspond to two general types of inference: *conditional* and *unconditional* types of inference respectively. Conditional inferences take (at least some) beliefs as input, whereas unconditional inferences only take pieces of reasoning (sub-processes of suppositional reasoning) as input. Both of them output beliefs. This distinction is not supposed to imply that there cannot be inferences whose input is a mix of beliefs and pieces of suppositional reasoning––there surely are inferences like that (see Dogramaci 2016 for a similar observation). But it is also clear that these are *conditional* types of inference, for their conclusions are still grounded on other beliefs.

One might suspect that what I am calling an 'unconditional inference' here is only a conditional type of inference in disguise. But if e.g. my belief that *If Judy is a feminist philosopher then some feminists are philosophers* is actually grounded on another belief––instead of being grounded on a piece of suppositional reasoning––what would the content of that grounding belief be? It better not be a conditional whose antecedent is the content of my initial assumption and whose consequent is the

---

[3] Some readers will perhaps associate these examples with logical proofs from the empty set of premises in a proof-system. But presumably the informal reasoning in natural language comes first––proof-theory can be seen as an attempt to formalize *it*, or to abstract away from its irrelevant details. One does not need to be a proof-theorist to perform these inferences, and one does not need to explicitly employ rules of derivation to do that either.

conclusion I have drawn under the scope of that assumption. For in that case I would be using a belief in a content of the form *If φ then ψ* as a reason to believe that *If φ then ψ*. And surely no such viciously circular inference would give me knowledge of that conditional. Alternatively, one might suggest that the content of my grounding belief is something like *I have competently/validly derived 'Some feminists are philosophers' under the scope of my assumption that 'Judy is a feminist philosopher'* (see Wright 2014, p. 29 for a proposal along these lines and, relatedly, also Valaris 2016, p. 4). But that in turn only invites the question of how a reasoner can have knowledge of the fact that she performed the relevant piece of suppositional reasoning in a competent or valid way. Suddenly an otherwise ordinary reasoner needs to have knowledge of competence in reasoning and validity. There is no good reason to think that a reasoner without any of these theoretical types of knowledge could not engage in the practice of unconditional inference and gain knowledge through it.

4.

So much for (some of) the epistemic roles of deduction. I will now get back to the question: What makes a type of deductive inference reliable? Answering that question should help us understand how knowledge-expansion and knowledge-generation through deduction are possible. Given the distinction I have made above between two general types of inference, that question can now be split in two:

(1) What makes a *conditional* type of deductive inference reliable?
(2) What makes a *unconditional* type of deductive inference reliable?

Let me now say something about what is and what is not expected from the answers to these questions. First, good answers to (1) and (2) should allows us to explain *why* certain types of deduction are reliable. The envisioned type of explanation is not a *causal* explanation, however––e.g. the *explanans* need not antecede the *explanandum* in time. It is rather a *metaphysical* type of explanation: it makes reference to the properties of reliable types of deduction *in virtue of which* they are reliable.

One could say here that possession of the target properties on the part of reliable types of deduction is what *grounds* their reliability (on the notion of grounding in metaphysical explanation, see Dasgupta 2014). So

the fact that a type of deductive inference is reliable is supposed to follow from the fact that it has the relevant properties. Assuming that the truth-conduciveness of a reliable process-type is *modally robust*, the properties mentioned in that explanation should entail not only that instantiations of the relevant inferential processes *do actually* preserve truth or generate true beliefs––but also that *they would still do so* in certain counterfactual situations (e.g. scenarios belonging to some fixed class of possibilities, or perhaps scenarios that are in some sense 'close' to the actual world––see Goldman 1988, pp. 61-63 and also Baumann 2009 for critical discussion).

Second, we must clearly distinguish between reliability-entailing conditions, or conditions that entail that certain types of inference are reliable (hopefully also *explaining why* they are reliable), and warrant-entailing conditions, or conditions that entail that certain beliefs are warranted/justifiably held when they are formed by means of inference. This is best seen in the case of conditional inference. Consider my inferential belief that *I am in Europe*, as maintained on the basis of my belief that *I am in Germany and Germany is in Europe*. In order for my inferential belief that *I am in Europe* to be warranted or justifiably held, my grounding belief that *I am in Germany and Germany is in Europe* must itself be warranted or justifiably held. But surely this is not required for the type of inferential process I instantiate in this case to be reliable (this point was also made by Goldman 1979, pp. 13-14). The reliability of a belief-forming process does not depend on the epistemic status of my own beliefs.

This observation may sound quite trivial, but it actually carries theoretically relevant information. The present investigation is about the reliability-entailing properties of certain kinds of *processes*––more specifically, it is about those properties of reliable types of deduction that *explain why* they are reliable (in the metaphysical sense of explanation). Accordingly, the adequacy of an answer to question (1) should not be affected by the fact that warrant or justification fails to be transmitted through deductive inference in those cases I have mentioned in Section 2. E.g. suppose it is claimed that possession of property *P* explains why certain conditional types of deduction are reliable. Suppose, furthermore, that that view entails (among other things) that believing a conjunction on the basis of (one's beliefs toward) its conjuncts is a reliable process-type––for the process of believing a conjunction on the basis of (one's beliefs

toward) its conjuncts has the relevant property *P*. One cannot then use the fact that warrant sometimes fails to be transmitted through that type of inference as an objection to the claim that possession of property *P* on the part of a type of deduction explains why it is reliable. For the original claim is not about transmission of warrant through deduction *per se*––it is rather about one (and just one) of the necessary conditions for transmission of warrant through deduction: the reliability of the inferential process.

Third, one might expect that good answers to our questions (1) and (2) should also address the *generality problem* for reliabilism (see Conee and Feldman 1998) concerning inferential justification. But one should not expect that. Consider my inference that *There are logicians*, as made on the basis of my belief that *Ruth is a logician*. To a first approximation, one might tentatively describe the type of inference I instantiate in this case as follows: I transition from a belief in a content of type α *is* Φ to a belief in a content of type *There are* Φ*s* (α being a placeholder for singular terms and Φ a placeholder for predicates).

But this is at best a sketch of a description of the relevant type of inference. There are several ways of filling in the missing details here. Maybe in order for a singular content like *Ruth is a logician* to be a token of the relevant content-type α *is* Φ it must be the case that *Ruth* refers to an individual. So, e.g. if *Pegasus* fails to refer to an individual, then *Pegasus is strong* is not a token of α *is* Φ under that way of individuating the relevant content-types. Or perhaps there are restrictions on the existential quantifier: an existential claim is a token of *There are* Φ*s* only when its existential quantifier ranges over *concrete* things only, e.g. claims involving quantification over *abstracta* do not count as instances of the relevant content-type. So *There are prime numbers* is not an instance of *There are* Φ*s* under this way of individuating the relevant content-types. (To say that *There are prime numbers* is not a token of *There are* Φ*s* here is, among other things, to say that I would not instantiate that type of inference if my input-belief were e.g. a belief that *3 is a prime number* instead of a belief that *Ruth is a logician*).

The alternative ways of individuating content-types correspond to different types of inference. Which one of these types of inference is the one my actual cognitive process is a token of? It is clear that types of inference are *underdetermined* by particular inferences (process-tokens),

and that different inferential process-types will have different degrees of reliability. So how can we ever determine whether my particular inference was an instance of a reliable process? That surely is a hard problem––but it is not a problem for theories about *what makes deduction reliable* (my concern here). Theories of the latter kind are not necessarily committed to any claim to the effect that a reasoner does reason reliably in this or that particular situation (or class of situations). Furthermore, in order to metaphysically explain the reliability of reliable types of deduction one does not need to present a principled way of determining which types of inference are instantiated in which cases.

5.

In order to talk about the reliability of types of inference we make reference to *content-types*: the types of contents of the doxastic or suppositional attitudes involved in the inference. This is not a mere artifact. The fact that the *singular contents* I am actually reasoning with maintain positive alethic relations guarantees little or nothing concerning the reliability of the process-type I am instantiating. To illustrate, suppose I infer that *The World War II ended in 1945* on the basis of my belief that *World War II began in 1939 and ended in 1945*; suppose, however, that the type of inference I am instantiating in this case is such that, *if any other* claim carrying the information that WWII lasted for more than 5 years were to constitute the content of my input-belief, I would still infer that *World War II ended in 1945*. Never mind the fact that my cognition works in strange ways in this hypothetical scenario: the point is just that, even though the claim that *World War II began in 1939 and ended in 1945* entails that *World War II ended in 1945*, the type of inference my inferential process is a token of in this case is not a reliable process (the type of inference I am instantiating here is that of believing that *World War II ended in 1945* on the basis of any belief to the effect that WWII lasted for more than 5 years). So we cannot make sense of the unreliability of my inferential process by making reference to the content-*tokens* that are actually involved in it––for luckily in this case those tokens do maintain a good alethic connection (logical entailment), even though in many other cases they obviously would not.

Once we realize that mentioning content-*types* is crucial for determining the reliability of types of inference, and that logical

entailment between only some instances of those content-types does not by itself amount to reliability, a natural thought suggests itself: perhaps what makes a type of deduction reliable is the fact that the content-types of its input- and output-attitudes are such that *all instances* of the former maintain logical entailment relations with the corresponding instances of the latter. On the face of the constant association we make between the notion of *deduction* and the notion of *logical entailment* or *validity*, it might appear that this proposal is on the right track. The epistemic virtues of deduction would be intimately connected with highly general truths about what logically follows from what.

As it stands, however, the proposal is still vague. To make it more precise, let me represent a type of inference that goes from a set of beliefs in contents of types $\phi_1,...,\phi_n$ respectively to a belief in a content of type ψ as follows: $\{B\phi_1,...,B\phi_n\} \Rightarrow B\psi$. The symbol '$\Rightarrow$' here represents a reasoned state-transition between doxastic states. (Please note that '$\Rightarrow$' *is not* a conditional, and that '$\{B\phi_1,...,B\phi_n\} \Rightarrow B\psi$' *is not* a sentence. If anything, '$\{B\phi_1,...,B\phi_n\} \Rightarrow B\psi$' is a *definite description* of a type of inference).

Next, we need a mechanism to quantify over tokens of content-types, in such a way as to make the tokens of ψ (the content-type of the output-belief) uniform with the tokens of $\phi_1,...,\phi_n$ (the content-types of the input-beliefs).[4] So let me fetch a *type-token mapping*: a function t that maps pairs of ordinals ($k \geq 0$) and content-types into tokens of those content-types. Every application of that function has the form: $t(ordinal\ k,$ *content-type ϕ) = token of ϕ* or, to make it simpler: $t_k(content\text{-}type\ \phi) =$ *token of ϕ*. E.g., $t_1(\alpha\ is\ \Phi) =$ *Ruth is a logician* and $t_2(\alpha\ is\ \Phi) =$ *Thomas is German* are two tokens of the content-type $\alpha\ is\ \Phi$ and, accordingly, $t_1(There\ are\ \Phi s) =$ *There are logicians* and $t_2(There\ are\ \Phi s) =$ *There are Germans* are two tokens of the content-type *There are Φs* (this is just one example). Depending on what exactly the type $\alpha\ is\ \Phi$ is, however, the claim that *John is bald* may not be in the range of $t_k(\alpha\ is\ \Phi)$, i.e. there is no ordinal k such that $t_k(\alpha\ is\ \Phi) =$ *John is bald*––say, because only claims involving non-vague concepts are tokens of that type. (Relatedly, one might

---

[4] If we had assumed that content-types are just formula-schemes, their tokens being nothing but formulas or sentences of a language *L*, we could just use the typical translation mappings with uniform substitution that we find in logic textbooks. But then we would also be assuming that the contents of the doxastic attitudes involved in particular inferences are sentences or formulas, which is quite contentious––so we need something more general.

systematically infer/be disposed to infer claims of type *α is Φ or α is not Φ* where *Φ* stands for non-vague properties, and yet fail to infer claims of type *α is Ψ or α is not Ψ* where *Ψ* does stand for vague properties as well, e.g. baldness). Many details concerning these type-token mappings need to be worked out in order to make it more precise––but the rough idea will do for our present purposes.

When it comes to *conditional* types of inference, then, the proposal I informally described above can be formulated as follows (where $\phi_1,...,\phi_n, \psi$ are the content-types):

(L) If for all $k \geq 0$: $t_k(\phi_1),...,t_k(\phi_n)$ together logically entail $t_k(\psi)$, then $\{B\phi_1,...,B\phi_n\} \Rightarrow B\psi$ is a reliable type of inference.

Here the property of a type of deduction that accounts for its reliability is: all tokens of its content-types, when arranged in an argument form, are such that the premises logically entail the conclusion.

Notice that this by itself does not mean that content-types are the *logical forms* of their tokens (in which case some of their instances could be quite complex and computationally intractable for ordinary reasoners). What matters is that all instances of those content-types satisfy the logical criterion mentioned in (L). E.g. there might be a type of inference $\{B\phi, B(If \phi then \psi)\} \Rightarrow B\psi$ such that only contents up to a certain complexity are tokens of ϕ and ψ, and nothing more complex than that (the reasoner would not instantiate that type of inference when handling more complex contents). This restriction of course does not apply to logical forms: they have infinitely many instances. Still, assuming that for all k, $t_k(\phi)$, $t_k(If \phi then \psi)$ together logically entail $t_k(\psi)$ then, according to (L), that type of inference is reliable. (One might complain that this type of inference has a narrower scope of application, or that it is less content-neutral than one would perhaps otherwise expect––but that gives us no reason to think that it it is *unreliable*).

6.

What is stated in (L) certainly seems to be true, at least assuming that logical entailment is (perhaps among other things) necessary truth-

preservation:[5] any conditional type of inference that satisfies the antecedent of (L) will indeed be a reliable one. But it is also clear that, depending on what the correct account of logical entailment is, there might be many reliable types of deduction whose reliability is unaccounted for by (L).

Consider the popular notion of logical entailment as truth-preservation in virtue of logical form.[6] If the logic that properly characterizes the class of logically valid *argument-forms* is not inclusive enough, (L) will not enable us to explain the reliability of several types of reliable deduction.

E.g. suppose that there are indeed counterexamples to *modus tollens* (*If A then B, not-B*, therefore *not-A*), say, because it fails to preserve truth when probabilistic claims of a certain kind are plugged in the consequent position of the conditional *If A then B* (see Yalcin 2012 for an example of this kind). Yet here I am, reasoning as follows: *If it is 8 p.m. then it is already dark outside; but it is not dark outside; therefore it is not 8 p.m. yet.* There is no good reason to think that *just because* there are a few probabilistic counterexamples to the *modus tollens* argument-form, the type of inference I am instantiating in this case is not a reliable process-type. Maybe I *would not* instantiate that type of inference if my input-beliefs were about the relevant probabilistic claims instead (or: the content-types that best describe the way in which I am reasoning *do not* include the relevant probabilistic claims among their possible instances).

The bad apples are out of my basket: *all instances* of the content-types that are constitutive of the type of inference {B(*If ϕ then ψ*), B(*not-ψ*)} $\Rightarrow$ B(*not-ϕ*) that I am instantiating in this case are such that a token of *not-ϕ* is true whenever the corresponding tokens of *If ϕ then ψ* and *not-ψ* are true. (Here again the types *not-ψ*, *not-ϕ* etc. *are not* the logical forms of their instances). And yet, as far as (L) goes, the reliability of that type of deductive inference is unaccounted for (assuming that the *modus tollens* form is not a valid argument-form).

Of course, to the extent that it is controversial to say that there are counterexamples to *modus tollens*, this example is also controversial.

---

[5] This is compatible with taking logical entailment to be necessary truth-preservation *plus something else*, e.g. a relevance relation (see Mares 2004).

[6] See Shapiro (2005) for different notions of logical consequence.

But we need not make use of anything as controversial as that to make the point. Suppose I infer that $f(2) > f(8)$ (or that $f(2)$ *is bigger than* $f(8)$), where f is some well-defined function)[7] on the basis of my beliefs that $f(2) > f(4)$ and $f(4) > f(8)$ respectively. Yet, it is not the case that $f(2) > f(4)$ and $f(4) > f(8)$ together logically entail $f(2) > f(8)$––or at least not purely in virtue of logical form. There is a relation R such that both $\langle f(2), f(4) \rangle$ and $\langle f(4), f(8) \rangle$ are part of its extension but $\langle f(2), f(8) \rangle$ is not. But the type of inference I am instantiating here is (or at least it can be) reliable: the tokens of the content-types involved in it always involve the same relation $>$. I do not reason in that way with any old relational concept/predicate.

One might suggest, however, that my inference in this case is also based on a belief that *For all x, y and z, if x > y and y > z then x > z*. Adding this premise to the previous pool of premises would indeed give rise to an argument with a logically valid form. But surely one can perform the target inference without so much as believing any such additional premise. The inference is also guided by the semantic properties of $>$, which can be implemented in one's thought in a way other than in the form of a belief (by training, corrected use or exposition). It would furthermore appear that ascribing beliefs in universally quantified/conditional claims of the relevant type whenever the reasoner has a disposition to perform a certain type of inference somehow trivializes the notion of belief. Be that as it may, the point still stands: the type of inference I instantiate in this case *can be* reliable even if I do not hold a belief toward the relevant universally quantified claim––in which case the tokens of the content-types involved in it will not all maintain the logical entailment relation as a matter of logical form.

So there must be an alternative answer to question (a) that does a better job than (L) does. That answer would account for the reliability of conditional types of deduction without requiring all their content-tokens to maintain logical entailment relations, and it would account for the reliability of a larger class of deductive inferences.

7.

The notion of entailment as truth-preservation in virtue of logical form is not the only one out there. Could not there be a notion of

---

[7] E.g. f maps positive integers onto negative ones as follows: $f(1) = -1, f(2) = -2, f(3) = -3$, etc.

entailment that is unaffected by the possibility of counterexamples to otherwise valid argument-forms? The modal notion of entailment would seem to fit the bill: $A_1,...,A_n$ together entail B when it is impossible for all of $A_1,...,A_n$ to be true and B false (see Shapiro 2005, p. 655). Even if there are counterexamples to the *modus tollens* argument-form, it might still be impossible for both *If it is 8 p.m. then it is dark outside* and *It is not dark outside* to be true and *It is not 8 p.m.* false. Likewise, it would be impossible for both $f(2) > f(4)$ and $f(4) > f(8)$ to be true and $f(2) > f(8)$ false, it does not matter which extra premises are added to the argument.

As it stands, however, the modal notion of entailment is more like a *scheme* for a family of interpretations of the verb 'to entail': different notions of entailment can be defined by means of different notions of impossibility. The schematic use of that verb can then be used to flesh out an equally schematic reliability-entailing condition:

(E) If for all $k \geq 0$: $t_k(\phi_1),...,t_k(\phi_n)$ together entail $t_k(\psi)$, then $\{B\phi_1,...,B\phi_n\} \Rightarrow B\psi$ is a reliable type of deductive inference.

That certainly looks like an improvement––at least with respect to the goal of accounting for the reliability of a wider class of deductive inferences (as compared to the formal notion of logical entailment). Can *all* reliable types of conditional deduction have their reliability accounted for by means of (E) and the modal notion of entailment? Possibly, but a full defense of this claim should *either* establish that a single notion of impossibility/entailment can be used to explain the reliability of all reliable types of conditional deduction *or* it should present us with a principled way of determining which notions of impossibility/entailment are relevant to which reliable types of conditional deduction.

Using the notion of *possible worlds* to explicate the relevant modalities, the point can be made as follows: it is impossible for A to be true and B false iff B is true in every possible world in which A is true; but the universal quantifier here ('every possible world') can quantify over all *nomically* possible worlds––i.e. worlds in which the laws of nature are in operation––or all *metaphysically* possible worlds. So, if one is to claim that (E) can be used to explain the reliability of all reliable types of conditional deduction, one must either settle for just one of these domains or rather

establish which domains of quantification should be used to account for the reliability of which types of deduction.

It is not my goal here to defend that the reliability of all reliable types of conditional deduction can be explained by means of (E) and the modal notion of entailment. It is clear, however, that (E) already does a good job when the entailment relation is defined by means of the notion of *metaphysical* necessity/impossibility: members of a much more inclusive class of reliable types of deduction are entailed to be reliable under that interpretation, including the ones mentioned above (even if their content-tokens fail to constitute arguments with a logically valid argument-form). Metaphysical necessity is necessity that obtains in virtue of the nature or identity of things (see Fine 2002), and so it includes all logically necessary connections––e.g. *Some feminists are philosophers* therefore *Some philosophers are feminists*; *2 is prime and 3 is prime* therefore *2 is prime*, etc.––and in general all 'conceptually' necessary connections––e.g. *c is square* therefore *c is not round*.

Using (E) together with the metaphysical notion of entailment has yet another attractive feature. Consider again the psychological/cognitive properties that can be used to characterize deductive inferences (as opposed to non-deductive inferences): when the subject inferentially believes that ψ on the basis of her belief that ϕ through a process of deduction, she is disposed to revise ϕ upon learning that ψ is not the case, and/or she is disposed to believe that it is impossible for ϕ to be true and ψ false, and/or she takes ϕ to give infallible support to ψ. Arguably, these are indeed things that one should do if ϕ metaphysically entails ψ (and one has formed/maintained one's belief that ψ on the basis of one's belief that ϕ).[8] So that which makes a type of deduction reliable according to the metaphysical reading of (E) (the fact that all tokens of its content-types maintain a relation of metaphysical entailment) also normatively requires the reasoner to have exactly those dispositions/states that characterize her inference as a process of deduction.

---

[8] Of course, these are also things that one should do when ϕ logically entail ψ. But the point is that every logical entailment is also a metaphysical entailment, and therefore the latter allows us to make sense of the correctness of those psychological features in a wider range of cases.

8.

Now let me make some analogous points concerning the reliability of *unconditional* types of deduction. Let $[A\phi \Rightarrow A\chi] \Rightarrow B\psi$ represent a type of inference that outputs a belief in a content of type ψ on the basis of a piece of suppositional reasoning that goes from the assumption that a content of type ϕ is the case to the conclusion (derived under the scope of that assumption) that a content of type χ is the case. The input to this type of process is itself a reasoned state-transition, instead of a set of beliefs. The intensional operator A stands for suppositional attitudes, both stand-alone assumptions and assumptions that are held under the scope of previously made assumptions. There are even more complex types of unconditional inference, not only involving more than one piece of suppositional reasoning, but also pieces of suppositional reasoning that rely on more than just one initial assumption. But the simpler form will do for our present purposes.

A proposal along the lines of (E) can now be fleshed out that accounts for the reliability of unconditional types of deduction as well:

(E*) If for all $k \geq 0$: $t_k(\phi)$ entails $t_k(\chi)$ and, consequently, $t_k(\psi)$ is necessarily true, then $[A\phi \Rightarrow A\chi] \Rightarrow B\psi$ is a reliable type of deductive inference.

'Necessarily' here is intended to express the same type of modality as the one that is used to interpret 'entails' (which is, as before, a modal notion of entailment). Similarly, the phrase 'consequently' relating the fact that $t_k(\phi)$ *entails* $t_k(\chi)$ with the fact that $t_k(\psi)$ *is necessarily true* can be understood as expressing the same type of entailment relation as the one that holds between $t_k(\phi)$ and $t_k(\chi)$. Or at least that can be so under the metaphysical interpretation of entailment: the fact that $t_k(\phi)$ *metaphysically entails* $t_k(\chi)$ itself metaphysically entails that $t_k(\psi)$ *is necessarily true.*

Again, it is not my goal to defend that the reliability of *all* reliable types of unconditional deduction can be explained by means of (E*). Clearly, however, (E*) together with the metaphysical interpretation of entailment and necessity also allow us to account for the reliability of a large class of reliable types of unconditional deduction. In particular, it tells us that beliefs toward metaphysically necessary truths can be reliably

formed through unconditional ways of inferring, which includes both beliefs toward logical truths and also 'conceptual' or 'analytic' truths that are not themselves logical truths. As an example of the former, consider the unconditional type of inference:

[A(*Some Fs are Gs*) ⇒ A(*Some Gs are Fs*)] ⇒ B(*If some Fs are Gs then some Gs are Fs*)

Given some standard assumptions about the semantic properties of *Some* and *If... then...*, this type of inference clearly satisfies the antecedent of (E*): every instance of *Some Fs are Gs* metaphysically entails the corresponding instance of *Some Gs are Fs*, which in turn metaphysically requires that the corresponding instances of *If some Fs are Gs then some Gs are Fs* are necessarily true.

As an example of an unconditional type of inference that outputs a belief toward a 'conceptual' truth that is not itself a logical truth, consider:

[A(*a is a square*) ⇒ A(*a is not round*)] ⇒ B(*If a is a square then a is not round*)

Here again, assuming that *square* and *round* have the meaning they have in English when used to talk about geometrical figures, every instance of *a is square* metaphysically entails the corresponding instance of *a is not round* and, consequently, instances of *If a is a square then a is not round* are metaphysically necessary.

In addition to logical truths and non-logical 'conceptual' truths, (E*) also regards certain modes of reasoning that output beliefs toward necessity claims as reliable, e.g. as in the unconditional type of inference: [A(*not-φ*) ⇒ A*φ*] ⇒ B(*Necessarily φ*). (*What kinds* of necessity claims are judged to be true through a process like this depends on the correct interpretation of the necessity operator in *Necessarily φ*).

9.

Let $(E)_M$ /$(E^*)_M$ stand for (E)/(E*) under the metaphysical interpretations of entailment and necessity. Perhaps some philosophers will not be completely satisfied with the claim that $(E)_M$ and $(E^*)_M$ allow

us to explain the reliability of many reliable types of deduction, including all paradigmatic types of deduction. They might agree that the property of having content-tokens that satisfy the conditions mentioned in $(E)_M$ or $(E^*)_M$ does indeed explain the reliability of certain types of deduction—if only because *not all* reliable types of deduction can have their reliability accounted for in one of these ways.

Consider the case of conditional deduction. Perhaps there are reliable types of conditional deduction that fail to satisfy the antecedent of $(E)_M$. Some conditional types of deduction would not cease to be reliable just because *only some* of the instances of its content-types fail to maintain entailment relations. One way to motivate this would be by using Goldman's (1979) notion of *conditional* reliability, which is used by him to define the reliability of belief-dependent cognitive processes in general (including deductive and non-deductive inference): "A process is conditionally reliable when a sufficient proportion of its output-beliefs are true *given that its input-beliefs are true*" (1979, p. 13). One might similarly suggest that if a sufficient proportion of the tokens of ψ are entailed by the corresponding tokens of $\phi_1,...,\phi_n$, then the deductive type of inference $\{B\phi_1,...,B\phi_n\} \Rightarrow B\psi$ is reliable.

It is not immediately clear what a 'sufficient proportion' is supposed to be here. Is anything more than 50% of the total class of tokens maintaining entailment relations enough for the reliability of a type of deductive inference? Or does that somehow depend on the context? Etc. Putting that aside, however, insofar as different proportions of entailment relations between the tokens of ψ and the corresponding tokens of $\phi_1,...,\phi_n$ are taken to be enough for the reliability of a type of deduction, there would be types of deduction with different degrees of reliability. And so there would be reliable types of deduction that are less than maximally reliable––which some may think it is absurd. (As one might say: 'If a type of deduction is reliable at all, it is maximally so').

I do not have space here to discuss the idea that deduction admits of different degrees of reliability. Even assuming that it does, however, the antecedents of $(E)_M$ and $(E^*)_M$ will occupy a special place in the whole list of explanatory reliability-entailing conditions with respect to deductive process-types. The class of types of deduction that satisfy those conditions is such that all its members are maximally reliable *and* it includes all paradigmatic examples of reliable deduction. The paradigmatic examples

of conditional deduction are those in which we draw logical and more generally 'conceptual' consequences from our beliefs, and paradigmatic examples of unconditional deduction are those in which we gain knowledge of logical, mathematical and other 'conceptual' truths. That is a large and important chunk of the realm of claims that are traditionally taken to be knowable *a priori*, or 'from the armchair'.

(Of course, there are types of inference that do satisfy the antecedents of $(E)_M$ and $(E^*)_M$ and yet they *are not* paradigmatic examples of deduction, e.g. consider the conditional type of inference $B(x \text{ is water})$ $\Rightarrow$ $B(x \text{ is } H_2O)$. But if these are not really deductive types of processes, then they are also not 'on the market' for assessments as to what makes them reliable *as types of deduction*. I.e. they are not types of cognitive process that exhibit those psychological features that characterize deduction: disposition to revise one's attitudes toward the premises on the basis of revising one's attitude toward the conclusion, etc.)

Perhaps one would want to say that $(E)_M$ and $(E^*)_M$ individuate only two types of properties that are capable of grounding the reliability of deductive inferences––but there are many other properties of this kind. An investigation about all classes of such properties, paired with the kinds of deductive inferences whose reliability they account for, is something that I must leave for future work, however.

To the extent that $(E)_M$ and $(E^*)_M$ allow us to explain what makes certain types of deduction reliable (i.e. because all their tokens satisfy the relevant alethic relations), those reliability-entailing conditions also help us understand how knowledge-expansion and knowledge-generation through deduction are possible. In particular, if in some contexts (or for some types of contents) knowledge through deduction requires both necessary truth-preservation among the singular contents one reasons with *and* maximal reliability of the belief-forming process, then types of deduction that satisfy $(E)_M$ and $(E^*)_M$ are able to expand and to generate knowledge respectively in those contexts. So some progress is made in addressing the question of how deduction is able to play those epistemic roles.

References

Balcerak Jackson, Magdalena and Balcerak Jackson, Brendan (2013). 'Reasoning as a source of justification', *Philosophical Studies* 164(1): 113-126.

Baumann, Peter (2009). 'Reliabilism: Modal, Probabilistic or Contextualist', *Grazer Philosophische Studien* 79(1): 77-89.

Boghossian, Paul (2014). 'What is inference', *Philosophical Studies* 169(1): 1-18.

Brown, Jessica (2004). 'Wright on transmission failure', *Analysis* 64(281): 57-67.

Conee, Earl and Feldman, Richard (1998). 'The Generality Problem for Reliabilism', *Philosophical Studies* 89(1): 1-29.

Dasgupta, Shamik (2014). 'On the Plurality of Grounds', *Philosophers' Imprint*, 14: 1–28

Dogramaci, Sinan (2016). 'Reasoning Without Blinders: A Reply to Valaris', *Mind* 125(499): 889-893.

Duc, Ho Ngoc (1997). 'Reasoning about Rational, but not Logically Omniscient, Agents', *Journal of Logic and Computation* 7(5): 633-648.

Fine, Kit (2002). 'The Varieties of Necessity', In T. S. Gendler & J. Hawthorne (eds.) *Conceivability and Possibility*, Oxford: Oxford University Press.

Goldman, Alvin (1979). 'What Is Justified Belief?', in G. S. Pappas (ed.), *Justification and Knowledge*, Dordrecht: Reidel, pp. 1-25.

Goldman, Alvin (1988). 'Strong and Weak Justification', *Philosophical Perspectives* 2: 51-69.

Harman, Gilbert (1986). *Change in View*, Cambridge: MIT Press.

Jago, Mark (2014). *The Impossible: An Essay on Hyperintensionality*, Oxford: Oxford University Press.

Lasonen-Aarnio, Maria (2008). 'Single premise deduction and risk', *Philosophical Studies* 141(2): 157-173.

Mares, Edwin (2004). *Relevant Logic: A Philosophical Interpretation*, Cambridge: Cambridge University Press.

Meyer, John-Jules (2003). 'Modal Epistemic and Doxastic Logic', In D. Gabbay & F. Guenthner (eds.) *Handbook of Philosophical Logic* (2 ed.) Vol. 10, Dordrecht: Kluwer, pp. 1-38.

Richard, Mark (1997). 'Propositional attitudes' in B. Hale and C. Wright (eds.), *A Companion to Philosophy of Language*, Oxford: Blackwell, pp. 197-226.

Rumfitt, Ian (2008). 'Knowledge by Deduction', *Grazer Philosophische Studien* 77: 61-84.

Schechter, Joshua (2010). 'The Reliability Challenge and the Epistemology of Logic', *Philosophical Perspectives* 24(1): 437-464.

Schechter, Joshua (2013). 'Rational self-doubt and the failure of closure', *Philosophical Studies* 163(2): 429-452.

Shapiro, Stewart (2005). 'Logical Consequence, Proof Theory and Model Theory', in S. Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford: Oxford University Press, pp. 651-670.

Valaris, Markos (2016). 'Supposition and Blindness', *Mind* 125(499): 895-901.

Williamson, Timothy (2007). *The Philosophy of Philosophy*, Malden MA: Blackwell Publishing.

Wright, Crispin (2002). 'Anti-sceptics simple and subtle: Moore and McDowell', *Philosophy and Phenomenological Research* 65(2): 330-48.

Wright, Crispin (2014). 'Comment on Paul Boghossian, "What is inference"', *Philosophical Studies* 169(1): 27-37.

Yalcin, Seth (2012). 'A Counterexample to Modus Tollens', *Journal of Philosophical Logic* 41(6): 1001-1024.

Endereço postal:

Cologne Center for Contemporary Epistemology and the Kantian Tradition (CONCEPT),
Albertus-Magnus-Platz (Bauteil 4, R. 4207),
University of Cologne,
Cologne, 50923.