

Soundness arguments for consistency and their epistemic value: A critical note

MATTEO ZICCHETTI 

Faculty of Philosophy, University of Warsaw, Krakowskie Przedmieście 3, 00-047 Warsaw, Poland

Soundness Arguments for the consistency of a (mathematical) theory S aim to show that S is consistent by first showing or employing the fact that S is sound, i.e., that all theorems of S are true. Although soundness arguments are virtually unanimously accepted as valid and sound for most of our accepted theories, philosophers disagree about their epistemic value, i.e., about whether such arguments can be employed to improve our epistemic situation concerning questions of consistency. This article provides a (partial) negative answer to this question and argues that soundness arguments cannot be employed to justify their conclusion. Additionally, soundness arguments are unconvincing; they cannot be employed to overcome reasonable open-mindedness about their conclusion.

Keywords: Soundness Arguments; Consistency; Transmission of Justification; Underminers; Failure; Cogency Template.

I. Outline

Soundness arguments for consistency have been discussed in the philosophy of mathematics since the 1960s.¹ A soundness argument for the consistency of S is an argument for the consistency of S, which employs a *detour by arguing first* that S is sound, i.e., that all theorems of S are true. Informally, we can think of a soundness argument for the consistency of S as the following:

Corresponding author: Matteo Zicchetti (matteo.zicchetti@gmail.com)

¹See, for instance, (Myhill 1960), (Dummett 1978), (Girard 1987), (Wright 1994), and (Shapiro 1998).

If the axioms of S are true and the rules of inference of S are truth-preserving, then all theorems of S are true. Since all theorems of S are true, S does not prove any false statement. Therefore, S is consistent.²

Although it is virtually unanimously accepted that such arguments are valid and sound for most of our accepted theories, philosophers disagree about whether soundness arguments have *any epistemic value*. Girard, Dummett and Wright separately suggest (Girard 1987; 226), (Dummett 1978; 194), and (Wright 1994; 176) that soundness arguments have little epistemic value. On the other hand, Shapiro claims (Shapiro 1998; 505) that soundness arguments can be employed to improve our epistemic situation concerning the statement expressing the consistency of arithmetic.³ Of course, there are many ways of understanding the rather vague phrase ‘improving one’s epistemic situation’. An essential way of improving one’s epistemic situation about the consistency of S by means of a soundness argument is to use it to *justify* the consistency of S . This possibility is entertained by Horsten in his recent article when focusing on first-order arithmetic:

[Y]ou might go on to argue by mathematical induction, using a Tarskian compositional notion of truth, that all theorems of $[S]$ are true, and that therefore, since $\perp$ is not true, $[S]$ must be consistent. In this way, you might obtain justified belief in $\neg Bew_S(\perp)$.⁴ (Horsten 2021; 751)

Where $\neg Bew_S(\perp)$ is a standardly accepted statement expressing the consistency of S . An additional and equally important way of improving one’s epistemic situation is to use a soundness argument to overcome reasonable open-mindedness about the consistency of S . This article focuses on these two ways of improving one’s epistemic situation, and more explicitly, it investigates the following two questions:

- (†) Can soundness arguments, in principle, be employed to justify the consistency of S ?
- (‡) Can soundness arguments, in principle, be employed to overcome reasonable open-mindedness about the consistency of S ?

This article has the following structure: Section II introduces and briefly discusses the needed concepts from epistemology: The notions of transmission

²A similar version of this argument is presented by (Dummett 1978).

³Dummett and Wright argue that soundness arguments *prima facie* fail to be informative, whereas Girard limits himself to the claim that soundness arguments have little epistemic value without further qualification. On the other hand, Shapiro suggests that soundness arguments might have some explanatory force.

⁴In fairness, Horsten is quite speculative and is not interested in investigating the soundness argument. He proposes an alternative, so-called *entitlement* strategy to justify the consistency of arithmetic. Therefore, what I will say in this article is not against the approach in (Horsten 2021). For more information on the entitlement strategy, see (Wright 2004), (Pedersen 2021) and (Zicchetti 2022).

of justification, transmission failure, and the concept of an undercutting defeater. Section III investigates (†) and (‡) and provides negative answers to both: I argue that soundness arguments *fail to transmit justification*—in a standard sense to be introduced later—from their premises to their conclusion. As I will show, this amounts to a vicious *epistemic dependence* between the argument's premises and its conclusion. As I will argue, the existence of a justification for the argument's premises presupposes the truth of the argument's conclusion. As for (‡), I will show that soundness arguments are not *cogent*—for a standard notion of cogency to be given later in this article. More specifically, soundness arguments violate Crispin Wright's 'new' *Cogency Template* introduced in (Wright 2002). Finally, Section IV makes some more exploratory comments on the ramifications, philosophical implications, and generality of the results of Section III.

II. Transmission, failure, and underminers

A well-known distinction in logic and its philosophy is between *valid* and *invalid* arguments: An argument is valid just in case whenever its premises are true, its conclusion is also true. Additionally, valid arguments are *sound* just in case they have true premises (they are *unsound* otherwise). It is a nowadays accepted distinction that some valid arguments can be employed to justify their conclusion and that other valid arguments cannot. The former arguments are called *transmissive*:

Definition 1. (*transmission of justification*) *An argument D is transmissive just in case the following obtains:*

- (a) *D is valid,*
- (b) *D 's premises are justified, and*
- (c) *The justification for D 's conclusion is given by virtue of (a) and (b).*⁵

Transmission is a principle about the possibility, for a given valid argument D , of justifying its conclusion by virtue of its validity *and* the existence of a justification for its premises only. Following (Moretti and Piazza 2018), transmission is understood as being about *propositional* justification and not about *doxastic* justification. Importantly, propositional justification is a property of *propositions*, whereas doxastic justification is a property of *beliefs*.⁶ For our investigation, we focus on the following non-trivial case of failure: The case where (a) and

⁵This understanding of transmission is standard. See, for instance, (Wright 2002), (Dretske 2005), or (Moretti and Piazza 2018) for a recent presentation of transmission. Sometimes this notion is understood in terms of *raising our confidence* in its conclusion. For a discussion of transmission along these lines, see (Okasha 2004), (Chandler 2009) and (Moretti 2010).

⁶Both (Wright 2002) and (Dretske 2005) seem to focus on propositional justification, whereas (Tucker 2010), for instance, focuses explicitly on the transmission of doxastic justification.

(b) obtain, *whilst* (c) fails to obtain.⁷ We should say a few clarificatory words about the distinction between transmission and closure. It is well-known that transmission is a distinct epistemic phenomenon from *closure of justification*: An argument D satisfies closure of justification just in case, whenever D 's premises are justified, D 's conclusion is also justified. It is important to see that an argument can satisfy closure and fail to transmit justification.⁸ For illustration, let us look at a trivial model of transmission failure: the case of circular arguments. Valid circular arguments fail to be transmissive, while nevertheless satisfying closure, where a circular argument is an argument whose conclusion is exactly one of its premise(s). The following is a trivial example of such arguments:

- (i) p ,
- (ii) p implies p ,

$\therefore p$.

Where (i) and (ii) are the premises and the conclusion is obtained by performing *modus ponens* from (i) and (ii). Assuming that (i) and (ii) are justified, closure is satisfied; trivially, there is a justification for the argument's conclusion precisely because the argument's conclusion appears as one of the argument's premises.⁹ However, as (Moretti and Piazza 2018) explain, the argument fails to be transmissive because the justification for the conclusion is not provided by virtue of the premises being justified *and* of the argument's validity. If (i) is justified, the conclusion is also justified without the need for the argument's validity because (i) is the same as the conclusion.

There are additional, philosophically more fruitful ways other than plain circularity to diagnose transmission failure. Consider, for instance, the notion of an *underminer*. Consider some proposition p , some evidence e , and suppose that e (defeasibly) supports p . An underminer u of e is whatever statement whose truth (at least partially) defeats e 's support of p by making e less reliable (or even completely unreliable).¹⁰ Using the notion of an underminer, one

⁷As (Moretti and Piazza 2018) point out, cases where either (a) or (b) fail to obtain are trivial cases of failure.

⁸That closure and transmission are different epistemic phenomena has been recognized and made explicit, for instance, in (Wright 2002), (Dretske 2005) or (Tucker 2010). Recently (Wright 2022) called closure principles 'static' and referred to transmission principles as 'dynamic'. However, it should be pointed out that philosophers sometimes talk about closure, although they are discussing what I call 'transmission' here. Examples are (Williamson 2000) and (Hawthorne 2004; Hale and Wright 2005) and even (Dretske 2005). This has been also pointed out in (Tucker 2010).

⁹It is important to point out that, in the presented circular argument, the conclusion and one of the premises are identical. This means that either one of the premises also occurs as the argument's conclusion. Of course, the argument is not circular if its conclusion is logically equivalent to either one of the premises.

¹⁰This is in contrast to so-called overriding (or rebutting) defeaters. An overriding defeater for p would be any new evidence e' for the negation of p or for some proposition q inconsistent

can aim to evaluate some arguments as nontransmissive. For illustrative purposes, consider, as an example, one of the most (in)famous putatively non-transmissive arguments in epistemology: Moore's argument for the existence of the external world:¹¹

[MOORE]

M1. I have a hand.

M2. If I have a hand, then there is an external world

∴ There is an external world.

For the context of the present investigation, there is no need to venture into a discussion of Moore's 'proof'. However, it is helpful to look at the structure of the argument briefly. According to the usual reconstruction of [MOORE], the first premise (M1) is justified by *sense experience*. The second premise (M2) is justified *a priori* as one of the conceptual tenets of *common-sense realism*: All objects are material and external. Finally, Moore concludes that there is an external world by performing *modus ponens*. Suppose that (M1) is justified by experience, as in the relevant case discussed in the literature. One way of diagnosing a failure of transmission would be to argue that the evidence *e* provided by the experience of hands supports (M1), i.e., the statement that I have hands, only if, for some relevant underminers $u_1, \dots, u_n$, which would make *e* unreliable, none of $u_1, \dots, u_n$ obtain.

Intuitively, one of the underminers would be that there is no external world, for if that were to be the case, experience would hardly support any propositions about external objects. But then, the experience of hands supports that there are hands only if—amongst other things—there is an external world! In general, philosophers concluding that [MOORE] fails to transmit generally argue for some vicious *epistemic circularity*, grounded in a dependence relation between the existence of evidential support for (M1) by means of experience and the status of [MOORE]'s conclusion.¹²

Let me stress this point: We do not care about [MOORE]'s diagnoses or about whether the argument actually transmits, as our brief discussion only has an illustrative purpose. Our question is whether soundness arguments fail to transmit justification similarly to one of [MOORE]'s diagnoses. Using the

with *p*. For this reason, overriding defeaters are sometimes called 'first-order' (Carter 2016). The *locus classicus* for the notion of a defeater is (Pollock 1970). In Pollock's work, underminers were called undercutting defeaters. For a recent general discussion of defeaters in epistemology, see, for instance, (Moretti and Piazza 2017).

¹¹This argument has been extensively investigated in the literature on scepticism and transmission of justification. For a glimpse into this discussion, see, for instance, (Wright 2002 2003, 2004, 2014) and (Pryor 2000, 2004). For the original argument, see (Moore 1939).

¹²Other so-called templates to model this dependence are for instance Wright's *information dependence template* (Wright 2002, 2003, 2012), Wright's *disjunctive template* (Wright 2000), and Davies' *directness template* in (Davies 2009). For a presentation and discussion of these templates, see (Moretti and Piazza 2018).

presented notion of an underminer, the following section argues that soundness arguments fail to transmit justification.

III. (†): Transmission failure

As pointed out at the beginning of this article, a soundness argument for consistency aims to show the consistency of S by employing a detour using the fact that S is sound, i.e., that all theorems of S are true. Let us reconstruct the soundness argument for the consistency of S explicitly:

[SOUND]

S1. For any statement p , if p is a theorem of S , then p is true.

S2. For any p , if it is not the case that p is true, then p is not a theorem of S .

S3. For a given contradiction, $\perp$, it is not the case that $\perp$ is true.

$\therefore \perp$ is not a theorem of S . That is, S is consistent.

Employing the notion of an underminer from Section II, I will show that [SOUND] is not transmissive and that, therefore, cannot, in principle, be employed to justify its conclusion, i.e., the statement that S is consistent. To do so, we must specify the source of the justification for the argument's premises. From the literature on theories of truth, it is now standard knowledge that one can *prove* the formalized statement expressing that S is sound in a formal theory of truth T over S . This was part of the background from which Horsten made his remark in the quoted passage in Section I. Let us be more specific about the theory of truth involved in the proof. The theory of truth T over S extends S by expanding the language of S with a fresh, new predicate $\mathcal{T}r$ to play the role of the truth predicate, and with axioms expressing natural properties of truth: that truth commutes with the logical connectives and with the quantifiers. In the case of a mathematical theory S equipped with induction, one finally allows for the truth predicate $\mathcal{T}r$ to appear in instances of mathematical induction. For clarity, let us specify the following notation: For a theory of truth, I will denote with T the collection of the theory's truth-theoretic axioms. On the other hand, I will use the notation ' $T[S]$ ' to denote the extension of S by T 's truth-theoretic axioms, also allowing the truth predicate to appear in instances of S 's schemata (if S has any). In the context of theories of truth over Peano arithmetic PA , the classical theory of truth is known as $CT[PA]$ (for Compositional Truth). This is, in fact, the Tarskian theory of truth mentioned by Horsten in the previously quoted passage.¹³ The axioms of CT are virtually unanimously accepted. In fact, CT is one of the standard axiomatizations of a classical notion of truth over arithmetic. Moreover, in this case,

¹³For a presentation of the axioms of CT see, for instance, (Halbach 2014).

there are no worries about possible paradoxes of truth because the truth predicate of CT is *typed*, i.e., it applies only to formulas of the language without the truth predicate. In CT[S], one can prove the formalized statement expressing the soundness of S, i.e., that all theorems of S are true.¹⁴ For the sake of this investigation, we take *proofs* of the formalized statement expressing that S is sound as our evidence *e* to support the informal statement that S is sound. In the present context, [SOUND]'s premises are taken to be justified by the proofs of the respective formalized statements in our Tarskian theory of truth.

Before continuing, two clarifications about the theory of truth are in order: First, although CT is accepted as one of the standard truth theories, the question of the acceptability of CT is more delicate within the debate about deflationism about truth. In that context, the issue of the acceptability of CT is closely related to whether truth deflationists should accept proof-theoretically *non-conservative* theories of truth.¹⁵ However, that issue is not relevant to our present discussion since CT, or an equivalent theory of truth able to prove the soundness of S, is taken to be instrumentally accepted as a means to justify (or otherwise support) the premises of the soundness argument.¹⁶ Finally, one might wonder whether one should interpret the formalized statements provable in our theory of truth as expressing the informal statements one aims to justify. Although this is an interesting question, it should be pointed out that in our present context, it is assumed that the notion of provability in the theory of truth T has its natural properties so that one can naturally interpret the formalized statements as expressing their target informal counterparts.¹⁷ Against this background, we can now focus on the question of transmission: Does [SOUND] transmit justification to its conclusion?

For the sake of argument, suppose that proofs in the theory of truth CT[S] are accepted as evidence for mathematical statements and that, in particular, a CT[S]-proof of the formalized soundness statement for S provides some evidence *e* *supporting* the informal statement that S is sound. Presumably, proofs

¹⁴This statement is often called the *global reflection principle* for S. See (Halbach 2014, Theorem 8.39, 104) for a sketch of the proof. For general philosophical discussions of reflection principles, see, for instance, (Franzén 2004) and (Cieśliński 2017).

¹⁵A theory T is proof-theoretically non-conservative over its base theory S just in case S proves some statement *p* in the language of S, which is unprovable in S. Due to Gödel's Incompleteness, any consistent theory T that proves the consistency of S—in the language of S—must be non-conservative over S.

¹⁶Although a discussion of deflationism would be interesting, it exceeds the scope of this investigation. For a glimpse into the discussion about deflationism and conservativity, see (Ketland 2005, 2010), (Tennant 2005, 2010), (Cieśliński 2010), and more recently (Waxman 2017) and (Fujimoto 2017, 2021, 2022).

¹⁷These are the so-called Löb's derivability conditions. See, for instance, (Boolos 1993) and (Halbach 2014). One might still have sceptical worries about the possibility of interpreting formalized statements, as some target informal claims. However, this worry is quite deep and not limited to the present context. To investigate this worry would exceed the scope of this article.

in mathematical theories do not support or justify theorems independently of collateral information. For instance, a proof of a statement p , given in a trivial theory S , provides no support for p . This is so because if S is trivial, then S proves any statement—formulated in its language—and in particular, it also proves the negation of p . In general, for proofs in theories to support their theorems, one has to have some collateral information that at least the theory employed is *reliable*. Given the relation between the reliability of evidence and underminers (from Section II), this motivates the following *No Underminer Condition* (NUC):

(NUC) For any theory S , a proof in S of some statement p defeasibly supports p only if, for some relevant underminers $u_1, \dots, u_n$ of S , none of $u_1, \dots, u_n$ obtain.

In particular, for the reason given above, if S -proofs must be able to defeasibly support S -theorems, the statement that S is inconsistent cannot obtain.¹⁸ That is, S -proofs defeasibly support S -theorems only if, amongst other things, S is consistent.¹⁹

With (NUC) in place, we can show that [SOUND] *fails to transmit* justification to its conclusion. We argue for failure in the following manner:

(Context) The soundness argument is valid and its premises are justified by proofs in the (suitably formulated) theory of truth $CT[S]$.

- 1 Any proof in $CT[S]$ of a statement p supports p only if, for the relevant underminers u of $CT[S]$, u does not obtain. In particular, the inconsistency of $CT[S]$ cannot obtain. That is, $CT[S]$ must be consistent *at least*.
- 2 In particular, 1 holds for the $CT[S]$ -proof that S is sound.
3. Since we assumed that the premises of [SOUND] are justified by $CT[S]$ -proofs, we have by 2 that $CT[S]$ must be consistent.
4. But then, since S is a sub-theory of—in fact, it is syntactically contained in— $CT[S]$, the consistency of $CT[S]$ directly implies the consistency of S . So, we can conclude the following:
5. S must be consistent.
6. But then, [SOUND] fails to transmit justification.

[SOUND] fails to transmit justification because the existence of a justification for its premises *by means of proofs in* $CT[S]$ *requires* the consistency of S , which

¹⁸This is due to the fact that inconsistency implies triviality. This is standardly accepted, with the only exception being the *dialetheist* such as Graham Priest (Priest 2006). However, it should be pointed out that even dialetheists accept that *some* inconsistencies imply triviality—although they do not accept the conditional for *all* inconsistencies. As a side remark: We do not need to worry about dialetheism since the dialetheist is *prima facie* unable to employ the [SOUND] because they wouldn't accept both premises (S_2) and (S_3).

¹⁹This presupposes some weak form of determinacy of facts about the syntax of S , i.e., that if S is not inconsistent, then S is consistent. This is hardly problematic in the present context, where we are considering theories S formulated in first-order classical logic.

is expressed by [SOUND]'s conclusion! That is, [SOUND] cannot be employed to justify the consistency of S because its premises are justified only if—amongst other things—its conclusion is true. With this, we can conclude that soundness arguments cannot improve—in the way specified in (†)—our epistemic situation concerning questions of consistency. Some clarifications need to be made about this conclusion. However, since these clarifications apply also to the investigation of cogency, we postponed them to Section III.B and Section IV, and discuss the issue of cogency first.

III.A (†): Failure of cogency

The second question I consider is whether sound arguments can be employed to overcome reasonable open-mindedness about their conclusion, i.e. whether S is consistent. For this, we need the notion of *open-mindedness*. Following (Wright 2022), we understand open-mindedness about a statement in the following manner:

An agent is reasonably open-minded about p just in case the following is the case:

- (i) The agent has no opinion whether or not p ;
- (ii) Open-mindedness about p involves readiness to accept or reject p , should appropriate evidence turn up;
- (iii) One is of the opinion that so far, one does have no basis for a view, but one does not rule out the possibility that new information might mandate taking a view about p .

Importantly, Wright's notion of open-mindedness about p is weaker than agnosticism or scepticism about p due to (ii) and (iii); the agent's open-mindedness welcomes a readiness to accept new evidence. Moreover, the agent is open to the possibility of improving her epistemic situation about p . Agnosticism and scepticism presumably reject at least either (ii) or (iii). With this notion of open-mindedness in place, Wright motivates the following notion of cogency:

Definition 2. (*Cogency*) An argument D from premises $p_1, \dots, p_n$ to a conclusion q is cogent just in case rationally accepting evidential support for $p_1, \dots, p_n$ is compatible with rational open-mindedness about q .

Using the notion of the previous section, we can show that [SOUND] is not cogent. For the sake of the argument, we assume that the agent is rationally open-minded about whether S is consistent and that she aims to employ [SOUND] to overcome this open-mindedness. We argue in the following manner:

1. The agent's open-mindedness about whether S is consistent amounts (possibly upon reflection) to her open-mindedness about whether, for some relevant underminer u of the consistency of S , u obtains.
2. Since $CT[S]$ contains S , if u undermines the consistency of S , it also undermines the consistency of $CT[S]$.
3. From this, we can infer that the agent should be at least equally open-minded about whether the consistency of $CT[S]$ is undermined.
4. Intuitively, her rational open-mindedness about whether $CT[S]$ is consistent implies at least equal open-mindedness about whether one should rationally accept that $CT[S]$ -proofs support $CT[S]$ -theorems.
5. In particular, one should be equally open-minded about accepting $CT[S]$'s support of the statement that S is sound.

But then, one cannot be rationally open-minded about whether S is consistent whilst rationally accepting that $CT[S]$ -proofs support the premises of $[SOUND]$. That is, $[SOUND]$ violates Wright's Cogency Template. Therefore, we conclude that $[SOUND]$ cannot be employed to overcome reasonable open-mindedness about the consistency of S . Analogously to the case against transmission, some clarifications about my arguments and my conclusions ramifications are necessary. The remainder of this article focuses precisely on clarifying such matters.

III.B Generalizing failure

The previous section has shown that soundness arguments cannot be successfully employed to improve our epistemic situation concerning questions of consistency in two quite essential ways: They cannot be successfully employed to justify the consistency of S , and they cannot be successfully employed to overcome reasonable open-mindedness about the consistency of S . However, one might wonder about the generality of this analysis for the following reason: For both my arguments for transmission failure to go through the fact has been employed that $CT[S]$ *syntactically contains* our target mathematical theory. It was assumed as a part of the context that agents were employing a theory of truth CT containing S . This fact was crucial in step four of the argument for transmission failure when arguing that the consistency of $CT[S]$ *directly* gives us the consistency of S . Similarly, this fact was employed in step two of the argument against cogency when arguing that open-mindedness about the consistency of S implies equal open-mindedness about the consistency of $CT[S]$.

However, recent developments in truth theories showed that soundness arguments can be formulated so that the theory of truth proves the soundness of the target mathematical theory in such a way that avoids the syntactic containment of the target mathematical theory in the truth theory.²⁰ In the

²⁰Thanks to an anonymous referee of this journal for pointing this out to me.

formalization of soundness arguments, the assumption that the truth theory CT contains S can be lifted. Such a setting is known in the literature as the *disentangled setting*.²¹ For illustration, it is helpful to briefly introduce and discuss one example investigated in (Leigh and Nicolai 2013). In contrast to, for instance, the case of CT[PA], where CT extends Peano arithmetic, the authors employ a theory of truth called CTD[O] extending a base theory of syntax formulated in a language with different sorts to talk about different domains of discourse: The language of CTD[O] has a sort *o* to talk about ‘mathematical objects’, a sort *s* to talk about syntax, and a third, mixed sort *sq* ranging over variable assignments. This sort is mixed because the variable assignments associate elements of the syntactic domain of discourse with elements of the mathematical domain of discourse. Finally, the mixed theory CTD[O] has the expressive resources to express syntactic facts about the language of the sort *o*. Moreover, it can express the usual axioms for truth.²²

In (Leigh and Nicolai 2013), the authors show that the disentangled theory of truth proves the formalized *syntactic* statement that expresses the *objectual mathematical* theory O is sound. Therefore, it also proves that O is consistent. Crucially, their disentangled theory *does not syntactically contain* the objectual mathematical theory.²³ At this point, it is important to clarify that, for the present discussion, we are not interested in investigating specific proofs and the technicalities involved in the disentangled setting.²⁴ For the present epistemological discussion, the example of disentangled settings is relevant, as it provides us with the possibility of formalizing soundness arguments lifting the assumption that the truth theory syntactically contains the target mathematical theory, whose consistency we wish to prove. It should be noted that the versions of the arguments given in the previous section do not apply—at least directly—to such disentangled scenarios; the move from the consistency of the truth theory to the consistency of the target mathematical is blocked because there is no syntactic containment. Because of this fact, it is important to ask whether the previous arguments can be generalized to such scenarios. If such generalization were not possible, my proposed analysis would only have minor significance for the question of whether soundness arguments can, *in general*, be employed to improve our epistemic situation concerning questions of

²¹This has been investigated, for instance, in (Heck 2015) and (Leigh and Nicolai 2013).

²²For the precise formulation of CTD[O] and its language, see (Leigh and Nicolai 2013, pp. 619–20).

²³These are Proposition 3.2 (Leigh and Nicolai 2013; 622) and Proposition 3.8 (Leigh and Nicolai 2013; 626).

²⁴This would be relevant and important, for instance, in the context of truth deflationism. As (Leigh and Nicolai 2013) points out, the theory CTD[O]—and some natural variants—are conservative over O. See, for instance, Corollary 3.12. (Leigh and Nicolai 2013; 627). However, the addition of *prima facie* natural principles—so-called “coding axioms”—results in non-conservative theories. See, for instance, Proposition 3.18 (Leigh and Nicolai 2013; 631).

consistency.²⁵ Fortunately—as I will show in what follows—my argument can be adapted to this more general setting. Consider the following scenario:²⁶ An agent accepts Peano arithmetic PA and aims to show that PA is consistent. They wish to do so by employing a soundness argument for the consistency of PA:

[SOUND*]

S1* For any statement p , if p is a theorem of PA, then p is true.

S2* For any p , if it is not the case that p is true, then p is not a theorem of PA.

S3* For a given contradiction, $\perp$, it is not the case that $\perp$ is true.

$\therefore \perp$ is not a theorem of PA. That is, PA is consistent.

Similarly to the scenario in the previous section, we assume that the agent accepts proofs in some theory of truth as providing evidence supporting informal statements about PA. For instance, they accept that a proof in the theory of truth of the formalized soundness statement for PA provides a justification to believe S1*, the statement that PA is sound. The crucial distinction here is that the proofs are not given in CT[PA], some truth theory *syntactically containing* PA: Proofs are provided in some syntax theory, CT[S], which is rich enough to express syntactic facts about PA, and has ‘enough induction’ for the *syntactic language* to formalize the soundness argument. Nevertheless, CT[S] does not syntactically contain PA.²⁷

Against this background, we can investigate our previous questions, i.e., (†) whether the agent can successfully employ the soundness argument to justify the consistency of PA, and (‡) whether they can use it to overcome reasonable open-mindedness about PA’s consistency.

First, we focus on (†). We assume that the premises of the [SOUND*] are justified by proofs in CT[S] of their respective formalizations and check for transmission failure. Since there is no change in the role taken by the syntactic theory CT[S], we have the following by the (NUC):

i* Any proof in CT[S] of a statement p supports p only if, for the relevant underminers u of CT[S], u does not obtain. In particular, the inconsistency of CT[S] cannot obtain. That is, CT[S] must be consistent *at least*.

²⁵I am grateful to an anonymous referee of this journal for asking this question and making its importance for the current discussion explicit.

²⁶This scenario has been suggested by an anonymous referee of this journal.

²⁷As an anonymous referee of this journal correctly pointed out, CT[S] can be CT over some syntactic theory $\mathcal{S}$ interpreting the theory $\mathcal{I}\Sigma_1$, where $\mathcal{I}\Sigma_1$ has induction restricted to formulas of Σ_1 -complexity, i.e., formulas of the form $\exists x(\varphi(x))$, where φ is Δ_0 . φ is Δ_0 just in case all the quantifiers in φ (if there are any) are *bounded*, i.e., are of the form $\forall t < y$ or $\exists t < y$. For a presentation of this theory, see, for instance, (Heck 2015). (Heck 2015) has shown that CT[$\mathcal{I}\Sigma_1$] is sufficient to formalize the soundness argument. Recently, (Łełyk 2022) has shown that even CT[$\mathcal{I}\Delta_0 + exp$] suffices.

- 2* In particular, 1* holds for the $\text{CT}[\mathcal{S}]$ -proof that PA is sound.
- 3* Since we assumed that the premises of [SOUND*] are justified by $\text{CT}[\mathcal{S}]$ -proofs, we have by 2* that $\text{CT}[\mathcal{S}]$ must be consistent. At this point, we make the following remark.
- 4* $\text{CT}[\mathcal{S}]$ recovers full induction for the syntactic language. That is, there is a sub-theory of $\text{CT}[\mathcal{S}]$, let us call it $\mathcal{S}^*$, which is an extension of $\mathcal{S}$ by full induction.
- 5* By 4*, $\mathcal{S}^*$ must be consistent as well—since the consistency of $\text{CT}[\mathcal{S}]$ must be in place by 3*.
- 6* Crucially, the theories $\mathcal{S}^*$ and PA are *equiconsistent*. With 5*, this gives us that PA must be consistent if $\mathcal{S}^*$ is.
- 7* From 5* and 6*, we can conclude that PA must also be consistent.

Therefore, we can see that the generalized version of the soundness argument is also non-transmissive: The assumption that the premises of the soundness argument are justified by proofs in $\text{CT}[\mathcal{S}]$ require the consistency of PA. In other words, the consistency of $\text{CT}[\mathcal{S}]$, which must be in place by (NUC), gives us that the consistency of PA must be in place as well. If my generalized version of the argument is in good standing, then we can, in fact, diagnose soundness arguments as non-transmissive with greater generality. However, some clarifications need to be made concerning this generalized strategy. We first turn to the question of cogency and postpone the clarifications to the end of this section since—as I will show—a generalized argument against cogency can be easily formulated analogously.

In order to check whether the soundness argument is cogent, we assume that the agent is open-minded about the consistency of PA and wishes to overcome this open-minded by employing [SOUND*], where the premises of the soundness argument are taken to be justified by proofs in $\text{CT}[\mathcal{S}]$ of their respective formalizations. We argue in the following manner:

- 1* The agent's open-mindedness about whether PA is consistent amounts (possibly upon reflection) to her open-mindedness about whether, for some relevant underminer u of the consistency of PA, u obtains.
- 2* Since $\mathcal{S}^*$ is equiconsistent with PA, the agent should be at least equally open-minded about whether the consistency of $\mathcal{S}^*$ is undermined.
- 3* But then, since $\mathcal{S}^*$ is a sub-theory of $\text{CT}[\mathcal{S}]$, the agent should be equally open-minded about whether the consistency of $\text{CT}[\mathcal{S}]$ is also undermined.
- 4* Analogously to the reasoning employed in the previous version of this argument, reasonable open-mindedness about whether $\text{CT}[\mathcal{S}]$ is consistent implies at least equal open-mindedness about whether one should rationally accept that $\text{CT}[\mathcal{S}]$ -proofs support $\text{CT}[\mathcal{S}]$ -theorems.
- 5* In particular, one should be equally open-minded about accepting $\text{CT}[\mathcal{S}]$'s support of the statement that PA is sound.

With this, we can conclude that the argument violates Wright's 'new' Cogency Template: Reasonable open-mindedness about the consistency of PA amounts to reasonable open-mindedness about whether the premises of [SOUND*] are justified by CT[S]. In other words, [SOUND*] cannot be successfully employed to overcome reasonable open-mindedness about the consistency of PA.

With this proposed strategy, we can argue that soundness arguments cannot be employed to improve our epistemic situation concerning consistency—in the sense spelt out as (†) and (‡). However, some clarifications are necessary. First, we can see that, according to the proposed analysis, the “entangled” version of the soundness argument [SOUND] presented an *immediate* case of failure: The consistency of our theory of truth CT[S] gave us the consistency of the target theory S directly simply by the fact that S was contained in CT[S]. On the other hand, [SOUND*] presents a less immediate case of failure: The consistency of CT[S] directly gives us the consistency of $\mathcal{S}^*$ *only*. In order to conclude that PA must be consistent as well, we needed the additional information that $\mathcal{S}^*$ and PA are equiconsistent so that the consistency of $\mathcal{S}^*$ (and of CT[S]) is incompatible with PA being inconsistent.²⁸ The additional step involving the notion of equiconsistency is crucial: It gives us a way of generalizing the analysis to the target cases, where the syntactic containment assumption is lifted. Finally, it is essential to point out that the equiconsistency between $\mathcal{S}^*$ and PA is not some ‘external fact’ that *only we can appreciate*: It is provable in CT[S]. This is important because, for this strategy against transmission and cogency to be in good standing, we want the information about the equiconsistency to be accessible—upon reflection—to the agent accepting CT[S], and this is ensured by its provability in CT[S]. With this, we can naturally generalize our original diagnosis to soundness arguments *in general*. Where does this conclusion leave us? The remainder of this article makes some important yet exploratory remarks about my analysis's scope, consequences, and ramifications.

IV. Taking stock

This article focused on the purported epistemic value of soundness arguments, and, in particular, it investigated two central epistemological aspects of such arguments: The question of whether soundness arguments can be employed to justify or overcome reasonable open-mindedness about their conclusion. It has been argued that both questions should receive a negative answer. On one hand, a soundness argument fails to transmit justification due to a

²⁸Alternatively, we could also have argued that CT[S] interprets PA—so that PA must be consistent relative to the consistency of CT[S].

vicious epistemic dependence between the existence of evidential support for its premises and the truth of its conclusion. Moreover, when considering the possibility of overcoming reasonable open-mindedness about the consistency of S, I have shown that it is not possible to be rationally open-minded about their conclusion *whilst* rationally accepting that their premises are defeasibly supported.

In a sense, we have shown that, despite Horsten's entertainment with the possibility of acquiring a justification for the consistency of S *by virtue of* a soundness argument, there is little hope of obtaining an *inferential justification* for consistency by means soundness arguments. Moreover, if one focuses *exclusively* on justifications provided by proofs in accepted theories, our general argument seems to motivate the claim that *it is impossible for soundness arguments to be transmissive or cogent*. There is a sense in which this conclusion seems to point to some kind of *epistemic incompleteness* akin to the incompleteness phenomena: As long as soundness arguments are concerned, we cannot hope to obtain an inferential justification for the consistency of our accepted mathematical theories by means of proofs in other accepted theories. However, a necessary clarification must to be made: Although our diagnosis shows that there is no such hope when employing soundness arguments, it is still open whether we could obtain an inferential justification to believe consistency employing *different consistency arguments*.²⁹ Even though we should be rightly discouraged in our attempt to justify the consistency of our accepted theories by employing soundness arguments, the proposed diagnosis should not in any way support the idea that consistency statements cannot be justified *per se*. In other words, my analysis does not motivate any *scepticism* about consistency. On the contrary: The fact that soundness arguments for consistency are epistemically defective should put pressure on the question of what are *good grounds*, available to us, for believing the consistency of our accepted theory. Importantly, our key questions about transmission and cogency were relative to soundness arguments and a *particular envisaged justification* for its premises. If mathematical justification were exhausted by proofs in theories, then there might not be any justification *whatsoever* for the consistency of many theories we accept. However, I do not believe that justification is exhausted by proofs in our theories. In fact, with many philosophers and mathematicians, I believe we are, in fact, justified in believing in the consistency of many natural foundational theories. Of course, to argue for this would amount to providing a different

²⁹One such important example would be Gentzen's proof of the consistency of Peano arithmetic. Although the epistemological notions and principles employed in this article could be helpful in the analysis of the epistemic value of Gentzen's proof, a generalization of my approach to other types of consistency proofs exceeds the scope of the present investigation. I am grateful to an anonymous referee of this journal for suggesting to say a few words about this issue. For an epistemological investigation of Gentzen's proof, see Daniel Waxman ([unpublished manuscript](#)).

type of justification for such statements: Some mathematicians (such as Gödel) argued that we have some intuition-like faculty that might provide immediate justification for statements, such as the consistency of S. Others argued that we have some *default type* of justification for accepting consistency.³⁰ Finally, other philosophers (most notably Feferman) argued that reflection principles such as consistency statements (and even stronger reflection principles) are, in some sense, implicit commitments.³¹ This is important, as it places this investigation into a broader context within the philosophy and epistemology of mathematics.

Moreover, it should be added that even if it were impossible to acquire an inferential justification for consistency employing soundness arguments *relative to justifications provided by proofs*, one might acquire such a justification by different means. One example could be, for instance, the following scenario.³² One might argue that the truth-theoretic axioms of CT have a “special” epistemic status. For instance, they might be justified (i) *by default* or (ii) *by our understanding* of the concept of truth.³³ With this, one might try to argue that our justification to believe the consistency of PA results from our justification to believe the syntactic theory $\mathcal{S}$ —which does not contain PA—and our justification we have for the concept of truth axiomatized by CT. Whatever justification we have for $\mathcal{S}$ and CT does not seem to support the consistency of PA directly, i.e., *without employing the soundness argument*. Let me comment briefly on this possibility. Suppose that such an argument would be transmissive. Then, I believe this argument would ensure that our justification for believing in the consistency of PA would be quite minimal. For instance, our justification could result from some justification to accept a very small amount of induction, together with some justification to accept the concept of truth embodied by CT. However, I suspect that the soundness argument will end up being non-transmissive, as long as either (a) *conditions such as (NUC) are in place* or (b) the justification to believe the premises of [SOUND] is provided by *proofs of formalized statements* in the theory of truth. In the concrete scenario, (NUC) would give us that proofs provided in $\text{CT}[\mathcal{S}]$ justify the soundness of PA only if $\text{CT}[\mathcal{S}]$ is consistent. Therefore, my proposed argument for failure should apply. A possibility to argue that [SOUND] is transmissive would be to either reject (a) or (b). Concerning (a), one could, for example, *deny* or *weaken* the (NUC). With this, there is no *prima facie* immediate way to argue for failure.

³⁰For instance, (Pedersen 2021), (Horsten 2021) and (Zicchetti 2022). To argue that we have such independent justification for consistency statements exceeds the scope of our investigation.

³¹See (Feferman 1962, 1964, 1991). For a recent discussion of implicit commitments in this context, see (Dean 2014), (Horsten and Leigh 2017), (Nicolai and Piazza 2018), or (Lelyk and Nicolai 2022).

³²I am grateful to an anonymous referee of this journal for suggesting this scenario.

³³The idea that our acceptance of some concepts of truth might not need any justification, or might be justified by default, has been investigated, for instance, in (McGee 2005) and (Fischer et al. 2019).

However, this would amount to the rejection—to some degree—of the natural intuition that proofs justify theorems relative to the existence of collateral information. Equivalently, one would have to provide an argument as to why underminers should not play a role—or should play a lesser role—in determining the existence of justifications by means of proofs. With respect to (b), one possibility would be to argue that our justification to believe the premises of [SOUND] is not given by proofs in a theory of truth, but has a *different source*. For example, one might argue that our understanding of the concept of truth embodied by CT *alone* gives us a justification to believe the *informal* soundness statement without the need of any proof.³⁴

Finally, what should this diagnosis tell us about epistemology (of mathematics)? The proposed analysis provides more substance to the natural intuition expressed in (Dummett 1978), (Girard 1987), and (Wright 1994), helping us to make some substantive progress on the question about the *epistemic value* of soundness arguments. Moreover, my conclusions are in harmony with considerations in epistemology about transmission, failure, and cogency *in general*. Similarly to Wright's investigation of transmission and cogency in (Wright 2012), the issues investigated in (†) and (‡) are two separate epistemological questions, whose answers require deeper reflection on the epistemological background in place.³⁵ Moreover, similarly to [MOORE]'s diagnosis provided by Wright, the answers to (†) and (‡) converge—as I think they should.³⁶ This diagnosis suggests that *mathematical arguments* and the epistemology of mathematics might not be too different than usual, non-mathematical arguments and their epistemology, making some room for a more unified and encompassing future investigation of the Philosophy of Mathematics and its Epistemology.³⁷

³⁴Although I suspect that a rejection of either (a) or (b) would result in a transmissive argument, most of the philosophical work would be needed to explain why we should reject either (a) or (b). To investigate the possibility of employing a soundness argument for consistency in scenarios avoiding either (a) or (b) would exceed the scope and context of the present article.

³⁵My questions (†) and (‡) seem to be quite similar to (i) in (Wright 2012; 451) and (vi) in (Wright 2012; 459).

³⁶See, (Wright 2012; 460). However, (Wright 2012) seems to suggest that there is a possibility of (†) and (‡) to diverge, as he says the following: 'Moore's Proof, for instance, might indeed be serviceable for the provision of warrant to believe in an external world while useless for underwriting claim to such warrant' (Wright 2012; 471). I suspect that the answer to the question of whether or to what degree (†) and (‡) converge is going to depend on the kind of No Underminer Conditions adopted. I am grateful to an anonymous referee of this journal for asking to address why (†) and (‡) are treated separately.

³⁷Versions of this article have been presented at a number of research seminars and conferences in several places including the University of Konstanz, the University of Warsaw, the Munich Center for Mathematical Philosophy, the IUSS Pavia, Yonsei University, and King's College London. I am grateful to all the participants for their helpful comments and questions. In particular, my gratitude goes to Crispin Wright, Leon Horsten, Martin Fischer, Nikolaj Pedersen, Dan Waxman, and Henri Galinon for the many stimulating philosophical discussions about this article's topic. Moreover, I am grateful to all members of our research project in Warsaw,

References

- Boolos, G. S. (1993) *The Logic of Provability*, Cambridge University Press.
- Carter, J. A. (2016) 'Meta-Epistemic Defeat', *Synthese*, 195/7: 2877–96.
- Chandler, J. (2009) 'The Transmission of Support: A Bayesian Re-Analysis', *Synthese*, 176/3: 333–43.
- Cieśliński, C. (2010) 'Truth, Conservativeness, and Provability', *Mind*, 119: 409–22.
- Cieśliński, C. (2017) *The Epistemic Lightness of Truth, Deflationism and its Logic*, Cambridge University Press.
- Davies, M. (2009) 'Two Purposes of Arguing and Two Epistemic Projects', In: *Minds, Ethics, and Conditionals*, pp. 337–84. Oxford University Press: Oxford.
- Dean, W. (2014) 'Arithmetical Reflection and the Provability of Soundness', *Philosophia Mathematica*, 23/1: 31–64.
- Dretske, F. (2005) 'Is Knowledge Closed Under Known Entailment? The Case Against Closure', In: M. Steup and E. Sosa (eds.) *Contemporary Debates in Epistemology*, 13–26. Blackwell.
- Dummett, M. (1978) *Truth and other Enigmas*, London: Harvard University Press.
- Feferman, S. (1962) 'Transfinite Recursive Progressions of Axiomatic Theories', *The Journal of Symbolic Logic*, 27/3: 259–316.
- Feferman, S. (1964) 'Systems of Predicative Analysis', *The Journal of Symbolic Logic*, 29/1: 1–30.
- Feferman, S. (1991) 'Reflecting on Incompleteness', *The Journal of Symbolic Logic*, 56: 1–47.
- Fischer, M., Horsten, L., and Nicolai, C. (2019) 'Hypatia's Silence Truth Justification and Entitlement', *Noûs*, 55: 62–85.
- Franzen, T. (2004) *Inexhaustibility*, Lecture Notes in Logic. CRC Press.
- Fujimoto, K. (2017) 'Deflationism Beyond Arithmetic', *Synthese*, 196/3: 1045–69.
- Fujimoto, K. (2021) 'The Function of Truth and the Conservativeness Argument', *Mind*, 131/521: 129–57.
- Fujimoto, K. (2022) 'On the Logicity of Truth', *The Philosophical Quarterly*, 72/4: 853–74.
- Girard, J.-Y. (1987) *Proof Theory and Logical Complexity*. Neapel: Bibliopolis.
- Halbach, V. (2014) *Axiomatic Theories of Truth*, Cambridge, UK: Cambridge University Press, revised edn.
- Hale, B. and Wright, C. (2005) 'Logicism in the Twenty-First Century', In: S. Shapiro (ed.) *The Oxford Handbook of Philosophy of Mathematics and Logic*, pp. 125–134. New York.
- Hawthorne, J. (2004) *Knowledge and Lotteries*, Oxford University Press.
- Heck, R. G. (2015) 'Consistency and the Theory of Truth', *The Review of Symbolic Logic*, 8/3: 424–66.
- Horsten, L. (2021) 'On Reflection', *The Philosophical Quarterly*, 71/4: pqaa083.
- Horsten, L. and Leigh, G. E. (2017) 'Truth is Simple', *Mind*, 126/501: 195–232.
- Ketland, J. (2005) 'Deflationism and the Gödel Phenomena: Reply to Tennant', *Mind*, 114/453: 75–88.
- Ketland, J. (2010) 'Truth, Conservativeness, and Provability: Reply to Cieśliński', *Mind*, 119/474: 423–36.
- Leigh, G. E. and Nicolai, C. (2013) 'Axiomatic Truth, Syntax and Metatheoretic Reasoning', *The Review of Symbolic Logic*, 6/4: 613–36.
- Lelyk, M. (2022) 'Model Theory and Proof Theory of the Global Reflection Principle', *The Journal of Symbolic Logic*, 88/2: 738–79.
- Lelyk, M. and Nicolai, C. (2022) 'A Theory of Implicit Commitment', *Synthese*, 200: 284. <https://doi.org/10.1007/s11229-022-03601-5>
- McGee, V. (2005) 'Two Conceptions of Truth? Comment', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 124/1: 71–104.
- Moore, G. E. (1939) 'Proof of an External World', *Proceedings of the British Academy*, 25/5: 273–300.

Epistemic and Semantic Commitments of Foundational Theories (PI: Cezary Cieśliński), for their support, and to the National Science Centre, Poland (MAESTRO project number 2019/34/A/HS1/00399) for their financial support. Finally, I am thankful to two anonymous referees of this journal whose comments and suggestions improved the quality of this paper significantly.

- Moretti, L. (2010) 'Wright, Okasha and Chandler on Transmission Failure', *Synthese*, 184/3: 217–34.
- Moretti, L. and Piazza, T. (2017) 'Defeaters in Current Epistemology: Introduction to the Special Issue', *Synthese*, 195/7: 2845–54.
- Moretti, L. and Piazza, T. (2018) 'Transmission of Justification and Warrant', In E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2018 edn.
- Myhill, J. (1960) 'Some Remarks on the Notion of Proof', *The Journal of Philosophy*, 57: 461–71.
- Nicolai, C. and Piazza, M. (2018) 'The Implicit Commitment Thesis of Arithmetical Theories and its Semantic Core', *Erkenntnis*, 84: 913–37.
- Okasha, S. (2004) 'Wright on the Transmission of Support: A Bayesian Analysis', *Analysis*, 64/2: 139–46.
- Pedersen, N. J. L. (2021) 'Cornerstone Epistemology: Scepticism, Mathematics, Non-Evidentialism, Consequentialism, Pluralism', In N. J. L. Pedersen and L. Moretti (eds.) *Non-Evidentialist Epistemology*. Brill.
- Pollock, J. (1970) 'The Structure of Epistemic Justification', *American Philosophical Quarterly*, 4: 62–78.
- Priest, G. (2006) In *Contradiction*, Oxford University Press, 2 edn.
- Pryor, J. (2000) 'The Skeptic and the Dogmatist', *Noûs*, 34/4: 517–49.
- Pryor, J. (2004) 'What's Wrong with Moore's Argument', *Philosophical Issues*, 14/1: 349–77.
- Shapiro, S. (1998) 'Proof and Truth: Through Thick and Thin', *The Journal of Philosophy*, 95: 493–521.
- Tennant, N. (2005) 'Deflationism and the Gödel Phenomena: Reply to Ketland', *Mind*, 114/453: 89–96.
- Tennant, N. (2010) 'Deflationism and the Gödel Phenomena: Reply to Cieśliński', *Mind*, 119/474: 437–50.
- Tucker, C. (2010) 'When Transmission Fails', *Philosophical Review*, 119/4: 497–529.
- Waxman, D. (2017) 'Deflationism, Arithmetic, and the Argument from Conservativeness', *Mind*, 126/502: 429–63.
- Waxman, D. (unpublished) 'Did Gentzen Prove the Consistency of Arithmetic?', Unpublished Manuscript.
- Williamson, T. (2000) *Knowledge and its Limits*, Oxford University Press.
- Wright, C. (1994) 'About "The Philosophical Significance of Gödel's Theorem": Some Issues', In B. McGuinness and G. Oliveri (eds.) *The Philosophy of Michael Dummett*, pp. 167–204.
- Wright, C. (2000) 'Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Putnam's Proof', *Philosophical Issues*, 10: 140–163.
- Wright, C. (2002) '(Anti-)Sceptics Simple and Subtle: G.E. Moore and John McDowell*', *Philosophy and Phenomenological Research*, 65/2: 330–48.
- Wright, C. (2003) 'Some Reflections on the Acquisition of Warrant by Inference', In S. Nuccetelli (ed.) *New Essays on Semantic Externalism and Self-Knowledge*, pp. 57–78. MIT Press.
- Wright, C. (2004) 'Warrant for Nothing (and Foundations for Free)?', *Proceedings of the Aristotelian Society, Supplementary Volumes*, 78: 167–245.
- Wright, C. (2012) 'Replies Part IV: Warrant, Transmission and Entitlement', In A. Coliva (ed.) *Mind, Meaning, and Knowledge. Themes from the Philosophy of Crispin Wright*, pp. 451–86. Oxford University Press.
- Wright, C. (2014) 'On Epistemic Entitlement (II): Welfare State Epistemology', In D. Dodd and E. Zardini (eds.) *Scepticism and Perceptual Justification*, pp. 213–47. Oxford University Press.
- Wright, C. (2022) 'Closure and Transmission Again'. In D. Pritchard and M. Jope (eds.) *New Perspectives on Epistemic Closure*. Routledge.
- Zicchetti, M. (2022) *The Epistemology of Meta-theoretic Properties of Mathematical Theories: Consistency, Soundness, Categoricity*, Ph.D. thesis, University of Bristol.