

Rosenkranz's Logic of Justification and unprovability

Jan Heylen

April 4, 2020

Abstract

Rosenkranz has recently proposed a logic for propositional, non-factive, all-things-considered justification, which is based on a logic for the notion of being in a position to know (The Structure of Justification, *Mind*, 127(506), 309–338). Starting from three quite weak assumptions in addition to some of the core principles that are already accepted by Rosenkranz, I prove that, if one has positive introspective and modally robust knowledge of the axioms of minimal arithmetic, then one is in a position to know that a sentence is not provable in minimal arithmetic or that the negation of that sentence is not provable in minimal arithmetic. This serves as the formal background for an example that calls into question the correctness of Rosenkranz's logic of justification.

Keywords justification; being in a position to know; closure; unprovability

1 Introduction

There is a need for a logic of justification, as Artemov (2008, pp. 477–478) pointed out:

Meanwhile, commencing from seminal works (Hintikka, 1962; von Wright, 1951), the notions of Knowledge and Belief have acquired formalization by means of modal logic with atoms **KF** (*F is known*) and **BF** (*F is believed*). [...] The resulting Epistemic Logic has been remarkably successful in terms of developing a rich mathematical theory and applications (cf. Fagin et al., 1995; Meyer & van der Hoek, 1995, and other sources). However, the notion of justification, which has been an essential component of epistemic studies, was conspicuously absent in the mathematical models of knowledge within the epistemic logic framework. [...] This lack of a justification component has, perhaps, contributed to a certain gap between epistemic logic and mainstream epistemology (Hendricks, 2003, 2005).

Artemov's own Justification Logic is an attempt to fill the gap. Rosenkranz's logic of justification can be seen as a second attempt (Rosenkranz, 2018). Smith gives it a third shot (Smith, 2017). This article contains a critical discussion of Rosenkranz's logic of justification, which will be presented in the next section.

2 Rosenkranz's logic of justification

Rosenkranz does not propose a logic for every kind of justification, but only for propositional, non-factive, all-things-considered justification. *Propositional* justification is about the justification of propositions, whereas doxastic justification is about the justification of beliefs (in propositions). Let 'J' be a propositional, non-factive, all-things-considered justification operator. Rosenkranz uses as a criterion for *non-factive* justification that the following principle has to fail:

$$\mathbf{T}_J \ J\phi \rightarrow \phi.$$

In other words, Rosenkranz is concerned with a notion of justification that does not in general entail truth. The criterion for *all-things-considered* justification employed by Rosenkranz is that the following principle has to hold:

$$\mathbf{D}_J \ J\neg\phi \rightarrow \neg J\phi.$$

In other words, Rosenkranz is concerned with a notion of justification that rules out that there is justification for both a sentence and its negation.

Rosenkranz lays down two more constraints, which connect the notion of propositional, non-factive, all-things-considered justification with the notion of being in a position to know, which was introduced by Cohen (1999) and Williamson (2000, p. 95). About the latter notion Rosenkranz (2007, p. 69) writes the following (in a different context):

We may accordingly lay down three necessary conditions for one's being in a position to know p :

[(i)] p is true

[(ii)] One is physically and psychologically capable of knowing p

[(iii)] Nothing stands in one's way of successfully exercising these capabilities

Let 'K' be an operator expressing 'being in a position to know'. The first additional constraint is the following:

$$\mathbf{K-J} \ K\phi \rightarrow J\phi.$$

In other words, being in a position to know something entails that there is propositional, non-factive, all-things-considered justification for it. The second additional constraint is the following:

$$J\neg K\phi \rightarrow \neg J\phi. \quad (1)$$

Moreover, if there is propositional, non-factive, all-things-considered justification for the fact that one is not in a position to know a sentence, then there is no propositional, non-factive, all-things-considered justification for the sentence. With the above criteria in mind, Rosenkranz proposes an account of justification that comprises the following five principles:

E $J\phi \leftrightarrow \neg K\neg K\phi$

T_K $K\phi \rightarrow \phi$

RN_K If $\vdash \phi$, then $\vdash K\phi$.

RM_K If $\vdash \phi \rightarrow \psi$, then $\vdash K\phi \rightarrow K\psi$.

LUM $\neg K\neg K\phi \rightarrow K\neg K\neg K\phi$

In other words, there is propositional, non-factive, all-things-considered justification for a sentence if and only if one is not in a position to know that one is not in a position to know that sentence (E). Being in a position to know is factive (T_K), closed under theoremhood (RN_K) and it distributes over provable material implications (RM_K). If one is not in a position to know that one is not in a position to know something, then one is in a position to know that one is not in a position to know that one is not in a position to know that thing (LUM).

Note that RM_K is weaker than the ‘normal’ closure principle (K_K), which says that being in a position to know is closed under material implication:

K_K $K(\phi \rightarrow \psi) \rightarrow (K\phi \rightarrow K\psi)$.

The latter is deemed too strong for the notion of being in a position to know (Heylen, 2016; Rosenkranz, 2016). In particular, the following consequence of K_K and RN_K is deemed unacceptable:

$$(K\phi \wedge K\psi) \rightarrow K(\phi \wedge \psi). \quad (2)$$

I will not go into the details of the argument, but it is important for later purposes to know that the counterexample is of the form $\neg \kappa p$ (with κ the knowledge operator) for ϕ and p for ψ .

Clearly, postulates E and LUM are the substantives ones. Postulate T_K is generally accepted – see also condition (i) above. Rule RN_K is an idealisation: in case ϕ is a formula that contains a yottabyte or 1000⁸ proposition letters, it is doubtful that a *real person* is physically and psychologically capable of knowing ϕ , given that according to a recent estimate our brain can only contain one petabyte or

1000⁵ bytes of information. Rule RM_{κ} is also an idealisation: in case ϕ is the conjunction of the axioms of Robinson arithmetic and ψ is a theorem of Robinson arithmetic that contains a yottabyte of proposition letters, a real person is physically and psychologically capable of knowing ϕ , but it is doubtful that a real person is physically and psychologically capable of knowing ψ . Still, idealisations have their uses.

Principles D_J , $K\text{-}J$ and (1) are provable within the system. Even though he does not provide formal models, Rosenkranz claims that E explains why J is non-factive: $\neg K\text{-}K$ is not factive.

In Section 5 I will argue that Rosenkranz's logic of justification is incorrect. In order to prepare for that argument I will first present and discuss three assumptions in Section 3. Then I will prove in Section 4 a lemma that, under those assumptions, if one has positive introspective and modally robust knowledge of the axioms of minimal arithmetic and if one is in a position to know that one is not in a position to know a certain sentence (or its negation), then one is in a position to know that sentence (or its negation) is not provable in minimal arithmetic. This will be used in Section 5 to prove a theorem that says that, under the same assumptions and core principles of Rosenkranz's logic of justification, if one has positive introspective and modally robust knowledge of the axioms of minimal arithmetic, then one is in a position to know that a sentence is not provable in minimal arithmetic or that the negation of that sentence is not provable in minimal arithmetic. It will be argued that this presents a serious challenge to Rosenkranz's logic of justification.

3 The assumptions

In this section I will introduce and defend three quite minimal assumptions that go beyond Rosenkranz's logic of justification.

First, we will assume that, if one knows something, then one is in a position to know it:

$$\kappa\phi \rightarrow K\phi. \quad (3)$$

Recall that κ is the knowledge operator. It is easy to see that, if one knows p , then at the very least the three necessary conditions for being in a position to know it are fulfilled. Whatever other conditions one may deem sufficient for being in a position to know that p , surely knowing that p is one.

Let us now add the notion of being provable in an axiomatic system into the mix. In Section 5, the axiomatic theory of minimal arithmetic, Q , will be used. The axioms of Q are nine axioms for the successor function, the addition function, the multiplication function, and the is-smaller-than relation. See (Boolos et al., 2007, ch. 16) for an exposition of this theory. Since Q has finitely many axioms, let us use Q for the conjunction of those axioms. Suppose that ϕ is a theorem of Q . In symbols:

$$\vdash Q \rightarrow \phi.$$

Let us use the \Box -operator to express that something is logically provable. So, the above can be rewritten as follows:

$$\Box(Q \rightarrow \phi).$$

Let us abbreviate the above as:

$$\Box_Q \phi.$$

Under that interpretation of \Box_Q , the following is an analytic truth:

$$(Q \wedge \Box_Q \phi) \rightarrow \phi.$$

Note that, if one combines the above with RM_K , then one can derive the following:

$$K(Q \wedge \Box_Q \phi) \rightarrow K\phi.$$

However, given the intended meaning of \Box_Q and given RM_K , a stronger closure principle is also warranted:

$$(KQ \wedge \Box_Q \phi) \rightarrow K\phi. \quad (4)$$

The above is in the end nothing but a representation in the object language of rule RM_K (applied to the logical consequences of the axioms of Q), formulated in the metalanguage. Indeed, using \Box one could represent RM_K as follows:

$$\Box(\phi \rightarrow \psi) \rightarrow \Box(K\phi \rightarrow K\psi).$$

With ϕ being replaced by Q and ψ by ϕ , one obtains the following:

$$\Box(Q \rightarrow \phi) \rightarrow \Box(KQ \rightarrow K\phi),$$

which can be abbreviated as

$$\Box_Q \phi \rightarrow \Box(KQ \rightarrow K\phi).$$

Let us now invoke T_\Box , namely

$$T_\Box \Box\phi \rightarrow \phi,$$

which expresses that a sentence that is logically provable is also true. Assuming T_\Box , one can then derive that

$$\Box_Q \phi \rightarrow (KQ \rightarrow K\phi),$$

which is tautologically equivalent to (4). Principle (4) is the second assumption.

Now I will make use of a third assumption, which includes a *counterfactual* conditional that will be formalised with the help of $\Box \rightarrow$. The assumption is that,

- if one knows that ϕ and

- if one is in a position to know that ψ and
- if one were to know that ψ , then one would (still) know that ϕ ,

then one is a position to know that ϕ and ψ . Here is its formalisation:

$$(\kappa\phi \wedge K\psi \wedge (\kappa\psi \Box \rightarrow \kappa\phi)) \rightarrow K(\phi \wedge \psi), \quad (5)$$

First, note that (5) fits within a nice ordering of agglomeration principles. Consider the following variations on the antecedent of (2), including the antecedent of (2) itself and the antecedent of (5):¹

- $\kappa\phi \wedge \kappa\psi$;
- $\kappa\phi \wedge K\psi \wedge (\kappa\psi \Box \rightarrow \kappa\phi)$;
- $\kappa\phi \wedge K\psi$;
- $K\phi \wedge K\psi \wedge (\kappa\phi \Box \rightarrow K\psi)$;
- $K\phi \wedge K\psi$.

With the above formulas and $K(\phi \wedge \psi)$ one can form (2), (5) and the following other agglomeration principles:

$$(\kappa\phi \wedge \kappa\psi) \rightarrow K(\phi \wedge \psi) \quad (6)$$

$$(\kappa\phi \wedge K\psi) \rightarrow K(\phi \wedge \psi) \quad (7)$$

$$(K\phi \wedge K\psi \wedge (\kappa\phi \Box \rightarrow K\psi)) \rightarrow K(\phi \wedge \psi) \quad (8)$$

One can prove that:

- (2) entails (8);
- (8) entails (7);
- (7) entails (5);
- (5) entails (6).

All one needs to prove the above entailments are conjunction elimination, $\kappa\phi \rightarrow K\phi$, cf. (3), and that $\theta \wedge \rho$ entails $\theta \Box \rightarrow \rho$.

Second, (5) is immune against the kind of counterexamples given in (Heylen, 2016) and (Rosenkranz, 2016). Recall that the counterexample involves the following substitution: $\neg\kappa p$ for ϕ and p for ψ . Not only will that substitution deliver a counterexample against (2) but also against (7) and, therefore, also against (8). Yet, it does not produce a counterexample against (5). Let ϕ be $\neg\kappa p$ and let ψ be

¹For brevity's sake, I leave out the 'mirror' case in which one starts 'lowering' via the *second* conjunct. If one were to add those case, one would get a kite-shaped structure.

p . Then $\kappa\psi \Box\rightarrow \kappa\phi$ is $\kappa p \Box\rightarrow \kappa\neg p$. The latter is false, if $\kappa\psi$ is true. Note that, if $\kappa\psi$, then it is possible to know that ψ :

$$\kappa\psi \rightarrow \Diamond\kappa\psi, \quad (9)$$

with the possibility operator \Diamond perhaps made relative to one's physical and psychological capabilities (Heylen, 2016, p. 64). So, if κp , then $\Diamond\kappa p$. Therefore, one does not have to consider the case in which the counterfactual conditional is trivially true, because its antecedent is impossible. Clearly, in the closest worlds in which p is known, it is not known that p is not known, since otherwise in the closest worlds in which it is known that p , it would also be unknown that p (because of factivity).

Third, it is risky to try to argue against (5), since it is 'close' to (6). Assume that the antecedent of (5) is true, i.e., assume that:

$$\kappa\phi \wedge \kappa\psi \wedge (\kappa\psi \Box\rightarrow \kappa\phi).$$

It follows by simplification from the antecedent of (5) that

$$\kappa\psi.$$

In other words, one is in a position to know that ψ . Since one's being in a position to know that ψ entails that one is physically and psychologically capable of knowing that ψ and that nothing stands in one's way of successfully exercising those capabilities, it seems that not only is it possible that one knows that ψ – cf. (9) – but also that it is a *close possibility* that one knows that ψ . This can be expressed as follows:

$$\neg(\kappa\psi \Box\rightarrow \neg\kappa\psi) \quad (10)$$

The above is true if and only if among the closest worlds in which one is in a position to know that ψ , which includes the actual world, there is at least one world at which it is known that ψ . It also follows by simplification from the antecedent of (5) that

$$(\kappa\psi \Box\rightarrow \kappa\phi).$$

In other words, the closest worlds in which one knows that ψ are worlds in which one knows that ϕ as well. Therefore, it follows from the antecedent of (5) and (10) that, among the closest worlds, there is at least one possible world in which one knows both ϕ and ψ . On the basis of (6) one may then infer that, among the closest worlds, there is at least one possible world in which one is in a position to know that ϕ and ψ . Hence, it follows from (6) and (10) that:

$$\neg((\kappa\phi \wedge \kappa\psi \wedge (\kappa\psi \Box\rightarrow \kappa\phi)) \Box\rightarrow \neg\kappa(\phi \wedge \psi)). \quad (11)$$

Of course, a close possibility is not the same as an actual truth, so we have not derived (5). However, it does mean that anyone who wants to accept (6) and (10) but reject (5) will have to make sure the putative argument against (5) does not

also work against (11), because then (6) or (10) would have to be rejected as well. We have already seen a reason to accept (10). Rejecting (6) may put someone who is also endorsing RN_K and RM_K in a bit of a pickle, since those other closure principles require idealised agents, who have the physical and psychological ability to know all the logical truths and (single-premise) logical consequences of those, and ideal circumstances, in which nothing stands in the way for those agents to successfully exercise those abilities.

With (3), (4) and (5) in place, we are in position to start building the case against Rosenkranz's logic of justification and being in a position to know.

4 The lemma

We are ready to prove the lemma that says that, under the assumptions of section 3, if one has positive introspective and modally robust knowledge of the axioms of minimal arithmetic and if one is in a position to know that is not in a position to know a certain sentence (or its negation), then one is in a position to know that sentence is not provable in minimal arithmetic or that the negation of that sentence is not provable in minimal arithmetic.

Lemma 1. *If (3), (4) and (5) and if they are closed under RM_K , then*

$$(\kappa\kappa Q \wedge K\neg K\phi \wedge (\kappa(\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \kappa\kappa Q)) \rightarrow K\neg \Box_Q \phi$$

and

$$(\kappa\kappa Q \wedge K\neg K\neg\phi \wedge (\kappa(\neg KQ \vee \neg \Box_Q \neg\phi) \Box \rightarrow \kappa\kappa Q)) \rightarrow K\neg \Box_Q \neg\phi.$$

Proof. Suppose that one knows that one knows (the conjunction of) the axioms of Q:

$$\kappa\kappa Q \tag{12}$$

Suppose furthermore that one is in a position to know that one is not in a position to know that ϕ :

$$K\neg K\phi. \tag{13}$$

It follows by tautological reasoning from (4) that:

$$\neg K\phi \rightarrow (\neg KQ \vee \neg \Box_Q \phi).$$

An application of RM_K yields the following:

$$K\neg K\phi \rightarrow K(\neg KQ \vee \neg \Box_Q \phi). \tag{14}$$

Assumption (13) together with (14) entails that one is in a position to know that one is not in a position to know the (conjunction of the) axioms of Q or ϕ is not logically provable from the (conjunction of the) axioms of Q:

$$K(\neg KQ \vee \neg \Box_Q \phi). \tag{15}$$

Next, assume that, if one were to know that (15), then it would (still) be known that one knows (the conjunction of) the axioms of Q :

$$\kappa (\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \kappa \kappa Q. \quad (16)$$

Given (5) and (12)–(16), one can then derive that:

$$K ((\neg KQ \vee \neg \Box_Q \phi) \wedge \kappa Q).$$

Note that

$$\vdash ((\neg KQ \vee \neg \Box_Q \phi) \wedge \kappa Q) \rightarrow \neg \Box_Q \phi,$$

because of (3). Therefore, one can apply RM_K to derive that

$$K \neg \Box_Q \phi.$$

Chaining everything together leads to the following preliminary conclusion:

$$(\kappa \kappa Q \wedge K \neg K \phi \wedge (\kappa (\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \kappa \kappa Q)) \rightarrow K \neg \Box_Q \phi. \quad (17)$$

The second conjunct of the consequent of the lemma follows by substitution. \square
 \square

In other words, if

- one knows that one knows Q and
- one is in a position to know that one is not in a position to know ϕ ($\neg \phi$) and
- if one were to know that one is not in a position to know Q or that ϕ ($\neg \phi$) is not logically provable from Q , then one would still know that one knows the axioms of Q ,

then one is in a position to know that ϕ ($\neg \phi$) is not logically provable from Q .

5 The theorem and its philosophical significance

Building on Lemma 1 I will now prove a theorem that says that, under the assumptions of section 3 and core principles of Rosenkrantz's logic of justification, if someone has positive introspective and modally robust knowledge of the axioms of Q , then that person is in position to know that a sentence is not provable in Q or that the negation of that sentence is not provable in Q .

Theorem 1. *If (3), (4) and (5) and if they are closed under RM_K and E and D_J , then*

$$((\kappa \kappa Q \wedge (\kappa (\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \kappa \kappa Q)) \wedge (\kappa (\neg KQ \vee \neg \Box_Q \neg \phi) \Box \rightarrow \kappa \kappa Q)) \rightarrow (K \neg \Box_Q \phi \vee K \neg \Box_Q \neg \phi).$$

Proof. It is a tautological consequence of Lemma 1 that

$$\begin{aligned} &(((\mathbb{K}\mathbb{K}Q \wedge (\mathbb{K}(\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \mathbb{K}\mathbb{K}Q)) \wedge (\mathbb{K}(\neg KQ \vee \neg \Box_Q \neg \phi) \Box \rightarrow \mathbb{K}\mathbb{K}Q)) \wedge \\ &(\neg K\neg \Box_Q \phi \wedge \neg K\neg \Box_Q \neg \phi)) \rightarrow (\neg K\neg K\phi \wedge \neg K\neg K\neg \phi). \end{aligned}$$

One can use E to deduce from the above with the help of tautological reasoning that

$$\begin{aligned} &(((\mathbb{K}\mathbb{K}Q \wedge (\mathbb{K}(\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \mathbb{K}\mathbb{K}Q)) \wedge (\mathbb{K}(\neg KQ \vee \neg \Box_Q \neg \phi) \Box \rightarrow \mathbb{K}\mathbb{K}Q)) \wedge \\ &(\neg K\neg \Box_Q \phi \wedge \neg K\neg \Box_Q \neg \phi)) \rightarrow (J\phi \wedge J\neg \phi). \end{aligned}$$

One can use D_J to deduce from the above with the help of tautological reasoning that

$$\begin{aligned} &(((\mathbb{K}\mathbb{K}Q \wedge (\mathbb{K}(\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow \mathbb{K}\mathbb{K}Q)) \wedge (\mathbb{K}(\neg KQ \vee \neg \Box_Q \neg \phi) \Box \rightarrow \mathbb{K}\mathbb{K}Q)) \wedge \\ &(\neg K\neg \Box_Q \phi \wedge \neg K\neg \Box_Q \neg \phi)) \rightarrow (J\phi \wedge \neg J\phi). \end{aligned}$$

From the latter the lemma follows by tautological reasoning again. \square \square

The consequent of Theorem 1 rules out that anyone has positive introspective and modally robust knowledge of the axioms of Q while not being in a position to know that a sentence is not provable in Q and not being in a position to know that the negation of that sentence is not provable in Q. Yet, it is possible that there are such people.

Alice is a bright and self-confident young woman who knows the concepts of knowledge and being in a position to know. She has taken an introductory course in logic. She has done very well. She decides to take an intermediate course in logic. The teacher, Bob, has introduced her to the idea of axiomatic proofs. For illustration, Alice and her fellow students are presented with a simple axiomatic theory, namely the theory of minimal arithmetic, Q. Bob is a reliable teacher and the axioms seem self-evident, so Alice comes to know the (conjunction of the) axioms of Q. Better still, Alice reflects on this and she comes to know that she knows the (conjunction of the) axioms of Q, i.e. (12) is the case.

The teacher has a strategy. First Bob gives them arithmetical sentences that are *provable* in Q, e.g.

$$\begin{aligned} 2 + 3 &= 3 + 2 \\ 4 \cdot 5 &= 5 \cdot 4 \\ \exists z (6 &= z'). \end{aligned}$$

In each case Alice managed to prove those sentences. But then he writes down an arithmetical sentence that is neither *provable* nor *refutable* (i.e. undecidable) in Q, viz.

$$\forall x \forall y (((x + y = y + x) \wedge (x \cdot y = y \cdot x)) \vee (x = \mathbf{0} \vee \exists z (x = z'))). \quad (18)$$

Bob tells the students neither that (18) is unprovable in minimal arithmetic nor that its negation is also unprovable in minimal arithmetic. The secret pedagogical goal is to let the students, including Alice, fail at proving (18) and, subsequently, teach them about *undecidability*, standard and non-standard models of arithmetic, and what to add to Q so that (18) can be decided (viz. the principle of induction). This time Alice does not find a proof within the allotted time.

How could Alice be in a position to know that (18) is not a theorem of Q other than by being able to construct a (non-standard) model in which (18) is false and the axioms of Q are true? However, Alice has not been introduced to models yet, let alone non-standard models of arithmetic in which the axioms of Q are true but (18) is false (in contradistinction to the standard model of Q). In particular, Alice has not been made acquaintance with the notions of *ordinal addition*, *ordinal multiplication* and *ordinal succession*. So, we have the following:

$$\neg K \neg \Box_Q (18). \quad (19)$$

Furthermore, if Alice were to know she is not in a position to know the (conjunction of the) axioms of Q or that (18) is not a theorem of Q , then she would (still) know that she knows the (conjunction of the) axioms of Q :

$$K (\neg K Q \vee \neg \Box_Q (18)) \Box \rightarrow KKQ \quad (20)$$

For the closest worlds in which Alice comes to know that she is not in a position to know the (conjunction of the) axioms of Q or that (18) is not a theorem of Q are quite plausibly those worlds in which she has been given *further instruction* by Bob and she comes to know that (18) is not a theorem of Q , from which she then deduces the disjunction. It is highly implausible that Alice has *in the course of further instruction* lost her (positive introspective) knowledge of the (conjunction of the) axioms of Q .

How could Alice be in a position to know that $\neg(18)$ is not a theorem of Q other than by being able to prove that (18) is true if the induction scheme is added? However, Alice has not been introduced to mathematical induction yet. So, we have the following:

$$\neg K \neg \Box_Q \neg (18). \quad (21)$$

Furthermore, if Alice were to know that she is not in a position to know the (conjunction of the) axioms of Q or that $\neg(18)$ is not a theorem of Q , then she would (still) know that she knows the (conjunction of the) axioms of Q :

$$K (\neg K Q \vee \neg \Box_Q \neg (18)) \Box \rightarrow KKQ \quad (22)$$

For the closest worlds in which Alice comes to know that she is not in a position to know the (conjunction of the) axioms of Q or that $\neg(18)$ is not a theorem of Q are quite plausibly those worlds in which she has been given *further instruction* by Bob and she comes to know that $\neg(18)$ is not a theorem of Q , from which she then deduces the disjunction. Again, it is highly implausible that Alice has *in*

the course of further instruction lost her (positive introspective) knowledge of the (conjunction of the) axioms of Q .

If the above is right, then the consequent of Theorem 1, namely,

$$\left((KKQ \wedge (K(\neg KQ \vee \neg \Box_Q \phi) \Box \rightarrow KKQ)) \wedge (K(\neg KQ \vee \neg \Box_Q \neg \phi) \Box \rightarrow KKQ) \right) \rightarrow (K\neg \Box_Q \phi \vee K\neg \Box_Q \neg \phi).$$

is false for $\phi = (18)$. So, at least one of the antecedents of Theorem 1 is false. The options are to deny (3), (4) or (5), their closure under RM_K , E or D_J . Giving up on (3) would be to give up on a conceptual truth about the notion of being in a position to know, if anything is. Similarly, D_J is absolutely essential for a notion of all-things-considered justification. Let us now discuss the other options.

Plausibly, (4) and RM_K stand or fall together, since the first is nothing but a representation in the object language of rule RM_K (restricted to logical derivability from the axioms of Q). It is clear that (4) and RM_K are idealisations. Rosenkranz is of course aware of this. He makes the following remarks about them:

For present purposes, and with hindsight, it may suffice to point out that we are here ultimately interested in an account of propositional justification in terms of K . As long as we do not assume that whenever a given proposition is justified in one's situation, one can easily avail oneself of its justification, irrespective of any further assumptions about one's epistemic capacities, the idealisations inherent in the characterisation of K need not automatically render the account of propositional justification irrelevant for the case of less than ideal subjects. (Rosenkranz, 2018, p. 11)

Whatever the merits of the idealisations, RM_K is necessary to prove a lot of the theorems in (Rosenkranz, 2018), including theorems that are seen as principles that have independent plausibility. Here are a few examples from (Rosenkranz, 2018, pp. 15–16):

$$\begin{aligned} J(\phi \wedge \psi) &\rightarrow J\phi \\ J\phi &\rightarrow J(\phi \vee \psi) \\ \neg J(J\phi \wedge \neg \phi) \\ \neg J(\phi \wedge \neg J\phi) \end{aligned}$$

So, giving up on RM_K would completely cripple the logic of justification. It would amount to saving something by destroying it. Just as giving up on RM_K would break the back of the logic of justification, denying E would cut the heart out of that logic, since it is the only bridge principle for the notions of being in a position to know and justification.

That leaves (5). As has been argued in Section 3, (5) fits within a nice ordering of agglomeration principles. Furthermore, it is immune against the counterexamples given in (Heylen, 2016) and (Rosenkranz, 2016). Moreover, it is risky to deny

(5), since it brings one close to a denial of (6). The latter seems to be hard to deny, if one allows a modicum of idealisation: one has to be physically and psychologically capable of knowing the conjunction of what one already knows and nothing has to stand in the way of one successfully exercising the capability to combine two items of knowledge. It will be tough to reject that idealisation but retain the idealisations required for RM_K . In sum, there does not seem to be an attractive option here for Rosenkranz.

Rosenkranz's logic of justification faces a serious challenge. But maybe there is a way out. Rosenkranz (2018, p. 320) writes the following:

We have already idealised away physical and psychological deficiencies, conceptual limitations and limitations on reasoning power. So [E] is not already refuted by trees, dogs, small infants, or madmen.

Although Alice has not been trained in Peano Arithmetic and model theory, she does not belong to the above list (trees, dogs, small infants, madmen). For one, she does have the concepts of knowledge and being in a position to know. She has even learned how to give logical arguments and axiomatic proofs. Furthermore, she knows the axioms of minimal arithmetic, so she has at least minimal semantic competence with respect to the vocabulary of (18). So I submit that the case of Alice does not belong in the company of trees, dogs, small infants, or madmen. But suppose that this train of thought is carried further. One might say that knowing and understanding mathematical induction is fundamental for full competence with respect to the language of arithmetic. In that case I would respond, first, by rehearsing the arguments in Section 4 but replacing minimal arithmetic with the theory of second-order arithmetic, which contains the axioms of minimal arithmetic together with the second-order *axiom* of induction (Boolos et al., 2007, ch. 22). Then I would go over the counterexample presented in this section again, but now I would replace minimal arithmetic with second-order arithmetic and replace (18) with an arithmetical sentence that is undecidable even in second-order arithmetic (e.g. the Gödel sentence for that theory). In this case the counterexample still stands. In fact, it is even stronger.

6 Summary

In Section 3, I have made three quite weak further assumptions. First, I have assumed that, if one knows something, then one is in a position to know it, cf. (3). This is a conceptual truth. Second, I have assumed that, if one is in a position to know the (conjunction of) the axioms of minimal arithmetic, Q , and if ϕ is logically provable from the (conjunction of) those axioms, then one is in a position to know that ϕ , cf. (4). This is nothing more than a representation in the object language of rule RM_K (applied to the logical consequences of the axioms of Q), formulated in the metalanguage. Third, I have assumed that, if

- one knows that ϕ and

- one is in a position to know that ψ and
- one were to know that ψ , then one would (still) know that ϕ ,

then one is in a position to know that ϕ and ψ , cf. (5). In Section 3 it was argued, first, that it fits within a nice ordering of agglomeration principles, second, that the principle is not vulnerable to counterexamples given to other agglomeration principles, and, third, that denying it comes uncomfortably close to rejecting the weakest agglomeration principle, which threatens to undercut the idealisations behind the closure principles in Rosenkranz's logic of justification.

Based on these assumptions and Rosenkranz's logic of justification I have in Section 4 proved a lemma (Lemma 1). The latter says that if

- one knows that one knows Q and
- one is in a position to know that one is not in a position to know ϕ ($\neg\phi$) and
- if one were to know that one is not in a position to know Q or that ϕ ($\neg\phi$) is not logically provable from Q , then one would still know that one knows the axioms of Q ,

then one is in a position to know that ϕ ($\neg\phi$) is not logically provable from Q . The lemma used the three assumptions mentioned above and, in addition, the principle that says that being in a position to know is closed under provable material implication (RM_K).

Going further, I have then in Section 5 proved a theorem (Theorem 1). The consequent of the latter says that, if someone has positive introspective and modally robust knowledge of the axioms of Q , then that person is in a position to know that a sentence is not provable in Q or that the negation of that sentence is not provable in Q . This theorem depends on the assumptions of the lemma, the principle that says that there is propositional, non-factive, all-things-considered justification for a sentence if and only if one is not in position to know that one is not in a position to know that sentence (E), and the principle that says that one cannot have all-things-considered justification for both a sentence and its negation (D_J).

Next, we have considered the case of Alice, who is logically competent and who knows that she knows that Q , but who lacks knowledge of non-standard models of arithmetic and who lacks knowledge of mathematical induction. The case involved a sentence that is undecidable in minimal arithmetic. Alice does not have the means to realize that the sentence is undecidable in minimal arithmetic. Yet, she knows that she knows the (conjunction of the) axioms of Q and this knowledge is modally robust in the sense that it is retained, even if some other relevant knowledge were to be acquired. (At the end of Section 5 an even stronger case, which involves an undecidable sentence in second-order-arithmetic, was briefly outlined.)

If it is accepted that a person can fail to be in a position to know that a sentence is unprovable in arithmetic and fail to be in a position to know that the negation of that sentence is unprovable in arithmetic, then one should deny (3), (4), (5), RM_K , E or D_J . This basically amounts to denying conceptual truths, rejecting a closure principle that would cripple the logic, rejecting an object-language formulation and application of that closure principle, rejecting another closure principle that threatens to undercut the idealisations behind the first closure principle, or rejecting the only bridge principle for the notions of being in a position to know and propositional, non-factive, all-things-considered justification. This is the challenge that Rosenkranz faces.

Acknowledgements Earlier versions of this paper have been presented at the Third Leuven-Bristol Workshop (Leuven, 15 September 2017) and the First Flemish Epistemology Workshop (Leuven, 18–19 May 2018). I would like to thank the audiences at those events for their comments and questions. I would like to thank Sven Rosenkranz, who was present at the First Flemish Epistemology Workshop, in particular. In addition, I would also like to thank Felipe Morales Carbonell, Harmen Ghijsen and Lars Arthur Tump for their feedback. Last but not least, I would like to thank the editor and two anonymous reviewers.

References

- Artemov, S. (2008). The logic of justification. *Review of Symbolic Logic*, 1(4):477–513.
- Boolos, G., Burgess, J., P., R., and Jeffrey, C. (2007). *Computability and Logic*. Cambridge University Press.
- Cohen, S. (1999). Contextualism, skepticism, and the structure of reasons. *Philosophical Perspectives*, 13:57–89.
- Fagin, R., Halpern, J. Y., Moses, Y., and Y., V. M. (1995). *Reasoning About Knowledge*. MIT press.
- Hendricks, V. F. (2003). Active agents. *Journal of Logic, Language and Information*, 12(4):469–495.
- Hendricks, V. F. (2005). *Mainstream and Formal Epistemology*. Cambridge University, New York.
- Heylen, J. (2016). Being in a position to know and closure. *Thought: A Journal of Philosophy*, 5(1):63–67.
- Hintikka, J. (1962). *Knowledge and Belief*. Cornell University Press, Ithaca.
- Meyer, J. C. and van der Hoek, W. (1995). *Epistemic Logic for AI and Computer Science*. Cambridge University Press, New York.

- Rosenkranz, S. (2007). Agnosticism as a third stance. *Mind*, 116(461):55–104.
- Rosenkranz, S. (2016). Being in a position to know and closure: Reply to Heylen. *Thought: A Journal of Philosophy*, 5(1):68–72.
- Rosenkranz, S. (2018). The structure of justification. *Mind*, 127(506):309–338.
- Smith, M. (2018). The logic of epistemic justification. *Synthese*, 195(9): 3857–3875.
- von Wright, G. H. (1951). *An Essay in Modal Logic*. North-Holland Pub. Co., Amsterdam.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.