

# IF YOU DON'T KNOW THAT YOU KNOW, YOU COULD BE SURPRISED

Forthcoming in *NOÛS*

---

Eli Pitcovski

*Tel-Hai Academic College*

Levi Spectre

*The Open University of Israel*

**Abstract:** Before the semester begins, a teacher tells his students: “There will be exactly one exam this semester. It will not take place on a day that is an immediate-successor of a day that you are currently in a position to know is not the exam-day”. Both the students and the teacher know – it is common knowledge – that no exam can be given on the first day of the semester. Since the teacher is truthful and reliable, it seems that the students can know that what he says is true. However, in that case, assuming the students can know that they know whatever it is they know (KK) and assuming their knowledge is closed under entailment (closure), the students can reason from what they know to the conclusion that no exam will take place during the semester. This conclusion contradicts what they supposedly know: that there will be an exam. This puzzle, we argue, gives rise to a new consideration for the rejection of KK. We discuss unique features of the argument, especially in comparison to Timothy Williamson’s rejection of KK in light of other versions of the surprise exam paradox.

## 1 Introduction

The surprise exam paradox can be introduced (informally) as follows:

**SURPRISE:** A teacher announces “On one of the days next week (Monday to Friday) there will be a surprise exam: you will not know on the morning of the exam day that the exam will take place on that day”. A bright student reasons as follows: “the surprise exam cannot take place on Friday, because if it has not taken place by Friday morning, we will know on Friday morning that it will take place that day. But then it cannot take place on Thursday, because given that it cannot take place on Friday, if it has not taken place by Thursday morning we will know then that it will take place that day...”. She continues eliminating backwards, until she reaches the conclusion that the surprise exam cannot take place at all. The following Wednesday, to the student’s surprise, the teacher gives the exam. As the teacher first announced, the students did not know on Wednesday morning that the exam would be given that day.

---

According to standard solutions, the bright student's tenuous assumption is that knowledge is stable, in the sense that attained knowledge is accessible from any epistemic state in which the basis for it is still accessible (for example: the student knows that there will be a surprise exam as long as the teacher's announcement is not forgotten).<sup>1</sup> In other words, the student falsely assumes that, following the teacher's announcement, even if the test will not take place between Monday and Friday morning, she will continue to know on Friday morning that

TAS (Teacher Announcement SURPRISE): There will be an exam on one of the weekdays and I will not know on the morning of the exam day that the exam will take place on that day.

But given information the student will possess on Friday morning, it is argued, TAS will be impossible for her to know. This is because (together with the hypothetical knowledge that the exam has not taken place so far) on Friday morning TAS implies: "there will be an exam today, and I do not know that there will be an exam today." Due to their structure, sentences of this form ('p and I do not know that p') are unknowable:<sup>2</sup> One's knowledge of the first conjunct, p, contradicts the truth of the second conjunct (that one doesn't know that p). The student's reasoning is undermined by the false assumption that on Friday morning she will be able to deduce from known premises (i.e. TAS) that there will be an exam that day, and hence come to know this conclusion.

In this paper, we introduce a new version of the paradox, designed to resist the standard solutions. In light of this, the paradox we present requires an unorthodox solution focused on other aspects of the student's reasoning. We assess the student's reasoning, step by step. After considering other possible responses, we argue for the rejection of KK, the principle that if S knows that p then (it is a priori implied that) S can know that S knows that p. Importantly, we do not reject the standard solution to SURPRISE, and we do not suggest that SURPRISE calls for the rejection of KK. But we do think that there is a general lesson following both from the standard solution to SURPRISE and from our solution to the new paradox. In outline: both solutions expose a crucial difference between relying on true premises and relying on known premises.

---

<sup>1</sup> The idea was already present in Saul Kripke's recorded lecture at the Cambridge University to the Moral Sciences Club in 1972 (and perhaps earlier elsewhere). An edited version of this lecture can be found in his *Philosophical Troubles* (2011). Following Kripke, many locate the fallacy in the student's assumption that she will retain her knowledge that the teacher spoke truly throughout the week (see Wright and Sudbury (1977) for instance). However, Sorensen (1982) has presented a version of the paradox that removes major temporal aspects of the original paradox, and to which accounts relying on a notion of temporal retention fail to apply. The more general, currently prevalent solution, applying both to Sorensen's paradox and to the original paradox, relies on showing that the student falsely attributes knowledge to a subject in some states from which this knowledge is inaccessible. Sorensen (1988) recognizes contingent blindspots: propositions that are knowable for some agents but are not knowable for other (suitably informed) agents, in different epistemic states. Sorensen (1988, ch.8) relies on this notion to reject the first step of the student's argument. Kripke doesn't consider such cases, but his solution is structurally similar. Whether the rejection of the student's argument comes at the first or the second stage depends on further assumptions that do not change the basic features of the prominent solution to the paradox.

<sup>2</sup> This is a familiar Moore paradoxical sentence. Assuming (single premise) closure – in particular that knowledge distributes over conjunctions – knowing the Mooreian conjunction would mean knowing that p and knowing that one doesn't know that p:  $K(p \wedge \neg Kp) \models Kp \wedge K(\neg Kp)$ . Since knowledge is factive, a contradiction of the form " $Kp \wedge \neg Kp$ " follows.

---

After outlining our new version of the paradox and explaining how it differs from SURPRISE (2), we show that it generates an argument against KK (3). We then discuss differences and similarities between this argument and former attempts to argue against KK via versions of the surprise exam paradox, in particular Williamson (2000, ch.6) (4). In the final section, we briefly discuss ways in which our main argument complements the insights gained by the standard solution to SURPRISE with respect to reasoning with knowledge-to-fact conditionals (of the form ‘If S knows that p then q’) (5).

## 2 Backwards Backwards-Induction

A new version of the surprise exam paradox can be introduced (informally) as follows:

FORWARD: A semester in some faraway land lasts  $N$  days straight, no breaks, no weekends.<sup>3</sup> Before the beginning of the semester, a teacher announces “There will be exactly one exam this coming semester. It will not take place on a day that is an immediate successor of a day which you know, prior to the beginning of the semester, is not the exam-day.” It is well known by all parties that due to school policy, an exam cannot be given on the first day of the semester. A bright student reasons as follows: “There will not be an exam on the first day of the semester, and I currently know this (school policy). So, following the teacher’s announcement, there will not be an exam on the second day, and I currently know this. So, following the teacher’s announcement, there will not be an exam on the third day...” (She goes on iterating this reasoning, and concludes that there will not be an exam during the semester, as none would meet the conditions announced by the teacher.) However, since the day of the exam,  $i$ , was randomly and secretly chosen, and since the teacher knows that  $i$  is sufficiently high (so that the students cannot possibly know that  $i-1$  is not the exam day), when the teacher actually gives the exam on day  $i$ , his announcement is true. The students do not know, prior to the beginning of the semester, that day  $i-1$  is not the exam-day.<sup>4</sup>

In order for the puzzle to have any force, it needs to be clear that the teacher can in fact carry out the announcement, i.e. make it the case that first, there is exactly one exam in the semester and second, that

TAF (Teacher-Announcement FORWARD): for any day  $n$  throughout the semester, if the students know (before the beginning of the semester) that the exam is not on day  $n$ , then the exam will not take place on the following day ( $n+1$ ).

Moreover, it needs to be clear that the students can know that TAF is true, and that their potential knowledge would not make the teacher’s plan to give the exam (in line with TAF) impossible to carry out. It is easy to imagine scenarios on which TAF would not be known to be true despite the teacher’s announcement. For instance, if the teacher were to announce TAF and the students were

---

<sup>3</sup> The size of  $N$  is discussed below.

<sup>4</sup> A version of an infinite backward induction paradox has been presented by Sorensen (1999). As opposed to our version, Sorensen’s version depends on common knowledge – similar to the so called “centipede” – that the exam will be given on the first day on which it will surprise the students. Sorensen offers a wholly different analysis from the one we provide here.

---

to know (independently) that an exam can only be given on the 7<sup>th</sup> day of the semester, TAF (being false) would be impossible to know. Yet, we argue, FORWARD is not such a case.

Suppose that the teacher is committed to giving the exam on a day determined by a lottery he secretly draws between possible exam days (day 2 through day N). Suppose that the results of the lottery determine that the exam should take place on day 60 (that is,  $i=60$ ). It seems uncontroversial to say that in this situation, prior to making the announcement and on the basis of his knowledge of the actual result of the lottery, the teacher knows that the students do not know that the exam will not take place on day (i1) i.e. on day 59. At this stage, then, the teacher knows that the exam will not take place on a day that immediately follows a day the students know is not the exam day. The first question is: can the teacher convey this knowledge to the students by announcing TAF?

Since we can stipulate that the testimonial exchange is otherwise unproblematic (the teacher is reliable, he has no intention to deceive, he would not announce TAF unless the designated exam day was far from the beginning of the semester,<sup>5</sup> etc.), the only reason to worry that TAF cannot be known to the students upon hearing the teacher's announcement is the suspicion that TAF cannot be known together with other things that the students know: 1. that there is exactly one exam and 2. that it will not take place on day 1. We will come back to this worry in section 4. For now, appealing to an analogous case (based on Kripke (2011, p.29))<sup>6</sup> we can demonstrate that knowledge of the sort in question together with the same kind of background knowledge is unproblematic, and can even be attained without testimony. Suppose I stick the queen of hearts somewhere near the middle of a deck of cards (knowing that before doing so, the deck was missing the queen of hearts). I will flip the cards one by one, starting from the top, until I find the queen of hearts. Before flipping any of the cards I seem to know: 1. that there is exactly one queen of hearts in the deck 2. that the queen of hearts is not the top card 3. that before flipping the cards, there is no way for me to know of the card right on top of the queen of hearts that it is not the queen of hearts. In other words, I know that

CARDS: with respect to any card  $n$ , if I know by any means available to me now that  $n$  is not the queen of hearts, then the card immediately underneath it,  $n+1$ , is not the queen of hearts.

I know, for instance, that if I know that the top card is not the queen of hearts, then the second card from the top is not the queen of hearts.

Given the structural similarity between CARDS and TAF (and the similarity of the background knowledge in both cases), if CARDS is knowable in these circumstances – and it would be a hard sell to claim it is not – we see no reason to worry that the students cannot know that TAF is true. Later we will show that, if anything, the student's knowledge of TAF is easier to acknowledge than

---

<sup>5</sup> Below we show the teacher can make his announcement truthfully even under weighty assumptions about the degree of higher order knowledge the students have with respect to the school policy and even with a fairly early designated exam day (exactly how early is discussed below).

<sup>6</sup> Kripke's original case demonstrates that the students can know the teacher's announcement in cases like SURPRISE, at least before the week begins. We have adjusted the case to match our purposes.

---

my knowledge of CARDS.<sup>7</sup> If knowledge of TAF is possible in the students' state, then given the teacher's truthfulness and reliability, it seems that the teacher has no reason to worry that upon hearing his testimony the students will still not know TAF. Since it is impossible for them to know that day 59 is not the exam day, the teacher can safely give the exam on day 60 and by doing so, fulfill his announcement.

If all of this is true, the student's reasoning is problematic: it leads her to the false conclusion that there will be no exam. Before analyzing the student's reasoning in search of the source of the problem, let us provide a clearer view of two unique characteristics of our puzzle.

1. Most importantly, in FORWARD, the student does not in any way ascribe to herself knowledge of a proposition that is essentially, due to its structure, impossible to know (in the relevant epistemic state). This is so with respect to any of the epistemic states she considers. More specifically, she does not consider epistemic states in which the teacher's announcement implies anything of the form 'p and I do not know that p'. In fact, as opposed to SURPRISE, the student is only considering her own actual epistemic state: no relevant contingent information is added to that state after the teacher's announcement. There is nothing to affect the knowability of TAF in her epistemic state (prior to the beginning of the semester). Some of the more familiar solutions to SURPRISE, then, are ruled out as solution to FORWARD.<sup>8</sup>
2. Intuitively, in cases like SURPRISE, the larger the number of possible exam days, the easier it is to accept the teacher's announcement, and the less information it contains. Somewhat similarly, in FORWARD: the greater  $i$  – the number of the actual exam day – is, the easier it is, from the theorist's vantage point, to accept the teacher's announcement. (Likewise, the greater  $N$  – the number of possible exam days – is, the easier it ought to be for the students to believe the announcement, and the greater the value of  $i$ , the easier it is for this belief to count as knowledge.) Following Kripke, we consider solutions that account for this intuition to be preferable (this will become important).<sup>9</sup> Note that in contrast with SURPRISE, the present puzzle does not depend on there being a finite number of days in

---

<sup>7</sup> Kripke's case is also dissimilar to FORWARD in important respects. But all that is important for us are two things. First, that it is a case in which despite the background knowledge being relevantly similar to what the student knows in FORWARD, something in the form of TAF, viz. CARDS, can be known (or at the very least that an argument is required to show that it cannot. We will come back to this in section 4). Second, that in this case, an analogous form of reasoning (concluding that there is no queen of hearts in the deck, by utilizing my knowledge of CARDS) seems trivially fallacious. The dissimilarities are, if anything, in favor of FORWARD.

<sup>8</sup> Aside from the standard solution considered above, this includes Sorensen's (1988) account. According to Sorensen, knowing the truth of TAS implies knowledge of a disjunction of "blindspots" of the form "Either the exam will be on first day and I will not know this on the morning of the first day, or... or it will be on the last day and I will not know this on the morning of the last day". With every day that passes, the students discard the relevant disjunct and hence by the last day, they must lose their knowledge. Since TAF does not imply a disjunction of blindspots, nor does it leave room for contingent information that supports discarding particular elements, Sorensen's solution is unavailable.

<sup>9</sup> See Kripke (2011, ch.2). Kim and Vasudevan (2017) have nicely accounted for this intuition with respect to SURPRISE. As will be clear below, our view accounts for this intuition, as it shows that every iteration of the student's argument requires additional knowledge.

the semester. So there is strictly no limit to how compelling we can make the teacher's announcement by raising the value of  $i$  in the example. Conversely, there is literally no end to how implausible the student's reasoning can be here.

Based on her belief in TAF and the school policy, the student reasons forward and mistakenly eliminates possible exam days. Her reasoning takes her from true assumptions to false conclusions. To investigate why this takes place, we should take a closer look at the student's reasoning. Given the last point (2), to somewhat level the playing field, we will focus on a limited window of opportunities for the teacher to give the exam. In the version we consider in the next section,  $N=6$ . That is, there are six known consecutive semester days altogether. Our goal will be to find out how far the student can get with proper reasoning and what assumptions are needed to ground the reasoning she carries out.

### 3 KK And FORWARD

Before tracking the problematic element of the student's reasoning, let us unfold the student's reasoning process from a theorist's point of view. We can begin by clearly presenting the relevant facts and what the student knows. We will then be able to observe what follows from the student's knowledge and the facts, with the aid of which principles. At times we will consider the teacher's strategy for fulfilling his announcement, suppressing the strategy of determining the exam day by chance. We adopt Kripke's (2011) notation with some (minor) changes:

$e_n$  = The exam is given on day  $n$  (where  $n \in \{1, 2, 3, 4, 5, 6\}$ )

$K(A)$  = The students know that  $A$  (on all and any relevant day)

$K_m(A)$  = The students know on day  $m$  that  $A$  (where  $m \in \{0, 1, 2, 3, 4, 5, 6\}$  and day 0 is a time after the teacher's announcement but before the beginning of the semester).

$K^n(A)$  = The students know that they know...(n times altogether) that  $A$ .

For instance:

$K^2(A)$  = The students know that they know that  $A$ .

We can make the following assumptions. Firstly, there is exactly one exam. This is to prevent the teacher from giving an exam that does, and another (earlier) exam that doesn't, comply with his announcement

1. For exactly one  $n \in \{1, 2, 3, 4, 5, 6\}$ ,  $e_n$

This assumption is known as background knowledge, so we will assume that at least after the announcement:

- 1\*. The students know that for exactly one  $n \in \{1, 2, 3, 4, 5, 6\}$ ,  $e_n$

We also have the school's strict rule, and the student's knowledge of it:

2.  $\neg e_1$
- 2\*.  $K_0(\neg e_1)$

Next, we will assume that the teacher will act in accordance with his announcement and that

from the announcement, the students know that if they know (prior to the beginning of the semester) that no exam will take place on some day  $n$ , then no exam will take place on the following day,  $n+1$ .<sup>10</sup>

$$3. \quad \text{TAF: } K_0(\neg e_n) \supset \neg e_{n+1}$$

$$3^*. \quad K_0(K_0(\neg e_n) \supset \neg e_{n+1})$$

Let us see how far the bright student can get with what we have so far.

From an instantiation of 3 (TAF) we have:

$$4. \quad K_0(\neg e_1) \supset \neg e_2$$

From 2\* (knowledge of school policy) and 4, we infer (by modus ponens):

$$5. \quad \neg e_2$$

The teacher, then, will not deliver the exam on the second day of the semester. This follows from TAF and the student's knowledge that the exam will not be given on day 1. But these two assumptions do not entail that the student knows that the teacher will not deliver the exam on the second day of the semester.<sup>11</sup> We need a general claim about knowledge to move forward with what we have. We will assume closure under multi-premise deductive inferences (and that the student comes to believe all the relevant propositions that she properly deduces from propositions she in fact knows).<sup>12,13</sup> (Closure will be written as a schematic assumption, without necessity operators or quantification. The "⊨" sign refers to a priori entailment widely conceived)

$$6. \quad \text{CLOSURE: } K(A \supset B) \models K(A) \supset K(B)$$

CLOSURE and what we have so far still does not entail that the student knows that the teacher will not deliver the exam on the second day of the semester. Assuming CLOSURE and the following instance of 3\*:

---

<sup>10</sup> Another way to present this assumption is to turn 3\* to a conjunction of its instances:  $(K_0(K_0(\neg e_1) \supset \neg e_2)) \wedge (K_0(K_0(\neg e_2) \supset \neg e_3)) \wedge (K_0(K_0(\neg e_3) \supset \neg e_4)) \wedge (K_0(K_0(\neg e_4) \supset \neg e_5)) \wedge (K_0(K_0(\neg e_5) \supset \neg e_6))$ .

<sup>11</sup> We have seen that 5 follows from the premises, i.e. it is true. However, for the student to know that it is true it isn't enough that it follows from the premises. It must be somehow placed within the closure schema (6, below). One might wonder, "Can 5 not be learned from this paper if it shows that 5 follows from the premises?" The answer is that we can learn that 5 follows from the premises and if we know the premises are true, we know – assuming closure (but not necessarily assuming that we know closure) – that 5 is true.

<sup>12</sup> This is the strongest form of epistemic closure. Strong as it is, it is shared by all normal modal epistemic logics. It is equivalent to the idea that one can know the strongest proposition that follows from everything that is known by the subject, i.e., the conjunction of all the individual propositions she knows. (To see why  $K(A \supset B) \models K(A) \supset K(B)$  is equivalent to conjunction introduction, i.e.,  $K(A), K(B) \models K(A \wedge B)$ , consider the tautology  $A \supset (B \supset (A \wedge B))$ . By the epistemic necessitation principle (the second axiom of normal modal logics), we have,  $\models K(A \supset (B \supset (A \wedge B)))$ . Since the following are two instances of the closure:  $K(A \supset (B \supset (A \wedge B))) \models K(A) \supset K(B \supset (A \wedge B))$  and  $K(B \supset (A \wedge B)) \models K(B) \supset K(A \wedge B)$ , closure entails  $K(A), K(B) \models K(A \wedge B)$ ). See Pagin (1990) for a relevant critique of Sorensen's (1988). See Sharon and Spectre (2013) for an employment of this argument with regard to multi-premise knowledge closure. We will later discuss whether there is any point in adopting a weaker closure principle, and argue that, in this context, there isn't.

<sup>13</sup> We are assuming that CLOSURE cannot break down due to the student's failure to carry out some deductively valid inference. In any event, no adequate solution to our puzzle could depend on such a failure because it would still leave us with the question, "What if the students had made that valid inference?"

---


$$7. K_0(K_0(\neg e_1) \supset \neg e_2)$$

we can merely deduce:

$$8. K_0(K_0(\neg e_1)) \supset K_0(\neg e_2) \quad (\text{that is: } K_0^2(\neg e_1) \supset K_0(\neg e_2))$$

Proceeding from 8 requires us to strengthen assumption 2\*. It needs to be assumed that the students have higher-order knowledge that the exam will not take place on day 1. Since this knowledge is independent (the school policy (2) is known to the student on independent grounds), at least with respect to day 1 we can assume higher-order knowledge of the form (say):<sup>14</sup>

$$9. K_0^7(\neg e_1)$$

(A degree higher than  $K_0^7$  will obviously not be required.) This means that the student knows that she knows that she knows ... (seven times altogether) that the first day is not the exam day. From 8 and 9, it finally follows that the student knows that the exam will not take place on the second day:

$$10. K_0(\neg e_2)$$

The next step depends on the following conditional (another instantiation of 3):

$$11. K_0(\neg e_2) \supset \neg e_3$$

Now, from the theorists' point of view, we can infer (from 10 and 11):

$$12. \neg e_3$$

It is interesting to see that we can get this far. We might not be surprised that the teacher's announcement cannot be true if he delivers a second-day exam. But it is indeed surprising that unless the teacher knows that the students *do not* have higher-order knowledge of 2 (the school policy), he cannot administer a third-day exam (without rendering the announcement unknowable to the students).<sup>15</sup> So we already have an explanation for why a three-day semester renders the teacher's announcement problematic, assuming that he wants the students to know that his announcement is true.

But what now? Can the student expand her knowledge further? Can we possibly arrive at  $K_0(\neg e_3)$  with what we have so far? If we can, we can repeat the 10-11 maneuver (and discharge  $(K_0\neg(e_3) \supset \neg e_4)$ , another instantiation of 3). CLOSURE and the following instance of 3\*:

$$13. K_0(K_0(\neg e_2) \supset \neg e_3)$$

jointly entail:

$$14. K_0^2\neg(e_2) \supset K_0(\neg e_3)$$

But here we seem to be stuck. For proceeding requires the student to possess higher-order

---

<sup>14</sup> It is important to note that the students do not know that they have exactly  $K_7(A)$  levels of higher-order knowledge. That would mean that they have at least  $K_8(A)$  levels of higher-order knowledge.

<sup>15</sup> Note that the students don't actually need to know that they know that the first day will not be the exam day. This is because not knowing whether or not the students know that they know, or merely know, the teacher will be unable to determine whether he can truthfully execute a third day exam that will comply with his announcement. Assuming he needs to know that the students don't know that day two is not an exam day in order to know that his announcement will be unknowable to the students as well, a third day exam is at the very least entangled with serious complications.



knowledge about the exam not taking place on day 2 (i.e. it requires the truth of the antecedent of 14).<sup>16</sup>

More generally: noticing the pattern, in order for the student's reasoning to continue to unfold with the aid of instances of 3\* and CLOSURE, it will have to be assumed that the student not only knows each of the consequences she arrives at, but that she knows that she knows it. More specifically: to assume that given the student's knowledge of TAF (i.e. 3\*), arriving at knowledge with respect to some particular day  $K_0(\neg e_n)$  is enough to ground knowledge about the next day  $K_0(\neg e_{n+1})$  is to tacitly assume the KK-principle:<sup>17,18</sup>

$$15. \text{ KK: } K(A) \models K^2(A)$$

If KK is assumed, we could employ it to derive the desired higher-order knowledge directly from 10, which together with KK entails:

$$16. K_0^2 \neg(e_2)$$

If so, the student can clearly rule out a third day exam, i.e. we can derive the following from 14 and 16:

$$17. K_0 \neg(e_3)$$

The rest of the student's argument can be reconstructed with just a repetition of this pattern: use KK to turn first order knowledge  $K_0 \neg(e_n)$ , to second order knowledge  $K_0^2 \neg(e_n)$ . Then use 3\* and CLOSURE to derive:  $K_0^2 \neg(e_n) \supset K_0(\neg e_{n+1})$ . From here, derive (by modus ponens) first order knowledge about the next day  $K_0(\neg e_{n+1})$ . Following this process through, it can be shown that with respect to any day  $n \in \{1, 2, 3, 4, 5, 6\}$ , the student knows that it is not the exam day. Given the factivity of knowledge

$$18. \text{ FACTIVITY: } K(A) \models (A)$$

it follows that an exam will not take place on any day  $n \in \{1, 2, 3, 4, 5, 6\}$ , which contradicts 1.

At this point we have a choice. Without giving up the idea that the students can know TAF,<sup>19</sup> we have to choose between CLOSURE and KK. Before deciding one way or another, it is worth noting the following considerations.

First: It should be noticed that no fixed degree of higher-order knowledge of 2 and 3 (the school policy and TAF accordingly) can play the role of KK. To demonstrate: notice, first, that an arbitrarily

---

<sup>16</sup> We will shortly discuss the insufficiency of (what we have so far and) an arbitrarily high degree of knowledge with respect to day 1 to deliver higher-order knowledge with respect to day 2.

<sup>17</sup> We could — though we are not doing this here — restrict KK in familiar ways to make it more plausible. But there is no restriction that would make a difference with regard to the argument. At each stage, the student needs to appeal to the principle to gain knowledge that they do not have given the other assumptions alone.

<sup>18</sup> One might think that the antecedent of 14 is derivable from CLOSURE, 4 and 9, without KK. This amounts to the claim that 14 follows from these premises in normal modal (reflexive) logics weaker than S4, i.e., without the transitivity of the epistemic accessibility relation. But this is not the case. To see this think about whether or not there is a world where  $K^2(p)$ ,  $K(p \supset q)$ ,  $K(q)$ ,  $\neg K^2(q)$  are all true.  $K(p \supset q)$  only entails  $K(p) \supset K(q)$ , it does not entail  $K^2(p) \supset K^2(q)$ .

<sup>19</sup> We will reconsider this possibility in the next section (over and above our discussion in section 2).

high degree of higher-order knowledge of 2 is not all by itself sufficient. As presented above, the student does have some higher-order knowledge to spare ( $9$  reads —  $K_0^7(\neg e_1)$ ). However, this higher-order knowledge can only deliver  $K_0^2(\neg e_2)$  (which is essential for deriving  $K_0(\neg e_3)$ ) with the aid of something like:

$$19. K_0^3(\neg e_1) \supset K_0^2(\neg e_2)$$

The following might seem an initially promising mediator:

$$20. K_0^2(K_0(\neg e_1) \supset \neg e_2)$$

For 19 follows from 20 via CLOSURE. But for 20 to be true we don't merely have to assume that the student knows TAF. We must assume that the student also knows that she knows it. So higher-order knowledge of 2 is insufficient. More importantly: even if we do assume 20, this only allows the student's argument to go one step further. For while it can be employed to derive  $K_0(\neg e_3)$  (so if the students have higher-order knowledge of the school policy and TAF they can safely rule out a third day exam), it can definitely not be employed to derive  $K_0^2(\neg e_3)$ , which is necessary for the student to rule out a fourth day exam,  $K_0(\neg e_4)$ . Similarly,  $K_0^3(K_0(\neg e_1) \supset \neg e_2)$  (or for that matter:  $K_0^2(K_0(\neg e_2) \supset \neg e_3)$ ) also quickly exhausts itself after no more than two steps.<sup>20</sup> So rejecting KK is enough to rebut the student's reasoning.

Second: As noted earlier (section 2), the iterations in the student's argument are intuitively less and less plausible at each stage. Rejecting KK nicely explains this intuition. As we have just seen, if KK is false, the number of stages in which the student's argument can legitimately go on depends on the degree of higher-order knowledge she has with respect to the school policy (2) *and* with respect to TAF (3). Hence every new stage of her reasoning requires relevant knowledge of higher degrees. Rejecting CLOSURE, however, does not explain this intuition.<sup>21</sup> Since the same reasoning is applied by the student in every step, it is hard to see why it should become less and less plausible at different stages.<sup>22</sup>

Third: Having a look at the informal presentation of FORWARD, the most suspicious link of every

<sup>20</sup> Generally: it is true that CLOSURE and higher-order knowledge of TAF jointly entail the general schema  $K_0^n(\neg e_i) \supset K_0^{n-1}(\neg e_{i+1})$ , but its application too will run out in accordance with how many iterations of knowledge the students have with regard to the school policy (2). It is hard to see how allowing for the students to have arbitrarily high orders of knowledge of the school policy can be motivated without prior acceptance of KK. Besides, even allowing this much will translate into a significant challenge to the above argument – i.e. allow the students to rule out all the exam days – only if at least one of the following two claims are true (or something very much like them): (a) The number of possible exam days  $N$  is equal to (or smaller than) the number  $n$  of higher-order knowledge iterations. (b) Every added possible exam-day allows the students to have a further iteration of knowledge. But (a) leaves room for versions of the paradox (with a number of semester days sufficiently greater than  $N$ ), and (b) is no better than assuming KK and plausibly worse. If there is any correlation between  $N$  and  $n$  – a correlation already means we are not assuming KK – it seems untenable to allow a further knowledge iteration for every added day.

<sup>21</sup> This can also count against attempts to defend KK by objecting to 3\*. Unless rejecting 3\* comes with an explanation for why the student's reasoning seems less plausible every step, rejecting KK seems to have an advantage.

<sup>22</sup> With respect to FORWARD, there is no reason to suspect that the decreasing degree of plausibility has to do with the student relying on more and more logical calculations, thereby increasing the risk for mistake as it were. The form of reasoning just repeats itself.

stage of the student's reasoning is the 'and I know this' clause (in: "so the exam cannot take place on day...*and I know this*"). This clause is not supported by what the student knows when this is declared. And as we have seen, the principle grounding what otherwise seems like an artificial "momentum boost" is KK.

At first glance, rejecting CLOSURE could also seem attractive. After all, multi-premise closure is an especially strong form of closure, incompatible with many fallibilist epistemologies.<sup>23</sup> However, this can only be pertinent if there is some good reason to think that in this specific case multi-premise closure expands knowledge beyond its actual reach. For instance, that the evidence is sufficient for yielding knowledge of the premises and insufficient to support knowledge of their conjunction. But here no such argument immediately suggests itself.

Another problem is that if we only reject multi-premise closure, knowledge of the following:<sup>24</sup>

$$21. \quad K_0((K_0 \neg e_n \supset \neg e_{n+1}) \wedge K_0^6(\neg e_1))$$

still gives rise to the paradox. If 21 — the conjunction of the knowledge gained from the teacher's announcement and 9 — is known to the student, the student's reasoning can be carried out assuming only single-premise closure (and KK). This is the weaker closure principle that assumes that whatever is derived from a known proposition is known. So not much is gained by exchanging multi-premise for single-premise closure.

Denying closure altogether is also possible, of course, but an argument needs to be given for why closure is violated in this particular context. And here too, we cannot rule out the possibility that such an argument will be given, but unless it is provided, the rejection of KK seems warranted.<sup>25</sup>

#### 4 FORWARD: Another Glimpse (?)

Recapitulating, FORWARD motivated the renunciation of the idea that given any ordered list of propositions  $\langle A_1, A_2 \dots A_N \rangle$ , knowledge of

$$1. \quad K(A_n) \supset A_{n+1}$$

and

$$2. \quad K^N(A_1)$$

is sufficient to generate knowledge of  $A_1$ - $A_N$ . Since only free iteration of knowledge can ground the steps from discharging one instance of 1 (viz.  $K(A_1) \supset A_2$ ) to discharging indefinitely many others, our response was to reject KK. Without KK, knowledge of 1-2, reinforced with CLOSURE, can only generate knowledge of  $A_1$  and  $A_2$  (and the truth of  $A_1$ - $A_3$ ).

---

<sup>23</sup> A notable exception here is Smith's (2010, 2016) normalcy justification account. However, a drawback of Smith's account is that it will accept cases where one is justified in believing a conjunction despite the fact that on the evidence it has a low probability. Smith (2016) proposes a version of the view that doesn't have this implication, but it is a view that he doesn't pursue and, in any case, that version does not support a multi-premise closure principle.

<sup>24</sup> Williamson (2000, 117-8) makes this point in the context of his argument against the KK principle.

<sup>25</sup> For the kind of contexts where closure denial is more promising see Spectre (2009) and Sharon and Spectre (2017).

---

In light of recent defenses of KK and recent objections to Williamson's (2000, ch.5) well known reductio argument against KK,<sup>26</sup> our favored response to FORWARD will not likely be accepted without a challenge. In this section we will compare our argument to a version of Williamson's reductio argument. As we argue: the similarities between the arguments will explain why advocates of KK need to give an independent reason for rejecting 3\* (our assumption that the students know that the teacher's announcement is true) in case they hope to save KK in this way (4.1). The differences between the arguments will show that FORWARD gives rise to an objection to KK that is stronger than that resulting from Williamson's reductio argument. Accordingly, familiar counterarguments leveled against Williamson's examples do not extend to FORWARD (4.2). Given the differences between FORWARD and Williamson's examples, toward the end of the section (4.3) we also offer a tentative explanation of why KK fails in cases like FORWARD even if it cannot be assumed to fail in the types of examples Williamson employs. The explanation is not meant to rule out all other possible explanations, but only to illuminate what might be wrong with KK, given examples like FORWARD.

#### 4.1 The Glimpse

Williamson's best-known argument against KK is based on his example of Mr. Magoo. However, Williamson applies the same sort of reasoning to another example, closer to the paradoxes we have been addressing here. Given the structural similarity between this example and Williamson's Mr. Magoo argument, the dialectics will be the same, and it will be easy to show that counterarguments to the former apply to the latter. So the following example by Williamson (2000, p.135) will be the center of our focus in what follows:

GLIMPSE: A student gets a glimpse of her teacher's calendar. Based on this glimpse, she knows that there will be exactly one exam during the semester and that it will not take place near the last day of the semester ( $N$ ). She also knows that her glimpse-based knowledge is inexact in the sense that she knows that for any number  $n$  ( $0 \leq n < N$ ), if she knows now (at the very beginning of the semester) that it will not take place  $n$  days away from the last day of the semester,  $N$ , then it will not take place  $n+1$  days away from  $N$ . She reasons as follows: "I know that the exam will not be on the last day of the semester, and I know that if it is not  $n$  days from the end of the semester, it is also not  $n+1$  days from the end of the semester. Hence, I can conclude that the exam is not on the penultimate day. Moreover, ..." She goes on reasoning, to rule out every day of the semester as a possible exam day. But the exam does take place during the semester.

According to Williamson, the student knows that in her situation:

G: For any number  $n$ , ( $0 \leq n < N$ ) if the student knows that the exam will not take place  $n$  days from the end of the semester, then the exam will not take place  $n+1$  days from the end of the semester i.e., the day before.

Assuming KK, CLOSURE and her knowledge that an exam will not take place on the last day,  $N$ , she

---

<sup>26</sup> Inter alia Sharon and Spectre (2008), McHugh (2008), Stalnaker (2009; 2015), Cohen and Comesaña (2013a), Cresto (2012), Greco (2014; 2015; 2017), Das and Salow (2018), Salow (2018; 2019) and Dorst (2019).

---

could discharge a list of conditionals so as to derive the false conclusion she arrives at. (Roughly as before: employ KK to get second order knowledge that the exam will not take place on day N. Use CLOSURE and G to get a conditional of the form ‘if the student knows that she knows that N is not the exam day, then she knows that the penultimate day, N1, is not the exam day’. By modus ponens, it follows that the student knows that N1 is not the exam day. Repeat this process to reconstruct the student’s reasoning process.)

In response, Williamson rejects KK. Based on their response to Williamson’s Mr. Magoo argument, KK advocates will reject G.

Before going into any details about why they will reject it, it is important to note that Williamson does not consider G to be a general epistemological principle,<sup>27</sup> but only to concern knowledge that the student has with regard to the particular predicament she is in. G is a proposition – a local generalization – that the student has independent evidence for, evidence concerning her sight and evaluative powers, and hence something she knows. This is important because as a general knowledge principle (even as a principle restricted to cases of inexact knowledge), it stands in direct opposition to KK.<sup>28</sup> Williamson wants his argument to apply to a wider audience than those who already agree that (inexact) knowledge is governed by a margin-for-error principle.

KK advocates ought to accept the following: while assuming G as a general principle about inexact knowledge begs the question against them, it’s up to them to show that the student cannot know G in the particular situation described in GLIMPSE. They need to explain why the student cannot know this proposition, not even as a local fact regarding her particular state and what she can know about the limits of her perceptual and evaluative skills. KK advocates accept this challenge: familiar objections to Williamson aim to provide precisely this sort of argument.

In light of the similarity between the student’s problematic reasoning process in FORWARD and the problematic reasoning process in GLIMPSE, attempts to defend KK will most naturally focus on 3\* (the assumption that the students know that the teacher’s announcement, TAF, is true). This assumption is structurally similar to Williamson’s assumption that the student knows G in the case of GLIMPSE. Our first point, however, is that such attempts will face similar dialectics. Since by no means do we take the student’s knowledge of TAF to reflect any general epistemological principle that begs the question against KK, and since (as we have argued in section 2) TAF is not otherwise clearly problematic, defending KK by rejecting 3\* requires independent argumentation.

---

<sup>27</sup> Or, for that matter, its counterpart in the Mr. Magoo argument.

<sup>28</sup> To see this, consider knowledge in a (normal) modal logic setting. Given G, if a subject knows that the exam is not on a certain day, then there can be no accessible worlds where the exam is too close to that day. In light of this, if the student has any substantive knowledge at all, a margin-for-error principle structured like G requires the epistemic accessibility relation to be intransitive. For, employing a metric of possible worlds to reflect the student’s knowledge of possible exam days, if the accessibility relation is transitive (i.e. if KK is valid in this model), either the student has no relevant knowledge at all, or what she knows is true in the entire possible world model (which is assumed not to be the case). This is because knowing that one knows requires not only truth in close possible worlds but also truth in worlds close to those close worlds and further iterations of knowledge extend in this way indefinitely.

---

## 4.2 Attempts to Save KK

Our second point is that such an argument will not be easy to provide. Given that FORWARD makes no appeal to the inexactness of a source of knowledge, familiar objections to Williamson's arguments will fail to apply. We will not be able to discuss the whole range of objections in detail. To demonstrate, we consider two challenges to GLIMPSE, and show that they fail to extend to FORWARD.<sup>29</sup>

A first objection Sharon and Spectre (2008)<sup>30</sup> runs as follows: as argued above, any generalization the student is assumed to know will only concern her particular state and the limits of her perceptual and evaluative skills. So, for instance, it would be true to say that given the imperfect nature of her glimpse, for any number  $n$ , if the student knows *from the glimpse* that the exam will not take  $n$  days from the end of the semester, then the exam will also not take place  $n+1$  days from the end of the semester. However, unless Williamson is tacitly appealing to a general principle, the student is not in a position to know  $G$ , but only something weaker:

$G^*$ : For any number  $n$ , if the student knows *from the glimpse* that the exam will not take place  $n$  days from the end of the term, then the exam will not take place  $n+1$  days from the end of the term.

At some point – after the first round of inference via KK and CLOSURE, but surely after several rounds – the student's reasoning must employ knowledge that is no longer based solely on the glimpse, but (partly) on inference. This reasoning cannot be grounded in  $G^*$ . The student's reasoning can therefore be blocked by rejecting  $G$  rather than rejecting KK.<sup>31</sup>

Another objection to Williamson's arguments against KK, due to Cohen and Comesaña (2013a),<sup>32</sup> also targets the knowability of generalizations like  $G$ . On the model that Cohen and Comesaña put forward, whether  $G$  is true depends on how far the actual designated exam day is from the estimate based on the glimpse. Imagine, for instance, that the glimpse makes it seem as if the exam is roughly 5 days from the end of the semester. Let us assume that this means that the student knows that the exam will take place between 3 and 7 days from the end of the semester. (Let this represent how inexact the student's knowledge is, for the sake of the example.) Cohen and Comesaña's point, however, is that  $G$  does not follow from this inexactness. For suppose that the designated day of the

---

<sup>29</sup> Other challenges concerning inexact knowledge will not apply to FORWARD for similar reasons. As far as we can see, other arguments in favor of KK (like Greco (2014)) are also irrelevant.

<sup>30</sup> For a similar response to Williamson's anti-KK argument, see Dokic and Egré (2009).

<sup>31</sup> Williamson does claim that an argument against KK can also rely on SURPRISE. And clearly, unlike  $G$ , the teacher's announcements in SURPRISE, TAS, is not implicitly restricted: Knowledge of TAS is not about some inexact source. Nevertheless, the only motivation for resolving SURPRISE via the rejection of KK is that, on Williamson's view, this solution extends to GLIMPSE and "any adequate diagnosis of [SURPRISE] should extend to [GLIMPSE]" (2000, p.138). Given the difference just noted between GLIMPSE and SURPRISE, we think there is good reason to reject the requirement to extend any solution to SURPRISE to GLIMPSE. But independently of that, we think that as an analysis of what is wrong with the student's reasoning in SURPRISE, a rejection of KK has drawbacks. Resolving SURPRISE by rejecting KK leaves no role for the observation that by Friday morning the students will have received counterevidence that no longer allows them to continue to know that there will be a surprise exam. We thereby prefer the standard solution to SURPRISE, which is also available regardless of the validity of KK.

<sup>32</sup> See Stalnaker (2015) for a similar defense of KK.

---

exam is in fact 3 days from the end of the semester. If so, then, contrary to G, the student knows that the exam will not take place 2 days from the end of the semester even though it will take place 3 days from the end. While this may not be the case in the example Williamson constructs (we can stipulate that it is not the case, i.e. stipulate that G is true in the example) – the student (on any scenario) is not in a position to know this.<sup>33</sup> Since there is no way for the student to know that the actual exam will not take place at the very edge of her margin for error, she is not in a position to know that G (or G\*, of that matter) is true. She may learn more about her specific situation or about her perception in general, but this will at best narrow the gap in her knowledge. She cannot generate knowledge of G by herself and this is sufficient for blocking the reductio argument. There is no need to reject KK.<sup>34</sup>

To be clear, the point we wish to make is not that Williamson is mistaken or that any of these objections is correct. The point is rather that objecting to our analysis of FORWARD along the same lines is not an option. The natural point of departure for such an objection would be to question 3\*. But in the case of FORWARD there is no reason to restrict the student's knowledge along the lines of Sharon and Spectre (2008), and no analogous restriction that suggests itself. As opposed to knowledge of G, 3\* is not supposed to be based on the student's knowledge about the inexactness of her perception. There seems to be no reason to suppose that she knows any less than what the teacher tells her. Cohen and Comesaña's model also fails to apply. In cases like GLIMPSE, Cohen and Comesaña could describe a non-far-fetched scenario on which G is false, and argue that the student in GLIMPSE cannot rule out the possibility that she is in that scenario. In FORWARD, however, the teacher, who knows the exact date of the exam and knows what the students can independently know, is in a position to guarantee to the students that they are not in a scenario on which TAF is false. Cohen and Comesaña's argument relies on the subject's inability to evaluate the gap between actuality and each different possibility left open by perception. When it comes to knowledge from an outside source, no such evaluation is required (on behalf of the subject). If KK advocates wish to question 3\* they will need a novel argument. We cannot say that providing such an argument is impossible, but we also cannot see how it would go.

One last point in support of 3\* relies on a comparison between the teacher's announcement in SURPRISE (TAS) and the teacher's announcement in FORWARD (TAF). According to the standard response to SURPRISE, TAS is knowable in the epistemic state the students are in when they hear it. But in certain epistemic states the student considers in their reasoning, TAS implies an unknowable sentence of the form "p and I do not know that p". Since, as we argued in section 2, there is no similar problem with TAF, denying that TAF is known upon its announcement (i.e. rejecting 3\*) while not denying that TAS is known upon its announcement seems implausible. KK advocates can, of course, extend their denial of 3\*, and (in case they initially tended to favor the standard response to SURPRISE) retroactively deny that TAS is known to the student in SURPRISE in the first place.

---

<sup>33</sup> Interestingly, Cohen and Comesaña (2013b) themselves think of it as a "quasi-skeptical" scenario.

<sup>34</sup> Williamson (2013) objects that Cohen and Comesaña's model, which allows what he calls cliff-edge knowledge, is implausible. Their reply (2013b) is that although possible on their model, such cases are very unlikely ("quasi-skeptical" as noted in the previous note). We need not take a stand on this matter here.

---

Our point, however, is that accepting KK and the standard solution to SURPRISE is not a stable view.

The argument based on FORWARD fills an important gap as an argument against KK. Unlike arguments based on cases of inexact knowledge (like GLIMPSE), the reasoning it concerns is not based on knowledge of the limits of one's own perceptual and evaluative capacities. Unlike other cases that do not involve inexact sources (like SURPRISE), it leaves no option more attractive than the rejection of KK open. These considerations make this argument exceptionally powerful in comparison to current alternatives.

### 4.3. Why KK Fails

FORWARD is special in giving rise to an argument against KK without relying on the subject's knowledge of her limitations (or on a general principle about the nature of knowledge). But if the problem does not lie in margin for error or self-generated knowledge of the limits one's own perceptual capacities, why does KK fail in cases like FORWARD?

One possible explanation, the one we tentatively advocate here, is that in some cases the strongest thing we know (in a given context) is not something we have good reasons to believe that we know. The requirements for knowing that it is true are lower than the requirements for justifiably believing that we know that it is true, let alone knowing that we know that it is true.<sup>35</sup> The basic idea can nicely be demonstrated by a closer look at Kripke's cards example. In the example, I stick the queen of hearts card at some random place near the middle of the deck. Let ' $q_n$ ' express 'the queen of hearts is not anywhere between the first card and card number  $n$  from the top of the deck', where  $n \in \{1, 2, 3, \dots, 52\}$ . We can now construct an  $n$ -tuple of propositions ordered by their strength as follows:  $Q = \langle q_1, q_2, q_3, \dots, q_{52} \rangle$ . Many of the propositions in  $Q$  are not known to me (among those, many are simply false). But I do know some of the weaker propositions. (I know  $q_1$  for instance.) In any event, if I can know any of the propositions in  $Q$  before examining any of the cards, then there is one proposition that is the strongest element of  $Q$  that I can know then. But then, if I also know something like CARDS (or, alternatively,  $K(q_n) \supset (q_{n+1})$ ), it is impossible for me to have higher-order (de dicto) knowledge of the strongest element of  $Q$  that I can know. If I have higher-order knowledge of some element of  $Q$ , knowledge of CARDS would imply that it is not the strongest element of  $Q$  I can know.

The knowability of CARDS may, perhaps, be contested along the lines of Cohen and Comesaña (2013a). Nevertheless, a similar line of reasoning applies in cases where I know something very much like CARDS on the basis of testimony. This is what happens in cases like FORWARD. In such cases, the strongest proposition one knows depends on the degree to which one knows the relevant piece of testimony, and the degree to which they know other elements of the  $n$ -tuple in question. In section 2 we looked at a case in which the student has  $K_0^7(\neg e_1)$  and  $(K_0(K_0(\neg e_n) \supset \neg e_{n+1}))$ . Letting  $E = \langle \neg e_1, (\neg e_1 \wedge \neg e_2), \dots \rangle$  be an  $n$ -tuple of propositions about possible exam days ordered by strength, the strongest element of  $E$  that the student can know turned out to be  $(\neg e_1 \wedge \neg e_2)$ . If we

---

<sup>35</sup> Since we will only be considering contexts in which justified belief about whether one knows would amount to knowledge that one knows, we waive the distinction in what follows.



allow higher-order knowledge ( $K_0^2(K_0(\neg e_2) \supset \neg e_3)$  or  $K_0(K_0^2(\neg e_2) \supset K_0(\neg e_3))$  and  $K_0(K_0^2(\neg e_1) \supset K_0(\neg e_2))$ ), this would get the student one step further, viz. to possibly knowing ( $\neg e_1 \wedge \neg e_2 \wedge \neg e_3$ ). But whatever the strongest proposition she knows may be, in order to know that she knows it, the student would need to know the degree to which she knows  $\neg e_1$  (and, needless to say, this is not something she can know). What she knows does not support second-order knowledge of the strongest element of E she can know. (If she did have second-order knowledge of the proposition in question, it would no longer be the strongest element of E she can know.)

By way of explanation, since CLOSURE and TAF merely imply something of the form  $K^n(A_m) \supset K^{n-1}(A_{m+1})$ , the strongest proposition the student can know will always fall short of higher-order knowledge. A pattern like  $K^n(A_m) \supset K^{n-1}(A_{m+1})$  exhausts itself, such that the final justified step the student makes only delivers first order knowledge that some particular day is not the exam day (assuming she has sufficient higher-order knowledge to begin with, ( $K^n(A_m)$ )).

Denying this, worrying that the student can always rule out the day before the designated exam day as a possible exam day on the basis of  $3^*$ , will involve weighty assumptions about how much the student initially knows. The strength of these KK-driven assumptions will grow in proportion to the size of  $i$ . Since the semester in FORWARD can be infinite (and the designated exam day,  $i$ , can be as high as we wish), KK is shown to be extremely strong. Given any finite value of  $n$  in  $K^n(A_m)$  and any strong assumption  $K^{n-1}(K(A_m) \supset (A_{m+1}))$ , if  $N$  (the number of semester days) is greater than  $n$ , knowledge will run out at some point. Given the futility of even these strong assumptions regarding FORWARD, and the relatively restrained character of  $3^*$ , we have good reason to think that KK fails.

## 5 Conclusion

We conclude by pointing out ways in which our reaction to FORWARD complements standard reactions to SURPRISE.

According to classical accounts of the original surprise exam paradox (SURPRISE), to succeed in excluding the possibility of a Friday exam, it must be the case that TAS is not only true on Friday morning, but also known to the student at that time. Generally speaking, the problem in the student's reasoning is that they are discussing an epistemic state in which the relevant proposition is true but not knowable. Going on to reason as if it was not only true, but knowable at that time, they derive falsehoods.

Despite the fact that no relevant proposition in the present paradox (FORWARD) is structurally unknowable, something similar happens: while it is true that the student knows that there will not be an exam on the second day, and while it is true that there will not be an exam on the third day, the student fails to know these propositions. Going on to reason as if they were not only true, but known, she derives further falsehoods.

Both SURPRISE and FORWARD can be viewed as concerning reasoning based on discharging knowledge-to-fact conditionals ('(for any day  $n$ ) if you know on the morning of day  $n$  that day  $n$  is the exam day, then day  $n$  is not the exam day'; '(for any day  $n$ ) if you know before the semester begins that day  $n$  is not the exam day, then day  $n+1$  is not the exam day'). Prevalent solutions to

---

SURPRISE focus on the unknowability of the conditional in epistemic states in which the antecedent is assumed to be known (i.e. in any state in which the relevant conditional can be discharged). Our solution to FORWARD presupposes that knowledge of all instances of the relevant conditional is unproblematic. It focuses instead on the requirements for knowing the antecedents that are employed to discharge those instances. In this way it complements the insights about knowledge gained by resolving SURPRISE.<sup>36</sup>

## References

- Cohen, S. and Comesaña, J. (2013a). Williamson on Gettier cases and epistemic logic. *Inquiry*, 56, 15–29.
- \_\_\_\_\_. (2013b). Williamson on Gettier cases in epistemic logic and the knowledge norm for rational belief: A reply to a reply to a reply. *Inquiry*, 56, 400–415.
- Cresto, E. (2012). A defense of temperate epistemic transparency. *Journal of Philosophical Logic*, 41, 923–55.
- Das, N. and Salow, B. (2018). Transparency and the KK Principle. *Noûs* 52, 3–23.
- Dokic, J. and Egré, P. (2009). Margin for error and the transparency of knowledge. *Synthese*, 166, 1–20.
- Dorst, K. (2019). Abominable KK failures. *Mind*, 128, 1227–1259.
- Greco, D. (2014). Could KK be OK? *Journal of Philosophy*, 111, 169–97.
- \_\_\_\_\_. (2015). Iteration and Fragmentation. *Philosophy and Phenomenological Research*, 91, 656–73.
- \_\_\_\_\_. (2017). Cognitive mobile homes. *Mind*, 126, 93–121.
- Kim, B. and Vasudevan, A. (2017). How to expect a surprising exam. *Synthese*, 194, 3101–3133.
- Kripke, S.A. (2011). *Philosophical Troubles*. Oxford University Press.
- McHugh, C. (2008). Self-knowledge and the KK principle. *Synthese*, 173, 231–57.
- Pagin, P. (1990). Review of Roy Sorenson, *Blindspots*, Clarendon Press, Oxford 1988. *History*

---

<sup>36</sup> We would like to thank Dan Baras, Mark Bowker, Meir Buzaglo, David Enoch, Eran Fish, David Heyd, Aviv Hoffmann, Ran Lanzet, Eric Olson, Andy Peet, Saleet Pitcovski, and Roy Sorenson, for valuable input on earlier drafts. We also thank the audiences at the Tel-Hai philosophy seminar and the 22nd conference of the (New) Israeli Philosophy Association. Special thanks to a reviewer of this journal for invaluable input and challenges that significantly improved this paper.

Levi Spectre's research is supported by the Swedish Riksbankens Jubileumsfond as part of the 'Knowledge Resistance: Causes, Consequences, and Cures' research project (reference number: M18-0310:1).

- 
- and *Philosophy of Logic*, 11, 243–245.
- Salow, B. (2018). The externalist's guide to fishing for compliments. *Mind*, 127, 691–728.
- \_\_\_\_\_. (2019). Elusive externalism. *Mind*, 128, 397–427.
- Sharon, A. and Spectre, L. (2008). Mr. Magoo's Mistake. *Philosophical Studies*, 139, 289–306.
- \_\_\_\_\_. (2013). Epistemic closure under deductive inference: What is it and can we afford it? *Synthese*, 190, 2731–2748.
- \_\_\_\_\_. (2017). Evidence and the openness of knowledge. *Philosophical Studies*, 174, 1001–1037.
- Smith, M. (2010). What Else Justification Could Be. *Noûs*, 44, 10–31.
- \_\_\_\_\_. (2016). *Between Probability and Certainty: What Justifies Belief*. Oxford: Oxford University Press.
- Sorensen, R.A. (1982). Recalcitrant variations of the Prediction Paradox. *Australasian Journal of Philosophy* 69 (4), 355–62.
- \_\_\_\_\_. (1988) *Blindspots*. Oxford: Clarendon Press.
- \_\_\_\_\_. (1999). Infinite backward induction argument. *Pacific Philosophical Quarterly*, 80, 278–283.
- Spectre, L. (2009). *Knowledge Closure and Knowledge Openness: A Study of Epistemic Closure Principles*. PhD thesis, Department of Philosophy, Stockholm University.
- Stalnaker, R. (2009). Iterated belief revision. *Erkenntnis*, 70, 189–209.
- \_\_\_\_\_. (2015). Luminosity and the KK Thesis. In S. Goldberg (Ed.), *Externalism, Self-Knowledge, and Skepticism* (pp. 1–19). Cambridge: Cambridge University Press.
- Williamson, T. (2000). *Knowledge and Its Limits*, Oxford University Press.
- \_\_\_\_\_. (2013). Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson on Gettier cases in epistemic logic. *Inquiry*, 56, 77–96.
- Wright, C. and Sudbury, A. (1977). The paradox of the unexpected examination. *Australasian Journal of Philosophy*, 55, 41–58.