

# STILL A NEW PROBLEM FOR DEFEASIBILITY: A REJOINDER TO BORGES

John N. WILLIAMS

**ABSTRACT:** I objected that the defeasibility theory of knowledge prohibits you from knowing that you know that  $p$  if your knowledge that  $p$  is *a posteriori*. Rodrigo Borges claims that Peter Klein has already satisfactorily answered a version of my objection. He attempts to defend Klein's reply and argues that my objection fails because a principle on which it is based is false. I will show that my objection is not a version of the old one that Klein attempts (unsuccessfully) to address, that Borges' defence of Klein's reply fails and that his argument against my new objection leaves it untouched.

**KEYWORDS:** defeasibility theory of knowledge, *a posteriori*, Peter Klein

I objected that the defeasibility theory of knowledge prohibits you from knowing that you know that  $p$  if your knowledge that  $p$  is *a posteriori*.<sup>1</sup> Rodrigo Borges claims that Peter Klein has already satisfactorily answered a version of my objection.<sup>2</sup> He attempts to defend Klein's reply and argues that my objection fails because a principle on which it is based is false.

I will show that my objection is not a version of the old one that Klein attempts (unsuccessfully) to address, that Borges' defence of Klein's reply fails and that his argument against my new objection leaves it untouched.

## 1. The Old Objection, Klein's Unsatisfactory Reply and Borges' Defence of Klein

What might be called the *early* defeasibility theory of propositional knowledge may be formulated as follows.

---

<sup>1</sup> John N. Williams, "Not Knowing You Know: A New Objection to the Defeasibility Theory of Knowledge," *Analysis* 75 (2015): 214.

<sup>2</sup> Rodrigo Borges, "A Failed Twist to an Old Problem: A Reply to John Williams," *Logos and Episteme* VII, 1 (2016): 75-81, citing Peter Klein, "A Proposed Definition of Propositional Knowledge," *Journal of Philosophy* 68 (1971): 471-82. This is a more charitable description of Borges' strategy than his own. What he actually says is that "Klein himself answered a version of this objection in 'A Proposed Definition of Propositional Knowledge.' Williams' paper adds a new twist to the objection Klein answered more than forty years ago. I will argue that Williams' objection misses its target because of this new twist." Surely my objection cannot fail simply because it is a *different* objection from the one Klein addresses.

John N. Williams

You know that  $p$  just in case you have a justified true belief that  $p$  and there is no defeater  $D$  of your justification for believing that  $p$

where

$D$  is a defeater of your justification for believing that  $p$  just in case it is a truth such that believing it would render your belief that  $p$  unjustified.

This is essentially the theory that Klein proposes, calling a defeater, “a disqualifying proposition,”<sup>3</sup> and taking “you have a justified true belief that  $p$ ” as synonymous with “ $p$  is evident to you.”<sup>4</sup> We may represent this definition of knowledge as:

$S$  knows that  $p$  just in case

- (1)  $p$
- (2)  $S$  believes that  $p$
- (3)  $S$  is justified in believing that  $p$
- (4) There is no defeater of  $S$ 's justification for believing that  $p$ .

As Borges observes,<sup>5</sup> Klein considers the objection that “If the definition were accepted, it would never be true that  $S$  knows that he knows that  $p$  because he could never know that the fourth condition held.”<sup>6</sup> This is the old objection. Now, a *first way* to analyse  $S$ 's knowledge that she knows that  $p$  in terms of the early defeasibility theory is to substitute “ $S$  knows that  $p$ ” for “ $p$ ” in each of the four conditions above, yielding

---

<sup>3</sup> Klein, “A Proposed Definition,” 475.

<sup>4</sup> Klein (“A Proposed Definition,” 475) proposes that “ $S$  knows that  $p$  at  $t$  if and only if

- (i)  $p$  is true;
- (ii)  $S$  believes  $p$  at  $t$ ;
- (iii)  $p$  is evident to  $S$  at  $t$ ;
- (iv) there is no true proposition such that if it became evident to  $S$  at  $t$ ,  $p$  would no longer be evident to  $S$ .”

Roderick M. Chisholm, in *Theory of Knowledge* (Englewood Cliffs, N.J.: Prentice Hall, 1966), 22, has it that sufficient conditions for  $p$  to be evident to  $S$  are that it is more reasonable for  $S$  to believe  $p$  than to withhold belief in  $p$  and that there is no proposition such that it is more reasonable for  $S$  to believe it than it is for him to believe  $p$ . To these Klein adds a third, that  $S$  has no reason to believe that the situation is abnormal (“A Proposed Definition,” 473). Nonetheless, he takes “evident” as interchangeable with “justified” (“A Proposed Definition,” 473).

<sup>5</sup> Borges, “A Failed Twist,” 76.

<sup>6</sup> Klein, “A Proposed Definition,” 480.

- (1') *S* knows that *p*
- (2') *S* believes that she knows that *p*
- (3') *S* is justified in believing that she knows that *p*
- (4') There is no defeater of *S*'s justification for believing that she knows that *p*.<sup>7</sup>

This is precisely what Klein does in response to the old objection. He then says that "It seems quite clear that these conditions could be fulfilled; or rather, the definition itself does not rule out the possibility that these conditions are fulfilled."<sup>8</sup> This misses the point of the objection, which, to concentrate on the fourth condition, was *not* that *there can be no defeater of S's justification for believing that she knows that p*, but rather that she could not know (4), in other words that *she could not know that there is no defeater of her justification for believing that p*. Klein could try claiming that the definition does not rule out the possibility of *S* knowing (4), but that would merely beg the question. To be fair to Klein however, the objection itself gives no reason why *S* cannot know (4). It seems likely that Klein fails to answer the objection he anticipates because it is easy to confuse this first way of analysing *S*'s knowledge that she knows that *p* in terms of the early defeasibility theory with a *second way*. This is to analyse it as *S*'s knowledge of each of the four conditions, yielding:

- (1'') *S* knows that *p*
- (2'') *S* knows that she believes that *p*
- (3'') *S* knows that she is justified in believing that *p*
- (4'') *S* knows that there is no defeater of her justification for believing that *p*

The difference between (4') and (4'') is crucial, because the objection was not that (4') cannot be satisfied but that (4'') cannot be satisfied. That objection is premised upon this second way of analysis, not the first. This enables us to see that the objection relies upon the *knowing you know principle*:

If you know that you know that *p*, then you know the content of each necessary condition of your knowing that *p*.

We can now give the form of the *argument for the old objection*:

---

<sup>7</sup> Klein ("A Proposed Definition," 481) puts this result as

"1' *S* knows that *p*, ...

2' *S* believes that he knows that *p*;

3' '*S* knows that *p*' is evident to *S*...

4' There is no disqualifying proposition for '*S* knows that *p*.'"

<sup>8</sup> Klein, "A Proposed Definition," 481.

John N. Williams

- I. Given the early defeasibility theory, if you know that  $p$  then (4) there is no defeater of your justification for believing that  $p$ .
- II. Given the early defeasibility theory, if you know that you know that  $p$ , then (4'') you know that there is no defeater of your justification for believing that  $p$ . (from I plus the knowing you know principle)
- III. You cannot know that there is no defeater of your justification for believing that  $p$ .
- IV. Given the early defeasibility theory, you cannot know that you know that  $p$ . (from II and III)

We should note that no reason is given for III. Borges then attempts to “substantiate Klein’s reply by providing a logically possible case” in which (1')-(4') are all true.<sup>9</sup> This strongly suggests that Borges has in mind the second way of analysing second-order knowledge, which explains why he follows Klein in missing the point of the old objection. If he were to succeed in giving a case in which (1')-(4') are all true then this would show that that IV, namely the conclusion of the old objection, is false. But it would not show what goes wrong with the argument for that conclusion.

Borges then attempts to give the following case in which (1')-(4') are all true, namely *Computer Screen*. You appear to see a computer screen in front of you, which is at least part of your justification for believing that there is one in front of you. He stipulates that this is a normal case of perceptual experience, that you know that there is a computer screen in front of you and that each of (1)-(4), namely the conditions of the early defeasibility theory, are met. I have no quarrel with this. I agree that the early defeasibility theory may allow you first-order knowledge.<sup>10</sup>

Setting external-world scepticism aside,<sup>11</sup> there are many different sorts of possible truths that, if you believed them, would render you unjustified in believing that there is a computer screen in front of you. These include the possible truth that you are prone to hallucination when near electrical equipment, that there is a trick of the light that causes you to mistake a printer for a computer screen or that makes the screen appear to be in front of you when it is in fact to one side, that what you are looking at is a realistic facsimile of a computer screen, and so on. Since Borges has stipulated that this is a normal case of perception, these are likewise stipulated not to be truths, in other words that there is no

---

<sup>9</sup> Borges, “A Failed Twist,” 76-77.

<sup>10</sup> Williams “Not Knowing You Know,” 215.

<sup>11</sup> As I did in “Not Knowing You Know,” 214.

defeater of your justification for believing that there is a computer screen in front of you.

So far so good. Borges continues, "Suppose further that I reflect on whether I know [that there is a computer screen in front of me], realize that it is a normal case of perceptual experience, and come to believe I do know it."<sup>12</sup> He goes on to claim that it is clear that there is no "defeater of my justification for *believing that I know* that there is a computer screen in front of me" because "*by assumption*, nothing like that is true in this situation."<sup>13</sup>

This does not follow. Borges has stipulated that there is no defeater of your justification for *believing that there is a computer screen in front of you*, with the result that the early defeasibility theory allows you to know that there is one in front of you. But this is not to stipulate that there is no defeater of your justification for *believing that you know that there is a computer screen in front of you*. For it is possible that although there is no defeater of your justification for believing that *p*, there is nonetheless a defeater of your justification for believing that you *know* that *p*. Such a case arises in *Biology Teacher*.

Someone who looks exactly like your biology teacher tells you that the insect that you have captured is a Brazilian wandering spider. You know your teacher to be reliable, sincere and an expert on spiders and so you come to believe that she has transmitted her knowledge to you, with the result that you now know that the insect is a Brazilian wandering spider. But unbeknownst to you, (*D*) the person who has just told you this is the identical twin of your teacher who is no expert on spiders, but can still classify them fairly accurately (or to a degree of statistical accuracy that warrants belief without constituting knowledge-grade justification).

If you were to come to believe (*D*) then you would cease to be justified in thinking that you *know* that the insect is a Brazilian wandering spider, but you would still be justified in thinking that it *is* one.

A second problem that afflicts Borges' defence of Klein is that although your perceptual experience of what appears to be a computer screen in front of you might be normal, it is far from obvious how you can *realize* that it is normal merely by reflection. I will return to this point below in section 3.

Borges concludes his defence of Klein by saying that "Defeasibility does not make it impossible for there to be second-order knowledge."<sup>14</sup> Borges means that the early defeasibility theory does not prohibit you from knowing that you know

---

<sup>12</sup> Borges, "A Failed Twist," 77.

<sup>13</sup> Borges, "A Failed Twist," 78 (my italics).

<sup>14</sup> Borges, "A Failed Twist," 78.

John N. Williams

that  $p$ . But I nowhere claim that it does. Instead my new objection is only that it prohibits you from knowing that you know that  $p$  if your knowledge that  $p$  is a *posteriori*.<sup>15</sup> I now turn to my new objection.

## 2. My New Objection

My new objection starts with the *conceptual principle*.

If the satisfaction of a condition at least partly constitutes an instance of a concept, then knowing that such an instance obtains requires you to know that the condition is satisfied.<sup>16</sup>

For example, since the concept of a triangle is at least partly constituted by it being three-sided, you know that a figure is a triangle only if you know that it is three-sided. The early defeasibility theory is intended to be an analysis of the concept of knowledge, with the result that the satisfaction of condition (4) is supposed to at least partly constitute an instance of the concept of knowledge. So given the early defeasibility theory, by the conceptual principle, your knowing that there is an instance of your knowledge requires you to know that (4) is satisfied. To take Borges's own example, on the early defeasibility theory, knowing that you know that there is a computer screen in front of you requires you to know that

(A) There is no defeater of your justification for believing that there is a computer screen in front of you.

This is equivalent to

(B) There is no truth such that believing it would render you unjustified in believing that there is a computer screen in front of you.

This in turn is equivalent to

---

<sup>15</sup> Williams, "Not Knowing You Know," 214.

<sup>16</sup> Williams, "Not Knowing You Know," 215. Strictly speaking, you might not have the concept of the satisfaction of a condition, and so given the plausible principle that you can have a belief only if you have the ability to think the thought of its content, you could not believe, nor therefore know, anything about the satisfaction of conditions. A more cumbersome but more accurate formulation of the principle is as follows.

If the satisfaction of a condition that  $q$  at least partly constitutes an instance of a concept  $C$ , then knowing that  $p$ , where  $p$  reports an instance of  $C$ , requires you to know that  $q$ .

For ease of exposition I will stick with the less cumbersome formulation. Nothing turns upon this.

## Still a New Problem for Defeasibility: A Rejoinder to Borges

- (C) Every truth is not such that if you were to believe it then this would render you unjustified in believing that there is a computer screen in front of you.

But how could you possibly know (C)? You cannot know it *a priori*. You cannot tell in advance that (C) is true. Mere reflection will not allow you to foresee all threats to your justification of your belief that there is a computer screen in front of you. As we saw above in section 1, there are many different sorts of possible truths that, if you believed them, would render you unjustified in believing that there is a computer screen in front of you. Who is to say what all of these truths are? The only other way for you to know (C) is to know every truth and ascertain that each is not such that if you were to believe it then this would render you unjustified in believing that there is a computer screen in front of you. This is not a way that you can follow, for the simple reason that being less than omniscient, you cannot know every truth.

This problem for the early defeasibility theory arises when your first-order knowledge that *p* is *a posteriori*. This is because the justification for your belief that *p* comes from experience, with the result that reflection alone will not enable you to verify that there are no other empirical truths that would defeat your experiential justification should you believe them. Nor will experience enable you to verify this, because you cannot consider every empirical truth.

We can now give the form of the *argument for my new objection*:

- I'. Given the early defeasibility theory, the satisfaction of condition (4) at least partly constitutes an instance of the concept of knowledge.
- II'. Given the early defeasibility theory, if you know that you know that *p*, then (4'') you know that there is no defeater of your justification for believing that *p*. (from I' plus the conceptual principle)
- III'. There are only two ways in which you could know that there is no defeater of your justification for believing that *p*, namely by knowing *a priori* that every truth is not such that if you were to believe it then this would render you unjustified in believing that *p*, or by knowing every truth and ascertaining that each is not such that if you were to believe it then this would render you unjustified in believing that *p*.
- IV'. If your knowledge that *p* is *a posteriori*, then you cannot know *a priori* that every truth is not such that if you were to believe it then this would render you unjustified in believing that *p*, nor can you know every truth and ascertain that each is not such that if you were to believe it then this would render you unjustified in believing that *p*.
- V'. Given the early defeasibility theory, you cannot know that you know that *p* if your knowledge that *p* is *a posteriori*. (from II', III' and IV')

But any satisfactory theory must allow you not only to have *a posteriori* knowledge, but also to know that you have it. The early defeasibility theory does not allow this. So the theory is unsatisfactory.

### 3. Borges' Failed Reply to My New Objection

As in the argument for the old objection, one step in this argument is II. However, there are three important differences. First, my argument for my new objection appeals to the conceptual principle, whereas that for the old objection appeals to the knowing you know principle. Second, the argument for the old objection contains an important premise that is unsupported, namely III. In contrast, my argument for my new objection provides reasons why you cannot know that there is no defeater of your justification for believing that *p* when that justification is *a posteriori*. Third, the conclusion of the argument of the old objection is that given the early defeasibility theory, you cannot know that you know that *p*. In contrast, that of my new objection is only that given the early defeasibility theory, then you cannot know that you know that *p* if your knowledge that *p* is *a posteriori*.

Given these differences, it seems quite a stretch for Borges to describe my objection as a "new twist to the old objection."<sup>17</sup> My objection is substantially different from the old objection that both Klein and Borges fail to address.

Nor is it true, contrary to Borges, that my objection is that it is impossible for the early defeasibility theory to be true and also for one to acquire second-order knowledge. That is the old objection, not mine.<sup>18</sup>

In response to my objection Borges again appeals to *Computer Screen*. He argues that since you may realize by reflection that this is a normal case of perceptual experience, you may realize in advance that there is no defeater of your justification for believing that there is one in front of you. This is supposed to follow because he has already stipulated a normal case of perceptual experience to

---

<sup>17</sup> Borges, "A Failed Twist," 78.

<sup>18</sup> In fact it seems to me that the early defeasibility theory allows you to know that you know that *p* if your knowledge that *p* is *a priori*. Suppose that you reflect upon the concepts of 2 and addition and come to recognize, and hence believe, that  $2 + 2 = 4$ , because you realize that this could not possibly be otherwise. By introspection, you may realize that you believe this and so come to recognize that what you yourself now believe is a conceptual truth (your justification for what you believe). You may now reason that since  $2 + 2$  could not possibly be other than 4, there could be no truth that would deprive you of your justification for thinking that  $2 + 2 = 4$ , were you to believe it, because no truth could change the fact that  $2 + 2$  could not possibly be other than 4, as you could still realize. You can indeed know in advance that there is no defeater of your justification for believing that  $2 + 2 = 4$ .



be one in which there are no truths such as that you are prone to hallucination when near electrical equipment and so on.<sup>19</sup>

There are at least two problems with this response. First, it is difficult to see how mere reflection on your apparent experience of a computer screen in front of you could tell you that this is a normal case of perceptual experience and not one in which, for example, you are looking at a realistic facsimile of a computer screen. You might know that your past experiences have overwhelmingly turned out to be veridical, thus providing you with inductive justification for thinking that your present experience is veridical too. But this would at least involve your memory of how past cases have turned out, not just your reflection on the present case.

Second and most importantly, *realizing* that this is a normal case of perception in which you appear to see a computer screen in front of you, involves *knowing* that there is one in front of you, so Borges' response assumes that if you know that there is one in front of you, then you may know that there is no defeater of your justification for believing that there is one in front of you. This begs the question of whether (4) is a necessary condition of knowledge. I argue that it can't be, for if it were a necessary condition, then you could not know that you have *a posteriori* knowledge, yet you could indeed know this.

We should also note that the conceptual principle makes no mention of *a priori* knowledge. However Borges represents it as CLAIM:

If the satisfaction of a condition at least partly constitutes an instance of a concept, then knowing that such an instance obtains requires you to know *a priori* that the condition is satisfied.<sup>20</sup>

Borges writes

Although Williams does not explicitly formulate CLAIM as requiring *a priori* knowledge, one must read CLAIM in this way lest his argument against Klein be made invalid, for Williams explicitly requires that *S* know *a priori* that she satisfies the no-defeater condition in order for her to know that she knows.<sup>21</sup>

But I do not even *implicitly* require that *S* knows *a priori* that she satisfies the no-defeater condition in order for her to know that she knows! In my

---

<sup>19</sup> Borges says that "Suppose that I reflect on whether I know ... realize that it is a normal case of perceptual experience ..." ("A Failed Twist," 77) and that "my total evidence bearing on the issue of whether I am justified in believing ... includes ... also my knowledge that this is a normal case of perceptual experience, that I am not drugged or otherwise visually impaired, and so on." ("A Failed Twist," 77).

<sup>20</sup> Borges, "A Failed Twist," 78.

<sup>21</sup> Borges, "A Failed Twist," 78, note 11.

John N. Williams

argument, II makes no mention of *a priori* knowledge. Nonetheless my argument is valid as it stands.

This misunderstanding undermines Borges' two remaining criticisms of my argument. The first of these is that there is a counterexample to CLAIM, namely *Mathematician*:

Suppose that *S*'s ability to prove some mathematical theorems at least partly constitutes the instance of the concept *S is a mathematician*. Timmy is unable to grasp any concepts involved by a mathematical proof. His calculus professor tells him that she can prove many mathematical theorems.<sup>22</sup>

Borges observes that intuitively, Timmy knows *a posteriori* that the professor is a mathematician, since his knowledge is based on his experience of what she tells him. But he does not know *a priori* that she has the ability to prove some mathematical theorems, for knowing that is also based on what she tells him. This falsifies CLAIM. So it does, but my argument relies upon the conceptual principle, not CLAIM.<sup>23</sup>

Borges' remaining criticism is that applying CLAIM to the concept of knowledge results in two false predictions. He assumes that a condition of your knowing that *p* is that you have a justified belief that *p*. Satisfying this condition at least partly constitutes the concept of knowledge.<sup>24</sup> So CLAIM predicts that if you know that you know that *p*, then you know *a priori* that you have a justified belief that *p*. Given that you know this *a posteriori*, this falsifies CLAIM.<sup>25</sup> So it does, but my claim is not CLAIM, only the conceptual principle.

Likewise, a condition of your knowing that *p* is that *p*, and satisfying this truth-condition at least partly constitutes the concept of knowledge. So by CLAIM, knowing that an instance of your knowledge obtains (in other words, knowing that you know that *p*) requires you to know *a priori* that the truth-condition is satisfied (in other words, you know *a priori* that *p*). Thus if you know

---

<sup>22</sup> Borges, "A Failed Twist," 79.

<sup>23</sup> Another problem with this example is that it is not enough to just *suppose* that *S*'s ability to prove some mathematical theorems at least partly constitutes the instance of the concept *S is a mathematician*. To falsify the conceptual principle, or even CLAIM for that matter, Borges needs an *actual* case in which the satisfaction of a condition at least partly constitutes an instance of a concept. I also note in passing that there is a kind of circularity in the supposition, since the satisfaction of the concept *S is a mathematician* already involves the concept of mathematics. However I am unsure what to make of this.

<sup>24</sup> Of course there are non-defeasibilists who will argue that the concept of knowledge is in no way constituted by justified belief.

<sup>25</sup> Borges, "A Failed Twist," 80.

that you know that  $p$ , then you know *a priori* that  $p$ .<sup>26</sup> Given that you may know that you know that  $p$  in a case in which your knowledge that  $p$  is *a posteriori*, this again falsifies CLAIM.<sup>27</sup> Yet again it does, but once again, my objection to the early defeasibility theory in no way relies upon CLAIM.

#### 4. Concluding Remarks

Both Klein and Borges fail to address the old objection to the early defeasibility theory and Borges' attempt to defend Klein against it fails. In any case my new objection is substantially different from the old one. Borges' attempt to fault my objection fails largely because he attributes to me a principle that I neither hold nor need.

In my "Not Knowing You Know," I argue that my new objection counts not only against the early defeasibility theory, but also against what might be called the *later* defeasibility theory originating from Klein:

You know that  $p$  just in case you have a justified true belief that  $p$  and there is no undefeated defeater  $D$  of your justification for believing that  $p$

where

$D$  is defeated by  $D^*$  just in case  $D^*$  is a truth such that believing it, in addition to believing  $D$ , preserves your original justification for believing that  $p$ .<sup>28</sup>

---

<sup>26</sup> This may be put equivalently as follows. Either you do not know that you know that  $p$  or your knowledge that  $p$  is *a posteriori*. This is why Borges thinks that I face a dilemma, since for some instances of " $p$ ," I would not want to accept either disjunct. See Borges, "A Failed Twist," 80.

<sup>27</sup> Borges, "A Failed Twist," 80-81.

<sup>28</sup> See Peter Klein, *Certainty: A Refutation of Scepticism* (Minneapolis: University of Minnesota Press, 1981). The early theory incorrectly excludes cases of knowledge, as shown by the case of *Tom Grabit*, in Keith Lehrer and Thomas Paxson, "Knowledge: Undefeated Justified True Belief," *Journal of Philosophy* 66 (1969): 225-37. Here is an adapted version.

Using your reliable vision and memory, you see someone who looks just like Tom Grabit stealing a book at the library, and on this basis believe that he stole a book. Unbeknownst to you, Tom's mother claims that he is away on a trip and has an identical twin who is in the library. But still unbeknownst to you, she is demented. Tom did steal a book.

Surely you know that Tom stole a book, but there is a truth that would render you unjustified in believing that Tom stole one were you to believe it, namely that Tom's mother claims that he is away on a trip and has an identical twin who is in the library. In contrast, the later defeasibility theory explains why you know that Tom stole a book. There is also another truth, namely that she is demented. If you were to believe both truths, then your original justification for believing that Tom stole a book would be preserved. Your justification is ultimately undefeated.

John N. Williams

In this rejoinder to Borges I have only discussed the early defeasibility theory, since as he observes, nothing in his reply to me depends on this distinction.<sup>29</sup> Everything I have said in this rejoinder will apply equally to the later theory.<sup>30</sup>

---

<sup>29</sup> Borges, "A Failed Twist," 76, note 3.

<sup>30</sup> I am most grateful to Claudio de Almeida for his insightful discussion of previous drafts.