

## Bet Accepted: A Reply to Freitag

Word Count (not including the abstract: 3466)

### ABSTRACT:

Wolfgang Freitag (2015) claims to have developed a proposal that solves Goodman's famous New Riddle of Induction. His proposal makes use of the notion of "derivative defeat"; the claim is that in certain circumstances, the projection of some predicates is *derivatively defeated*, i.e. it is inductively invalid. Freitag develops the proposal using some compelling examples, and then shows that it likewise applies to the argument at the basis of the New Riddle. There, he alleges, the projection of "grue" is derivatively defeated, whereas the projection of "green" is not. We thus have a justification for preferring the green induction over the grue induction.

I do not think Freitag has solved the New Riddle. The problem is that his proposal tacitly assumes certain counterfactual claims that a grue-speaker would dispute. I thus argue that Freitag's proposal begs the question against the grue-speaker, and I show why the grue-speaker would reject it.

### 1. Introduction

Wolfgang Freitag has recently written perhaps the most entertaining paper on the New Riddle of Induction in the last several decades (see Freitag, 2015: "I Bet You'll Solve Goodman's Riddle", this journal). It is presented as a dialogue between a "mysterious stranger" and a professor who has recently taught the Riddle to his class. In broad summary, it goes like this.

The stranger submits a bet to the professor: if you do not solve the New Riddle by tomorrow, I owe you \$50, but if you do solve it, you owe me \$50. The professor, obviously, accepts the bet. The stranger now offers the professor the following argument. Every day we

have “sampled” so far has been before tomorrow. Therefore, every day we have sampled has either been before tomorrow or been a day on which the professor has solved the New Riddle. Tomorrow, however, is not before tomorrow. Therefore tomorrow the professor will solve the New Riddle.

The professor now undertakes, with some helpful prompts from the stranger, to show what is wrong with the stranger's argument. In doing so, he develops a general account of when arguments like this are (inductively) invalid.<sup>1</sup> Having done so, however, the stranger points out that the professor's account applies equally well to the classic argument about grue emeralds at the basis of the New Riddle. In other words, the professor has unwittingly solved the New Riddle by showing why the stranger's argument is invalid.

I admire Freitag's paper. Not, unfortunately, because it solves the New Riddle; there I think it clearly fails. But as an exercise in creative writing, it is both immensely enjoyable and a welcome change of pace. My task here, however, is to show that he does not succeed. For Freitag's proposal implicitly relies on the faulty assumption that it is fair game in this dialectic to appeal to claims about counterfactual dependence.

This paper proceeds as follows. In §2 I review Freitag's proposal for solving the New Riddle. In §3 I argue that his proposal does not succeed because it implicitly relies on counterfactual claims that a grue-speaker would not grant. I conclude in §4.

## 2. Freitag's Proposal

---

<sup>1</sup> Freitag adopts Goodman's somewhat idiosyncratic usage of “valid” – not the notion that we traditionally associate with *deductive* inference. Goodman sometimes describes himself as looking for a criterion of inductive validity, e.g. on page 84 of *Fact, Fiction, and Forecast*: “What we want...is an accurate and general way of saying which hypotheses are confirmed by, or which projections are validly made from, any given evidence.”

Freitag's proposal relies on the notion of a *discriminating predicate*.<sup>2</sup> Discriminating predicates are defined relative to an induction set  $I$ . An induction set  $I$  is a set comprised of, on the one hand, a set of sampled elements  $I_\alpha$ , and on the other hand a set of unsampled elements  $I_\beta$ .  $I_\alpha$  constitutes the *sample set*, and  $I_\beta$  the *target set*. Now a predicate  $P$  is a *discriminating predicate* with respect to induction set  $I$  if and only if it meets two conditions:

1. For all elements of the sample set  $I_\alpha$ , we know that they are  $P$ .
2. We know that there is at least one element of the target set  $I_\beta$  that is not  $P$ .

A discriminating predicate is therefore, as Freitag puts it, “fully supported by the inductive evidence” (p. 257), yet it would be foolish to try to project it onto the target set, since we know that it does not apply to all members of the target set.<sup>3</sup> As Freitag says, “the projection of a discriminating predicate  $P$  is undermined by a *defeater*, namely the knowledge that some  $\beta$ 's are not  $P$ ” (p. 257).

As an example, let  $P$  be the predicate “is a day before tomorrow,” and take our induction set to be the set of days. Then the sample set is the set of all days that have occurred so far, and the target set is the set of all future days. Clearly,  $P$  applies to all members of the sample set, yet we know that it fails to apply to at least one member (in this case, *all* members) of the target set. Neither tomorrow nor any later day is before tomorrow. So “is a day before tomorrow” is a discriminating predicate: we must not project it onto any future days.

The most trivial case of a discriminating predicate is that of “is a member of the sample set.” Freitag dubs this predicate “ $\Delta$ .”  $\Delta$  is a discriminating predicate relative to *any* induction set. As a consequence, Freitag says, “the projection of  $\Delta$  is always perfectly supported by the

---

2 Freitag articulates the notion of a discriminating predicate on pages 256-257 of his paper.

3 Freitag says that we “must not” project  $P$  onto the elements of  $I_\beta$ . Presumably he means to ignore cases in which, say,  $I_\beta$  has an infinite cardinality, and we only know that one member of it is not  $P$ , for in such cases it might still be rational to project  $P$  onto at least part of the target set.

inductive evidence...yet it is never valid” (p. 258).

Now, remember that in Freitag's dialogue, this proposal is developed by the professor in order to refute the argument advanced by the stranger. The stranger's argument contains the predicate “is a day before tomorrow or is a day on which the professor has solved the New Riddle.” Note that this is *not* a discriminating predicate: we do not know that this predicate applies to all members of the sample set and not to some member of the target set, for we do not know that tomorrow the professor will not solve the New Riddle. But it *is* a *disjunctive weakening* of a discriminating predicate: the disjunct “is a day before tomorrow” is itself a discriminating predicate, as we just saw. The professor's first suggestion, then, is that projections of disjunctive weakenings of discriminating predicates are *derivatively* defeated by the fact that they contain a discriminating predicate – a predicate whose projection is *directly* defeated – as one disjunct.

But this evidently will not work, as the stranger points out. First, note that *all* predicates have disjunctive weakenings with a discriminating predicate as one of the disjuncts. Take any predicate  $Q$  and consider its disjunctive weakening  $Q$ -or- $\Delta$ . If we rule out all projections based on disjunctive weakenings of discriminating predicates, then  $Q$ -or- $\Delta$  cannot be projected. But if  $Q$ -or- $\Delta$  cannot be projected, then neither can  $Q$ . For if the projection of  $Q$ -or- $\Delta$  is defeated, then so is the projection of  $Q$ . As the stranger puts it: “some disjunctive weakenings of discriminating predicates must be projectible, or nothing ever is” (p. 259).

So we need to qualify the claim that projections of disjunctive weakenings of discriminating predicates are derivatively defeated. Clearly this cannot be true in all cases. So there are two options on the table:

1. *Some* projections of disjunctive weakenings of discriminating predicates are derivatively

defeated.

2. *No* projections of disjunctive weakenings of discriminating predicates are derivatively defeated.

Accepting (2) would entail that the stranger's initial argument, utilizing the predicate “is a day before tomorrow or is a day on which the professor has solved the New Riddle,” is valid. Indeed, it would likewise entail that a parallel argument utilizing the predicate “is a day before tomorrow or is a day on which the professor has *not* solved the New Riddle” is also valid. It would therefore permit us to infer contradictory conclusions. So (2) clearly cannot be correct. Thus we *must* accept (1). Our task is then to articulate *which* projections of disjunctive weakenings of discriminating predicates are defeated.

The professor takes the night to think about this, and the next day he comes back with the following idea. A disjunctive weakening of a discriminating predicate  $P$  is derivatively defeated if our evidence that it applies to the sample set is itself *dependent on* or *grounded in* our evidence that  $P$  applies to the sample set. For example, suppose I have sampled a bunch of emeralds, but I have not been paying attention to their color. Thus I know that they have all been sampled (i.e. that they are all  $\Delta$ ), and therefore that they are all  $\Delta$ -or-green. Is it now legitimate for me to infer that future emeralds, given that they have not yet been sampled, must be green? Intuitively, the answer is *no*. For my evidence that the sampled emeralds are all  $\Delta$ -or-green is dependent on my evidence that the sample emeralds are all  $\Delta$ , and  $\Delta$  is a discriminating predicate: I know that it does not apply to any emeralds outside of the sample set. I could equally well have noted that all of the sampled emeralds are  $\Delta$ -or-blue, and therefore inferred that future unsampled emeralds are blue. Conversely, if I *had* been paying attention to the color of the sampled emeralds, then I *could* legitimately project  $\Delta$ -or-green to the unsampled emeralds, and infer that they must be

green. For in this case, my evidence that the sampled emeralds are all  $\Delta$ -or-green is not dependent on my evidence that they are all  $\Delta$ . As the professor puts it, then, “a dependent projection stands or falls with the projection it is dependent on” (p. 262).

This account can be used to show why the stranger's initial argument is invalid. We should of course grant that all sampled days have either been days that are before tomorrow or days on which the professor has solved the New Riddle. But our evidence that all sampled days satisfy this predicate is, quite plainly, dependent on our evidence that all of the sampled days are before tomorrow. And that, again, is a discriminating predicate. Thus it is illegitimate for us to project “is a day before tomorrow or is a day on which the professor solves the New Riddle.”

So the stranger's initial argument is invalid. The catch is now supposed to be that the account the professor has offered likewise resolves the New Riddle. Here's how. Letting  $t$  be some time in the future, the predicate “grue” is said to apply to things that are examined before  $t$  just in case they are green, and to other things just in case they are blue. All emeralds we have examined so far have been grue: they have been examined before  $t$  and they have been green. Thus, our problem is to say why we *cannot* legitimately infer that *all* emeralds, even the unsampled ones, are grue. The professor's account allegedly solves that for us. Our evidence that the sampled emeralds are grue is *dependent on* our evidence that the sampled emeralds are examined. For, as Freitag says, “we would not believe the samples to be grue if we did not hold them to be examined” (p. 264). But of course, being examined is a *discriminating predicate* – indeed, it is the most generic discriminating predicate  $\Delta$ ! So “grue” is a disjunctive weakening of a discriminating predicate, and our evidence that the sampled emeralds are grue is dependent on our evidence that they are  $\Delta$ . And this is exactly the kind of case the professor's account was designed to rule out. Inferences like this are invalid.

Conversely, the corresponding inference using “green” in place of “grue” is not ruled out. All sampled emeralds have been green. And our evidence that they have been green is not dependent on our evidence that they satisfy some other discriminating predicate. In particular, it is not dependent on our evidence that they have been examined: the sampled emeralds would still have been green even if they hadn't been examined. So there is nothing wrong with projecting the predicate “green” in this context.

This completes my review of Freitag's proposal for solving the New Riddle. In the next section I argue that it does not work.

### 3. The Problem with Freitag's Proposal

I previously mentioned that Freitag's proposal implicitly makes the incorrect assumption that is fair game in this dialectic to appeal to claims about counterfactual dependence. Let me illustrate what I mean.

Whenever someone proposes a solution to the New Riddle, i.e. an argument showing why the green induction is valid but the grue induction is not, it is prudent to ask whether a parallel solution could be offered by a grue-speaker in favor of the opposite conclusion. Consider, for example, how Goodman rebuts the proposal that “grue” is a defective predicate because it is *positional* or *time-dependent*. The tempting thought here is that grue involves a change at a particular time, and thus we are not justified in projecting it *across* that time. The traditional response to this, which Goodman articulates in *Fact, Fiction, and Forecast* (pp. 79-80), involves appeal to another grue-like predicate, “bleen.” Something is bleen if it is examined before  $t$  and is blue, or is green otherwise. Now certainly if we start with “green” and “blue” as basic, then “grue” and “bleen” will be defined with reference to a temporal term. But the grue-speaker will

start with “grue” and “bleen” as basic, and use them to define “*green*” and “*blue*” with reference to a temporal term.<sup>4</sup> Time-dependence, then, is language relative. Thus if we appeal to it in justifying our choice of predicates, the grue-speaker will have a precisely symmetric justification open to him in favor of *his* predicates. And parallel justifications for opposite conclusions are no justifications at all. The symmetry of the situation thus renders our attempt at justification ineffectual.

Now the question I want to consider is whether the grue-speaker is going to grant Freitag's claims. There is one claim in particular that I want to focus on: the claim that our evidence that sampled emeralds are grue is dependent on our evidence that they have been examined. As Freitag puts it, “we would not believe the samples to be grue if we did not hold them to be examined.” The thought here, I take it, is that we know that emerald color typically does not depend on whether the emerald has been examined. Thus if we hadn't examined the sampled emeralds, they would still be green, and therefore *not* grue. (To still count as grue and yet be unexamined, they would have had to be blue.)

Now, Freitag is certainly right that *we* would not believe the sampled emeralds to be grue if we did not think they were examined. But does the *grue-speaker* believe that? What is at stake here is the following counterfactual:

- 1) If the sampled emeralds had not been examined, they would not have been grue.

Furthermore, to allow the green induction to go through, we also need a second counterfactual to be true:

- 2) If the sampled emeralds had not been examined, they would still have been green.

To get some traction here, let's consider how we should go about evaluating such

---

<sup>4</sup> Something is green if it is grue and examined before *t*, or is bleen otherwise; something is bleen if it is blue and examined before *t*, or grue otherwise.

counterfactuals. Currently our best semantics for counterfactuals involves appealing to possible worlds. According to standard possible worlds semantics for counterfactuals<sup>5</sup>, we evaluate a counterfactual  $P \Box \rightarrow Q$  by (i) imposing a similarity metric on the space of possible worlds, (ii) going to the nearest world where the antecedent  $P$  of the given counterfactual holds, and (iii) checking whether the consequent  $Q$  also holds there. The counterfactual is true iff  $Q$  holds in the closest  $P$ -world.

Now suppose we use such a semantics to evaluate counterfactuals (1) and (2). So we impose a similarity metric on the space of possible worlds, and we go to the nearest world in which the emeralds were not examined. And our question is: what color are the emeralds there? If they are green, then they are unexamined and green, so they are not grue. If they are blue, then they are unexamined and blue, so they *are* grue. Clearly, then, whether (1) and (2) hold is going to depend on the similarity metric that we impose on the space of possible worlds. On our traditional understanding of similarity, the closest (most similar) world where the sampled emeralds were not examined will be one in which they are green and not blue (and, therefore, they are not grue). Surely, however, a grue-speaker will have exactly the opposite conception of similarity. After all, he thinks grue things all “look alike,” so it is only natural that his conception of similarity would reflect that. According to *his* conception, then, the most similar world in which the sampled emeralds were not examined is one in which they are blue and not green (and, therefore, they *are* grue). Thus the grue-speaker will quite naturally *deny* counterfactuals (1) and (2). Instead he will affirm the following counterfactuals:

3) If the sampled emeralds had not been examined, they would not have been green.

4) If the sampled emeralds had not been examined, they would have been grue.

---

5 As developed in, e.g., Lewis (1973).

And that will lead him to conclude that his evidence that the sampled emeralds are all grue is *not* dependent on the fact that they have been examined. Thus, according to him, the projection of “grue” is perfectly legitimate here. That is, the grue-speaker may perfectly well agree with Freitag's claims about the conditions under which a projection is derivatively defeated. He simply *disagrees* that the projection of “grue” meets those conditions. And that is because he believes that different counterfactuals are true.

In short, the situation here is exactly analogous to the situation with the time-dependence proposal we reviewed above. Counterfactual dependence, just like time-dependence, is language relative. Thus if we appeal to it in justifying our choice of predicates, the grue-speaker will have a precisely symmetric justification open to him in favor of *his* predicates. Again, the symmetry of the situation renders our attempt at justification ineffectual.<sup>6</sup>

#### 4. Conclusion

Appeals to counterfactuals like (1) and (2) are fairly common in the history of proposed solutions to the New Riddle. Jackson (1975) articulates probably the most famous counterfactual-based solution. Other noteworthy counterfactual-based accounts include Jackson and Pargetter (1980), Godfrey-Smith (2003), and Schramm (2014). There are particular problems unique to each of these proposals, but they all share the same general problem: they beg the question against the grue-speaker by assuming counterfactuals that he would reject. The fact that appeals to counterfactuals are question-begging has been emphasized by authors such as Roskies (2008) and Dorst (2016), but the continued appearance of proposals that rely on such appeals indicates

---

<sup>6</sup> It is of course tempting to protest: “No, counterfactuals (1) and (2) are obviously true, whereas counterfactuals (3) and (4) are obviously false.” I agree. The problem, of course, is that simply stating this is not going to convince the grue-speaker (which, after all, was the goal all along). He will simply turn around and affirm the obviousness of (3) and (4) rather than (1) and (2).

that the message is not being taken up. So by way of concluding, let me just emphasize the fact that any *genuine* solution to the New Riddle should not be subject to what we might call “gruification.” That is, it must not be possible for a grue-speaker to construct a parallel argument in favor of his own conclusion. It is thus not enough to justify the green induction over the grue induction from within our own “conceptual scheme,” as it were. One must do so in a way that also justifies the green induction over the grue induction in the *grue-speaker's* conceptual scheme. This is what makes the problem so challenging: we have to show the grue-speaker that *by his own lights* he is being irrational. Freitag has not done that in his proposal.

All that being said, I do not mean to suggest that Freitag's paper isn't worthy of attention. As I said above, it is an immensely entertaining piece. The problem is only that the entire dialogue could be rewritten as taking place between two grue-speakers, who in the end reach the conclusion that the *grue* induction, not the green one, is valid.

## References

- Dorst, Chris. (2016). "Evidence, Significance, and Counterfactuals: Schramm on the New Riddle of Induction." *Erkenntnis*. 81: 143-154.
- Freitag, Wolfgang. (2015). "I Bet You'll Solve Goodman's Riddle." *Philosophical Quarterly*. 65: 254-267.
- Godfrey-Smith, Peter. (2003). "Goodman's Problem and Scientific Methodology." *Journal of Philosophy*. 100: 573-590.
- Goodman, Nelson. (1983). *Fact, Fiction, and Forecast* (4th ed.). Cambridge: Harvard University Press.
- Jackson, Frank. (1975). "Grue." *Journal of Philosophy*. 72: 113-131.
- Jackson, Frank and Pargetter, Robert. (1980). "Confirmation and the Nomological." *Canadian Journal of Philosophy*. 10: 415-428.
- Lewis, David. (1973). *Counterfactuals*. Malden: Blackwell.
- Roskies, Adina. (2008). "Robustness and the New Riddle Revived." *Ratio*. 21: 218-230.
- Schramm, Alfred. (2014). "Evidence, Hypothesis, and Grue." *Erkenntnis*. 79: 571-591.