


DISCUSSION NOTE

# Henderson on Meta-Induction and the Problem of Induction

Gerhard Schurz 

Department of Philosophy, Heinrich Heine University Duesseldorf, Duesseldorf, Germany.  
Email: [schurz@hhu.de](mailto:schurz@hhu.de)

(Received 06 May 2023; revised 21 July 2023; accepted 03 August 2023; first published online 09 August 2023)

## Abstract

This note presents a short reply to Henderson’s critical discussion of Schurz’s approach to the problem of induction based on the optimality of meta-induction. Henderson objects that the meta-inductive a posteriori justification of object-induction rests on a certain premise—namely, an approximation condition—that she reveals as untenable. I reply that Henderson’s approximation condition is indeed too strong to be plausible, but it is not needed by the meta-inductive approach; a much weaker and highly plausible approximation condition is sufficient.

In a recent article (Henderson, 2023), Leah Henderson gives a beautiful presentation and a critical discussion of the approach of Schurz (2019) to the problem of induction based on the optimality of meta-induction. As elucidated by Henderson, Schurz’s approach consists of two steps:

1. The a priori justification of weighted meta-induction (wMI) by the mathematically provable predictive optimality of wMI in regard to all prediction methods accessible to a given epistemic agent. Thereby, wMI tracks the success records of all accessible prediction methods and predicts an optimally weighted average of them, with weights depending on the predictive successes of the methods (so-called attractivities).
2. The meta-inductive a posteriori justification of object-induction (induction at the level of events, abbreviated as OI), based on the fact that object-inductive prediction methods have so far been significantly more successful than noninductive methods. This fact justifies, by meta-induction, the use of object-inductive methods for future prediction tasks.

I agree with almost all that Henderson says until the last two pages of her article, where her major objection against the second part of the meta-inductive approach

begins. Henderson argues that the a posteriori justification of OI requires a certain premise—namely, the approximation condition Approx—which she reveals as being untenable. In the following, I outline my major reason for rebutting Henderson’s objection, which, in a condensed form, says that Henderson’s approximation condition is indeed too strong to be plausible, but it is not needed by Schurz’s approach; a much weaker and highly plausible approximation condition suffices.

Henderson’s objection starts with her reconstruction of the meta-inductive a posteriori justification of the application of OI to a given prediction situation  $s$ , which includes the observed past events and success records of the accessible methods and their predictions for the next prediction task. Henderson’s reconstruction takes the form of the following argument—for better understanding, I have rewritten premise 2 in a more explicit form:

Premise 1 (optimality result): wMI is a priori justified.

Premise 2 (empirical premise): Until the present prediction situation  $s^*$ , OI has been more predictively successful than every alternative accessible method, to such a degree that the weight that wMI attributes to OI is so close to 100% that wMI’s prediction in  $s^*$  approximately coincides with that of OI.

“Premise” 3 (superfluous, follows from premise 2): Application of OI in  $s^*$  yields approximately the same prediction as wMI.

Premise 4 (approximation condition Approx): If the application of a method  $M_1$  to a situation  $s$  is justified and the application of a method  $M_2$  to situation  $s$  yields approximately the same result (i.e., prediction) as  $M_1$ , then the application of  $M_2$  to situation  $s$  is justified, too.

Conclusion: Application of OI to  $s^*$  is justified, too.

Note that “premise” 3 in my presentation of Henderson’s argument is superfluous because it follows from premise 2; this “premise” should better be understood as an intermediate reasoning step. I now turn to my major critique of Henderson’s objection—namely, the condition Approx. First of all, Henderson’s condition Approx seems to be incompletely formulated: it is not enough that the prediction of method  $M_2$  in situation  $s^*$  is de facto approximately identical with that of  $M_1$  in  $s^*$ ; we must also be *justified* in believing this, based on the relevant evidence. Is this amended formulation of the condition Approx plausible? In her major objection, Henderson argues that this is not the case because  $M_2$  could be an “otherwise crazy” method whose prediction in situation  $s^*$  happens to coincide with that of  $M_1$  by sheer accident or luck. In such a case, we should not infer that the application of the crazy method  $M_2$  to  $s^*$  is justified.

To illustrate Henderson’s point by a variation of her example, assume  $M_1$  is a justified weather-forecasting method (possibly but not necessarily a meta-inductive method) that predicts whether it will rain the next day or not, and  $M_2$  is a “crazy” method that bases its weather prediction on flipping a coin. Assuming statistical independence, the chance that  $M_2$ ’s prediction coincides with that of  $M_1$  is 50%, but should we therefore say that every second time, on average, the application of the coin-flipping method to weather forecasting is justified? Henderson says no, and I agree with her. But consider how we could be *justified* in believing that the coin-flipping method  $M_2$  will make the same prediction as  $M_1$  in a given situation: only by observing  $M_2$ ’s prediction and checking whether it agrees with that of  $M_1$ ; if it does,

we predict it, and otherwise, we don't. Obviously, this is a way of cheating: we cannot say that this procedure is really an application of method  $M_2$ . Rather, it is an application of the meta-method "observe  $M_2$ 's prediction and check whether it coincides with that of a justified method and if it does, predict it." Our decision to apply method  $M_2$  in a given situation  $s$  must be made *before*  $M_2$  delivers its prediction. So, our reasons for applying  $M_2$  to situation  $s$  also must not include  $M_2$ 's prediction; these reasons should only include our conceptual and observational evidence about method  $M_2$  before it delivers its prediction. More specifically, this evidence should only include (i) the method's track record until the given situation  $s$  and (ii) the method's algorithm together with the algorithm's input information in situation  $s$ , in the case that the method is given by an input-dependent algorithm rather than by an external agent (methods given by an external agent do not possess an accessible algorithm but just deliver predictions). We call this evidence the *justificatory-relevant* evidence about the method in situation  $s$ . Moreover, the conclusion inferred from this evidence concerning the approximate agreement of the predictions of two methods,  $M_1$  and  $M_2$ , must not presuppose OI because otherwise, the justification of OI by the approximation condition would become circular; thus, this inferred conclusion must be *entailed* by the evidence. In conclusion, I propose reformulating the approximation condition as follows:

Approximation condition Approx\*: For all situations  $s$  and prediction methods  $M_1$  and  $M_2$ : if the application of  $M_1$  to a situation  $s$  is justified and the justificatory-relevant evidence about  $M_1$  and  $M_2$  entails that the result of this application, that is, the prediction of  $M_1$  in  $s$ , is approximately identical with the prediction of  $M_2$  in  $s$ , then the application of  $M_2$  to situation  $s$  is justified, too.

Condition Approx\* is weaker than Approx because its antecedent is stronger. More importantly, condition Approx\* avoids the accidentality objection of Henderson because the justificatory-relevant evidence about the method excludes accidental features. If method  $M_2$  generates its predictions in a random way, then the justificatory-relevant evidence about  $M_2$  and  $M_1$  cannot entail information about the closeness of their predictions in the given situation  $s$ ; thus, the application of  $M_2$  to  $s$  cannot be justified by the agreement of  $M_2$ 's prediction with that of  $M_1$  in  $s$ . The same is true if  $M_2$  is a different kind of "crazy" method—for example, the method of a purported clairvoyant whose predictions, like that of the random method, sometimes agree with that of  $M_1$  and sometimes don't. Only if method  $M_2$  is given by a deterministic algorithm whose application to  $M_2$ 's input in situation  $s$  will demonstrably approximate the prediction of  $M_1$  in  $s$ , will the condition Approx\* be satisfied. In accordance with Henderson, I argue that in this case, the justification indeed transfers from the application of  $M_1$  to that of  $M_2$  because now this transfer is built solely upon nonaccidental features of the two methods, as illustrated in the third-to-last paragraph of Henderson's section 6. For example,  $M_1$  could be a complicated extrapolation function and  $M_2$  a simpler function whose output (prediction) demonstrably agrees with that of  $M_1$  in application to the particular kind of input data in situation  $s$ . In this case, we can legitimately say that because of the justification of  $M_1$  and the demonstrable agreement of  $M_2$ 's prediction with that of  $M_1$  in situation  $s$ , we are justified in applying the simpler method  $M_2$  instead of the more complicated method  $M_1$ ; cases of this sort occur frequently in science.

In conclusion, the weakened condition Approx\* holds, and the meta-inductive a posteriori justification of OI can be based on it, so the latter justification is established. Note, however, that condition Approx\* is still more general than really needed for the meta-inductive a posteriori justification of OI. In application to this context,  $M_1$  is not an arbitrary justified method, but the method wMI, and the justificatory-relevant evidence about wMI in a situation  $s$  consists of wMI's weighted prediction algorithm together with the success records of all accessible methods that determine their weights. In this context, the justificatory-relevant evidence about wMI already includes the track record of method  $M_2$  because  $M_2$  is supposed to be accessible to wMI. More importantly, the relevant evidence from which we can deductively infer that  $M_2$ 's prediction in situation  $s$  must approximately coincide with that of wMI—without knowing  $M_2$ 's prediction—is given by the fact that wMI attributes to  $M_2$  a weight close to 100%, so wMI's prediction must approximately coincide with that of  $M_2$ . If this fact obtains, then  $M_2$  is a posteriori justified, and according to the empirical premise 2, this fact obtains for  $M_2 = \text{OI}$ . In conclusion, the following still-weaker approximation condition Approx\*\* suffices to warrant the meta-inductive a posteriori justification of OI:

Approx\*\*: For all situations  $s$  and prediction methods  $M_1$  and  $M_2$ : if  $M_1$  is a justified wMI method, and the weight that  $M_1$  attributes to  $M_2$  in situation  $s$  is close to 100%, then the application of  $M_2$  to  $s$  is justified, too.

Condition Approx\*\* is a special instance of Approx\* and at least as successful in the refutation of Henderson's accidentality objection as Approx\*.

Only a tough strengthening of the accidentality objection could be launched against the weakened conditions Approx\* or Approx\*\*. Although I doubt that Henderson would go for this strengthening, let me discuss it for the sake of completeness. This tough objection could criticize that even if method  $M_2$  had a superior success record in the past,  $M_2$  could nevertheless be some version of random guessing whose superior track record was based on an extremely improbable "cosmic" amount of luck. Because this possibility cannot be excluded, the meta-inductive a posteriori justification of method  $M_2$  fails. But arguing in this way would overshoot the mark. That the outcomes of a method or mechanism are *randomly* distributed is, of course, a general *inductive* hypothesis about the statistical propensity or frequency limit of this method or mechanism. Because the epistemological justification of inductive prediction methods is at stake, no inductive hypotheses may be presupposed in the success evaluation of prediction methods, on pains of avoiding circularity. The success evaluation may only be based on the observed track record. And given that the track record of the method  $M_2$  in question is indeed superior and is based on a sufficiently large sample of predictive tests (or rounds of the prediction game), the hypothesis that this method is reliable and will be successful in the future is a posteriori justified—namely, justified by meta-induction. For illustration, assume that a person has successfully predicted tomorrow's weather (rainfall or not) by flipping a coin a hundred times with an average success of 100%. The likelihood of such a track record if it were obtained by a variant of random guessing is astronomically low ( $1/2^{100} > 1/10^{30}$ ). Therefore, people would consider this method as

a kind of “magic” coin flipping that—for reasons no one understands—yields trustworthy weather predictions, in accordance with the meta-inductive evaluation. This example reveals an important feature of meta-induction—namely, its undogmatic nature and openness to all kinds of epistemic possibilities. No background theory, even if it is extremely well confirmed, is immune against revision by a sufficiently large amount of conflicting evidence—in our example, this conflicting evidence is the track record of the method  $M_2$  that is prima facie considered as random or “otherwise crazy,” but after observation of its superior track record, it is meta-inductively evaluated as predictively trustworthy, at least for the time being.

This concludes my defense of the meta-inductive a posteriori justification of OI against Henderson’s objection. I conclude this reply with two further brief amendments to Henderson’s presentation. *First*, as remarked earlier, the a posteriori justification of OI is based on the empirical premise that until now, OI was more successful than any noninductive prediction method. Referring to Sterkenburg, Henderson writes (two paragraphs before her reconstruction of the argument) that what would really be needed is a stronger empirical premise—namely, that OI was by far *much* more successful than any other noninductive method—because only then can the weight of OI (attributed by wMI) come close to 100%. But this is not precisely true, for the reason that attractivity weights are historically *cumulative* and amplify sustainable success differences if time goes on. Even a small surplus average success of OI (compared to competitor methods) will produce a much higher weight of OI if this surplus success persists for a sufficiently long time. More precisely, let “ $w_n(M)$ ” designate the weight of method  $M$  at time  $n$  and “ $suc_n(M)$ ” designate the average success of method  $M$  until time  $n$ . Then, if  $suc_n(OI) \geq suc_n(M) + \varepsilon$  (for some  $\varepsilon > 0$ ) holds for all  $n \geq k$  (for some  $k \geq 0$ ), then for all  $n \geq k$ , the weight ratio  $w_n(OI)/w_n(M)$  increases with increasing  $n$ , and it approaches infinity for  $n \rightarrow \infty$ . In particular, if  $suc_n(OI) \geq suc_n(M) + \varepsilon$  holds for all competitor methods  $M$  of OI in the candidate pool,  $w_n(OI)$  will converge to 1 for  $n \rightarrow \infty$ . This is true for linear attractivity weights—defined as  $lin-w_n(M) = \max(suc_n(M) - suc_n(wMI), 0)$  (cf. Schurz 2019, 140)—as well as for the (more efficient) exponential weights  $exp-w_n(M) = \exp(\eta \cdot suc_n(M))$  (see Schurz 2019, 144).

*Second*, referring to Sterkenburg, Henderson remarks in the paragraph following her reconstruction of the argument that given the empirical premise, meta-induction provides a justification of OI merely “at this point in time.” But this comment of Sterkenburg is ambiguous. In one interpretation, it means that this justification is relative to the *present* evidence and could be overruled by future evidence. In this interpretation, Sterkenburg’s comment merely accentuates the a posteriori nature of the meta-inductive justification of OI, which is correct but not really a point of criticism. In this respect, the meta-inductive justification of OI is similar to the confirmation of *scientific theories*, which, too, is always relative to the actual evidence and can be overthrown by future evidence.

In an alternative interpretation, Sterkenburg’s comment says that all meta-induction can justify a posteriori is object-inductive predictions for the *next* time point but not in application to the more distant future. However, it is easily possible to extend the task of prediction games to the prediction of events in the more distant future or to the prediction of finite sequences or samples of events (Schurz 2019,

sec. 7.4). It is even possible to include post facto “predictions” that predict past events by events lying further in the past, provided the “predicted” events are *use novel* in the sense that they have not been used in the construction of the predictive method or theory (cf. Schurz 2024, sec. 6.5.2, 7.2). Examples of this sort are scientific predictions of ice ages or shifts of planetary orbits that have been meta-inductively trained by post facto predictions and are then applied to predictions of the distant future. I think that the meta-inductive justification strategy can also work for inductive generalizations, although the elaboration of this claim is beyond the scope of this short note.

**Acknowledgments.** For helpful communication, I am indebted to Leah Henderson, Igor Douven, Tom Sterkenburg, Peter Brössel, Jochen Briesen, and Daniel Herrmann.

## References

- Henderson, Leah. 2023. “Crossing Levels: Meta-Induction and the Problem of Induction.” *Philosophy of Science* 90 (4): 985–993. <https://doi.org/10.1017/psa.2023.2>.
- Schurz, Gerhard. 2019. *Hume’s Problem Solved: The Optimality of Meta-Induction*. Cambridge, MA: MIT Press.
- Schurz, Gerhard. 2024. *Optimality Justifications: New Foundations for Epistemology*. Oxford: Oxford University Press.