Philosophy of Science (2023), **90**, 985–993
doi:10.1017/psa.2023.2

PSA PHILOSOPHY OF SCIENCE ASSOCIATION

PHILOSOPHY OF SCIENCE

DISCUSSION NOTE

Crossing Levels: Meta-induction and the Problem of Induction

Leah Henderson

Department of Theoretical Philosophy, University of Groningen, Oude Boteringestraat 52, 9712GL Groningen, The Netherlands
Email: l.henderson@rug.nl

(Received 30 November 2022; revised 09 December 2022; accepted 03 January 2023; first published online 17 February 2023)

Abstract

Gerhard Schurz claims to have a solution to Hume's problem of induction based on results from machine learning concerning meta-induction. His argument has two steps. The first is to establish a justification for following a certain meta-inductive strategy based on its predictive optimality. The second step is to show how this justification can be transferred to object-induction. I unpack the second step and fail to find a convincing argument supporting the transfer of justification from meta-induction to object-induction. My conclusion is that the problem of induction has not yet been solved by appeal to meta-induction.

1. Introduction

Gerhard Schurz claims to have a solution to Hume's problem of induction (Schurz 2008, 2017, 2019). This is based on results from machine learning concerning meta-induction. The problem of induction traditionally concerns "object-induction," which is induction applied at the 'object-level' to ordinary events—for example, when we infer from the fact that the sun has always been observed to rise to the prediction that it will rise tomorrow. "Meta-induction," on the other hand, is the method of applying induction at the "meta-level" to the success rates of all accessible prediction methods. It predicts some combination of the predictions of those methods (both inductive and non-inductive) that were most successful in the past (Schurz 2019, 8–9).

Schurz's solution comes in two stages. The first stage is to establish a justification for following a certain meta-inductive strategy based on its predictive optimality. The second stage is to show how this justification can be transferred to object-induction. The second step of the argument is explicitly a posteriori, but Schurz claims that it nonetheless avoids the circularity that Hume was concerned with. Thus, Schurz's solution promises a subtle side-stepping of the notorious dilemma which Hume posed for justifications of induction.

The aim of this note is to clarify the second stage of Schurz's argument, and in particular exactly what conclusion can be drawn from it. Schurz presents his solution as providing a justification for the strategy of following object-induction. His approach has been criticized on the grounds that it does not provide a general justification for following object-induction as a method, but only for following the method of object-induction on the next time-step (Sterkenburg 2020). I argue that Schurz's argument provides no justification for the application of object-induction at all, whether in general, or restricted to the next time step.

2. Hume's problem

Hume asked for the reasoning behind the ready inferences that we typically make from past observed behavior to expectations about the future. For example, we draw inferences such as:

> All observed instances of bread (of a particular appearance) have been nourishing.
>
> The next instance of bread (of that appearance) will be nourishing.

Hume seeks a "chain of reasoning," which provides the foundation for carrying out such an inference. This reasoning, he says, would have to be based on a principle, often called the "principle of uniformity": "that instances, of which we have had no experience, must resemble those, of which we have had experience, and that the course of nature continues always uniformly the same (Hume 1739, 1.3.6.4)."

Hume's argument against the possibility of providing the foundation for inductive inferences goes as follows. Hume claims that there are two types of possible arguments: "demonstrative" arguments and "probable" arguments. Demonstrative arguments establish conclusions which cannot be conceived to be false, whereas probable arguments concern "fact and existence." A natural way to understand the distinction in modern terms is along the lines given by Samir Okasha: "A demonstrative argument, for Hume, has a priori premises and is deductively valid, while a probable argument has one or more empirical premises (Okasha 2001, 313)."

Hume then presents a dilemma. First, the principle of uniformity cannot be supported by a demonstrative argument, because there is no contradiction in supposing that the course of nature might change. And secondly, it cannot be supported by probable reasoning, because this would be circular. All probable arguments "proceed upon the supposition that the future will be conformable to the past (Hume 1748, 4.2.19)," and supposing this would be to "take for granted, which is the very point in question (Hume 1748, 4.2.19; see also Hume 1739, 1.3.6.7/90)." Since neither demonstrative nor probable arguments will serve, the preliminary conclusion PC is that:

> there is no chain of reasoning from the premises to the conclusion of the inference. (PC)

One can then add the Humean justification condition HJC:

> If there is no chain of reasoning from the premises to the conclusion of the inference, the inference is not justified. (HJC)

From PC and HJC, one can draw the conclusion that the inductive inference is not justified.[1]

3. What needs to be justified?

Hume clearly intended to raise a problem arising for a broad class of inferences. How should we characterize this class? One of the simplest suggestions is that the inferences results from application of a rule such as "simple enumerative induction":

All observed instances of A have been B.

The next instance of A will be B.

This may be generalized in a probabilistic setting to a rule such as what Reichenbach calls "the principle of induction," or "straight rule" (Reichenbach 1938):

If after n observations, a relative frequency of m/n is observed, the frequency to expect in the future is also m/n.

More sophisticated inductive methods can also be defined (Schurz 2019, section 5.9).

Following Schurz, we can define "object-induction" (abbreviated as OI)" to mean "all methods of induction that are applied at the level of events – the object-level" (Schurz 2019, 83). The general Humean problem of induction may be then seen as the problem of justifying the class of all inferences resulting from the application of object-induction methods. This allows for a reformulation of what the solution to Hume's problem requires. Rather than looking for specific "chains of reasoning" undergirding the inferences themselves, we can look to properties of the methods which make those methods justified. Inferences then acquire justification from the methods via a principle such as:

an inference *Inf* is justified when it is produced by an application of a method *M* and the method *M* is justified. (Inf-M)

4. Justifying a Method

What then does it mean to justify a method? According to the most straightforward way of reformulating Hume's problem, we could adopt the criterion that the method giving rise to the inductive inferences is justified if it is *reliable*. To be reliable, the method must mostly reach true conclusions, not only in the past, but also in the future. But then Hume's argument can be reprised: on the one hand, the reliability of a method cannot be established by a demonstrative argument, since it is quite conceivable that a method which has hitherto proved reliable stops being so tomorrow. On the other hand, we cannot establish the reliability by an argument

[1] This is a common way to understand Hume's line of argument, though it has been debated to what extent Hume intended to argue for inductive skepticism, and therefore whether he actually relied on a justification condition like HJC (see Henderson 2018/2022, sec. 2). The exact intentions of Hume are not so important for our purposes here.

https://doi.org/10.1017/psa.2023.2 Published online by Cambridge University Press

from the empirical premise that the method has hitherto been highly successful because this would involve extrapolating from the past successes of the method to its probable future successes, which itself requires some kind of principle of uniformity. Thus such an argument would be circular.[2]

Among authors who take Hume's argument to be decisive against the possibility of showing that the inductive method is reliable, another strategy has been to argue that there are weaker conditions which can justify a method. Reichenbach argued, for example, that a sufficient condition for a method to be justified is that that method is a necessary condition of success:

If M is a necessary condition for success, then M is justified. (RJC)

"Success" here is defined as meeting the aim of induction. Reichenbach gives the following example as an analogy:

> A man may be suffering from a grave disease; the physician tells us: "I do not know whether an operation will save the man, but if there *is* any remedy, it is an operation". In such a case, the operation would be justified. Of course, it would be better to know that the operation will save the man; but, if we do not know this, the knowledge formulated in the statement of the physician is a sufficient justification. (Reichenbach 1938, 349)

Thus, for Reichenbach an alternative formulation of the statement that "M is a necessary condition for success" is "if any method will work (succeed), it will be M." This allows that there may be cases where no method will work, including M.[3]

Schurz proposes a justification criterion SJC that he sees as having its roots in Reichenbach's approach (Schurz 2019). The justification of a method is to be based, not on its reliability, but on its optimality—on its being the best available alternative. We are to be justified in deploying a method if it is the best one we've got, even if we have no grounds for confidence that it will actually deliver. The justification criterion for the method becomes:

If M is predictively optimal, then M is justified. (SJC)

More precisely, the notion of optimality is defined in relation to a class of "prediction games." A prediction game G=((e),Π) is a pair consisting of a stream of events (e)=(e_1,e_2 ...) and a set of prediction methods Π (called "players"). Then, "a method M^* is called optimal in relation to a class of prediction games iff for all streams of events (e), M^* is at least as good as M for all other $M \in \Pi$ (Schurz 2019, 95)."

[2] Some have argued that in this form, the type of circularity involved is not so pernicious since it is "rule-circularity" rather than "premise-circularity" (e.g. Cleve 1984).

[3] In order to get a full account, of course, more needs to be said about what is meant for a method to have 'success' or to 'work'. See Henderson (2018/2022, sec. 5.3) for more on how Reichenbach defines these notions.

The performance here is measured in terms of predictive success rates. Each prediction method in the set Π has a certain success rate, which is a cumulative score for how well that method's predictions have panned out.[4] "M^* is at least as good as M" can be defined in a long-run sense—meaning that in the long run as the number of events observed becomes very large, M^* always achieves a higher cumulative success rate than M. It can also be defined in a more short-run sense according to which other methods can do better than M^* in the short run, but the amount by which they do so converges quite quickly to zero as the number of events observed increases. In either sense of optimality then, what is to justify the method is its performance, relative to other available methods in a relevant class, over a long run of data for all possible data-streams. The justification for performing a particular inductive inference remains parasitic, via Inf-M, on the justification of the method M. Now however, the justification of the method is in terms of its optimality rather than its reliability.

5. Schurz's two-step justification of induction

In the first step of his argument, Schurz appeals to results from the regret-based learning framework of Cesa-Bianchi and Lugosi that show that there is a meta-inductive strategy that is predictively optimal among all predictive methods that are accessible to an epistemic agent (Cesa-Bianchi and Lugosi 2006; Schurz 2008, 2017). In general terms, "meta-induction" is the method of applying induction at the meta-level to the success rates of all accessible prediction methods. One kind of meta-inductive strategy is simply to imitate the most successful prediction method available. More sophisticated meta-inductive methods will produce a combination of available prediction methods weighted according to their past success. For example, the meta-inductive strategy "wMI" predicts a weighted average of the predictions of the accessible methods, where the weights are "attractivities," which measure the difference between the method's own success rate and the success rate of wMI. The main result is that the wMI strategy is long-run optimal in the sense that it converges to the maximum success rate of the accessible prediction methods. Worst-case bounds for short-run performance can also be derived. The optimality results form the basis for an a priori justification for the use of wMI. Namely, the thought is, it is justified to use wMI because it satisfies the justification criterion SJC.

The first stage of Schurz's argument can be reconstructed as follows:

1. If method M is predictively optimal, method M is justified. (SJC)
2. wMI is predictively optimal. (mathematical result)
3. wMI is justified

The second step of Schurz's argument is to show how the justification can be transferred from meta-induction to object-induction. Schurz claims that the a priori justification of wMI, given in the first step, together with the contingent fact that inductive methods have so far been much more successful than non-inductive methods, gives rise to an a posteriori justification of induction which is non-circular. "We know by experience that in our world, object-inductive prediction methods have

[4] With respect to some loss function.

https://doi.org/10.1017/psa.2023.2 Published online by Cambridge University Press

been more successful than noninductive methods so far, whence it is *meta-inductively justified* to favor object-inductivistic strategies in the future (my emphasis, Schurz 2019, 85)."

The reasoning here is very compressed. In the next section, we explore how exactly to unpack the second stage of Schurz's argument.

6. Unpacking the argument of the second step

What exactly is meant here by "meta-inductively justified?" The first point to notice is that following object-inductive strategies is not justified according to the criterion SJC, since object-induction has not been shown to be predictively optimal. The result we have is that there are meta-inductive strategies which are predictively optimal. OI and MI are not the same strategy, and in general their predictions may not coincide. The meta-inductivist using wMI does not assign all the weight to the most successful method, but spreads it among the other competing methods. If the competitors are not very successful, this may mean that in practice wMI is fairly well-approximated by the most successful object-method. We may further observe that the most successful object-method up to now has been OI. But wMI remains a different method than OI. It can result in different predictions on the very next application if OI stops working and a non-inductive method does better. In that case, wMI would begin to favor the non-inductive method, and wMI would start to come apart from OI. Thus we have no argument that OI has the same longer-term properties that produce the optimality of wMI. We know by following meta-induction we are *en route* to doing as well as we can do predictively, but we have no similar guarantees for OI. Thus, we do not have an argument yet that we are justified in favoring object-inductivistic strategies in the future.

The argument for transferring justification to OI must be somewhat more elaborate. In fact, Schurz does not claim that OI is justified in exactly the same way as wMI. For one thing, the argument for OI is a posteriori, since it is based on an empirical premise EP:

> object-inductive prediction methods have been more successful than noninductive methods so far. (EP)

Given the empirical premise, the claim is that there is a non-circular argument for OI. The argument is non-circular because it depends only on steps of a priori reasoning once the empirical premise is given.

It looks as though the basis for transferring the justification from wMI to OI is the close approximation between OI and wMI. The methods are expected to produce approximately the same predictions. However, the empirical premise Schurz invokes—namely EP—does not in fact ensure such an approximation. If OI has been more successful than noninductive methods, that only means that OI should be weighted highly in the combination of methods that wMI produces on the basis of the past experience, not necessarily that it should produce a close approximation of wMI.

Sterkenburg suggests that one could invoke a stronger empirical premise EP* in order to ensure that the approximation holds:

> "As a matter of empirical fact, the strategy OI has been so much more successful than its competitors, that the meta-inductivist attibutes it such a large share of the total weight that its prediction (approximately) coincides with OI's prediction." (EP*) (Sterkenburg 2020, 538)

This is a fairly strong claim about what has actually been empirically shown so far. It requires not just that OI has proved more successful than other methods, but it has proved so much more successful that it gains the lion's share of the weight of the meta-inductive method.

Now let us see how the argument could be formulated using EP*. Let a "prediction situation s" include the observed past events, a candidate pool of methods, their track records and their actual predictions. Suppose the current situation s* is one in which EP* holds. Then we could use EP* to justify a claim that "application of OI in s* yields approximately the same result as wMI." We also need to add a premise ("Approx") to the effect that close approximation of methods allows for transfer of justification from one method to the other.

The argument would then go as follows[5]:

1. wMI is justified (from stage 1)
2. EP*
3. Application of OI in s* yields approximately the same result as wMI (given EP*)
4. 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 as M_1, then the application of M_2 to situation s is justified too. (Approx)

Conclusion: Application of OI in s* is justified too.

Sterkenburg allows this conclusion, given the assumption of EP*. However, he points out that it is not equivalent to Schurz's stated conclusion that "it is meta-inductively justified to favor object-inductivistic strategies in the future" (Schurz 2019, 85). Rather, Sterkenburg says, the argument gives "justification for sticking to strategy OI for now"—in other words, in situation s*. That is, it provides only a justification for following the prediction of the strategy OI "at this point in time (Sterkenburg 2020, 539)." However, he takes it that "this would still be an important result," since the problem of induction is important insofar as it it concerns whether we should follow scientific induction or its alternatives at the current moment (Sterkenburg 2020, 539).

However, in my view, the argument formulated above fails, and hence even the more limited conclusion cannot be drawn. The problem is that the premise Approx is not true. Suppose M_1 is the method of painstakingly applying careful scientific reasoning, and M_2 is the method of making a prediction by flipping a coin. It is possible that the application of M_2 to a situation yields approximately the same prediction as M_1, but we would not say in this case that the application of the coin-flipping method to the situation is itself justified. It is generally thought that one of the functions of justification is as a condition that enforces a certain kind of non-accidentality. It is an anti-luck requirement. Thus, if it simply happens that an otherwise crazy method happens to

[5] This formulation is informed by personal communication with Gerhard Schurz.

approximate a good method at a particular point, that in itself provides no reason to regard the crazy method as justified. The approximation is itself an accident, so you just got lucky by following the crazy method.

Is there a way to fix the argument to exclude these kinds of accidental approximation cases? It seems as though justification could transfer from M_1 to M_2, if we have a good reason for expecting M_2 to approximate M_1. Perhaps we know that the working of M_2 is relevantly similar to M_1, or even for example that they share some common way of working. Perhaps, for example, M_2 is a more coarse-grained way of making predictions, which omits some of the details that M_1 takes into account, but still achieves decent results. In such a case, it seems that we might allow that the fact that M_1 closely approximates M_2 is a reason to allow the justification for M_1 to transfer to M_2. However, the transfer of justification in this way would be one-way. The detailed working of the more fine-grained method would not be justified in any way by its coincidence with the coarse-grained method. The transfer of justification would only occur when you know that the reasons for the success of the first method are likely to be the same for the second method.

This is a condition which does not hold in the case where M_1 is wMI and M_2 is OI. The reasons for the success of wMI are its ability to make use of the successes of object-level methods such as OI. But the converse does not hold. The reasons for the success of OI, if it is successful, are something to do with the way that the world is structured, perhaps its uniformity in certain respects. Thus, although we might want to say that wMI could inherit justification from justification for OI, we would not want to put this the other way round. OI is not justified on the basis of an approximation to wMI, because that approximation does not come about because of the way that OI works in relation to wMI.

Furthermore, attempting to justify Approx as an empirical premise by appeal to scientific practice will also not work. Such a justification would be based on arguing that success in scientific practice is based in part on Approx, and then inductively projecting that success to the future. However, this presupposes that the success of approximation practices in the past will extend to the future, and since this is part of "the very point in question," this reintroduces the Humean circularity concern.

Thus, it seems that a plausible formulation of an argument for transferring justification from wMI to OI is still missing. We do not appear to have an argument which allows any conclusion about justification of the method OI either in general, or its application to the specific case.

7. Conclusion

Schurz's argument that there is a solution to the problem of induction based on meta-induction has two parts. Although the first part of the argument is clear, the second is very compressed. In this paper, I have attempted to spell out what the reasoning in the second part of the argument could be, and to examine the conclusion which should be drawn. Although Schurz presents the argument as supporting the conclusion that we are justified to "favor object-inductivistic strategies in the future," my attempts at reconstruction do not support such a conclusion. The justification for meta-induction does not transfer appropriately to a justification for object-induction, even for a single time-step.

Acknowledgements. I am grateful for helpful communication with Gerhard Schurz. Thanks also to Tom Sterkenburg for helpful comments on an early draft.

Competing interests. The author declares none

References

Cesa-Bianchi, Nicolo, and Gabor Lugosi. 2006. *Prediction, Learning and Games*. Cambridge: Cambridge University Press.

Cleve, James van. 1984. "Reliability, Justification, and the Problem of Induction." *Midwest Studies In Philosophy* 9 (1): 555–67.

Henderson, Leah, 2018/2022. "The Problem of Induction." In *The Stanford Encyclopedia of Philosophy*, edited by Ed Zalta. Stanford: Stanford University Press. https://plato.stanford.edu/entries/induction-problem/

Hume, David. 1739. *A Treatise of Human Nature*. Oxford: Oxford University Press.

Hume, David. 1748. *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.

Okasha, Samir. 2001. "What Did Hume Really Show about Induction?" *The Philosophical Quarterly* 51 (204): 307–27.

Reichenbach, Hans. 1938. *Experience and Prediction: an Analysis of the Foundations and the Structure of Knowledge*. Indiana: University of Notre Dame Press.

Schurz, Gerhard. 2008. "The Meta-inductivist's Winning Strategy in the Prediction Game: A New Approach to Hume's Problem." *Philosophy of Science* 75 (3): 278–305.

Schurz, Gerhard. 2017. "Optimality Justifications: New Foundations for Foundation-oriented Epistemology." *Synthese* 195 (9): 3877–3897.

Schurz, Gerhard. 2019. *Hume's Problem Solved: the Optimality of Meta-induction*. Cambridge, MA: MIT Press.

Sterkenburg, Tom. 2020. "The Meta-inductive Justification of Induction." *Episteme* 17 (4): 519–41.

Cite this article: 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