Good listeners, wise crowds, and parasitic experts

Jan-Willem Romeijn
University of Groningen

Tom Sterkenburg
CWI Amsterdam and University of Groningen

Peter Grünwald
CWI Amsterdam and Leiden University

Abstract

This article comments on the article of Thorn and Schurz in this volume and focuses on, what we call, the problem of parasitic experts. We discuss that both meta-induction and crowd wisdom can be understood as pertaining to absolute reliability rather than comparative optimality, and we suggest that the involvement of reliability will provide a handle on this problem.

1 Introduction

The article of Thorn and Schurz considers a central idea in social epistemology. It is the idea that "two heads are better than one": the judgment of a group is somehow of better quality than the judgment of an individual, or indeed of most individuals in the group. Thorn and Schurz reveal how this idea shows up in two prima facie different areas of study. On the one hand, they consider so-called *meta-induction*, an inductive method that combines or selects from a set of given predictive systems, or predictors for short, to arrive at comparatively optimal predictions. On the other hand, they are concerned with the aggregation of individual judgments towards a, typically more reliable, group judgment, often termed the *wisdom of the crowd*. Clearly these are two instances of the same idea, namely that the opinion of a collective, however it is spelled out, is of better quality or more likely to be correct than the opinion of an individual.

To readers who are familiar with meta-induction as well as crowd wisdom, their simultaneous treatment will not come as a surprise. Meta-induction can be viewed as a form of opinion aggregation, and as such it can subsume the standard accounts of the wisdom of the crowd. Take the example that Thorn and Schurz copy from Galton, about the estimation of the weight of an ox by the audience of a live stock show. If members of the audience are predictors, then the averaging of their answers is a meta-inductive prediction method. Or else take the standard setting for Condorcet's jury theorem, in which jurors can be considered predictors so that the majority rule functions as meta-inductive method. Of course many other settings and many other procedures for aggregating the opinions of members of a collective can be conceived. The point here is that all such procedures may be viewed as meta-inductive methods. Consequently, a discussion on meta-inductive methods has direct relevance for crowd wisdom.

Thorn and Schurz focus on a particular way in which the epistemic advantage of groups plays out when the two above areas are brought into contact. One of the main insights of meta-induction is that a suitable meta-inductive method, i.e. some combination of predictive systems, will outperform a large number of the predictive systems on which it is based, and will eventually outperform all or almost all of them. To illustrate, imagine that we invite a number of experts on a contested issue and ask them for a prognosis. Then, if we are clever enough in combining their opinions, we can ourselves pose as an expert that performs better than many if not all of the experts we invited. And this is so even though we do not ourselves bring any new information to the table! Our success as expert is completely parasitic on the quality of the experts we invited, but we do better than many of those experts. Our advantage is simply that we are prepared to listen to all the opinions.

This situation becomes paradoxical when we confront this with insights about the wisdom of the crowd. One of the main insights from this area of study is that crowds are more reliable when they are larger. Consequently, if we ask our panel of experts whether they would welcome an additional expert, their response should be positive because, as a group, they will benefit from the additional input. This response should certainly be positive if the additional expert is expected by everyone to be almost as good as the best expert in the group, although opinions will of course vary among the experts as to which of them is best. The point is that our additional expert could simply be using a meta-inductive method, which takes the already available opinions of experts as input. It seems strange that our panel of experts should be welcoming someone who only reflects what they are themselves already thinking, and who does not bring anything new to the table herself. What good can such a "good listener", who otherwise has no subject specific knowledge and who is perhaps better described as a "parasitic expert", really do?

In what follows, we will comment on the two themes highlighted above. First, we consider the idea that is central to both areas of study, the idea that two heads are better than one. After some remarks about the specific results on meta-induction that Thorn and Schurz invoke, we direct attention to the difference between a comparative and an absolute reading of this idea, which Thorn and Schurz call optimality and reliability respectively. We do not contest the optimality of meta-induction, which is that groups perform better than most individuals from that group. Instead we investigate the absolute reading of the results, according to which groups are more reliable or more likely to get things right than individuals are. Such an absolute reading presupposes that, one way or another, individuals have some inclination towards the truth. Second, we investigate the paradox concerning the addition of meta-inductive predictors to a wise crowd, as it was just outlined. We argue that the paradox is to some extent concerned with the performance of the group in an absolute sense and that, insofar as this is the case, a key role must be played by the inclination of group members to the truth. Specifically, dependent on the strength of this inclination, the addition of a meta-inductive predictor to a group may be beneficial or detrimental to the speed by which the group will converge to the correct opinion.

Before we enter a more detailed discussion on these two themes, we would like to applaud and advertise the article by Thorn and Schurz. It covers important aspects of social epistemology, a discipline that itself merits full attention. The social and institutional aspects of knowledge have long eluded philosophical study, which was traditionally focused on the individual. And insofar as these aspects did become the topic of philosophical investigation, they made their entrance primarily as part of historical and sociological studies of science and knowledge (e.g., Galison, 1987; Latour, 1987; Bloor, 1976). We think that it is high time for the parallel development of an analytic understanding of the social dimension of knowledge, and we believe that the article by Thorn and Schurz provides a valuable contribution to this.

2 Meta-induction and wise crowds

This section is devoted to a discussion of the central idea behind both meta-induction and wise crowds, to do with the epistemic advantages of groups. The focus of the section lies on meta-induction but, as will be recalled, it is thereby an indirect discussion of the wise crowd as well. After some technical considerations about meta-induction, we draw attention to the difference between comparative optimality and absolute reliability. We consider a way in which high comparative performance might aspire to be an absolute standard, following ideas from algorithmic information theory. The section ends with a sketch of some assumptions that would support absolute reliability.

It is noteworthy that the ideas and results of meta-induction have a very strong resemblance to work in machine learning, specifically the research on what is known as "prediction with expert advice" or "online prediction" (see for instance Cesa-Bianchi et al., 1997; Cesa-Bianchi and Lugosi, 2006). Although the latter work is mentioned by Thorn and Schurz in one instance, the general lack of reference to this large and flourishing area is somewhat unfortunate.[1] Possibly the authors are not aware of the extent of this rather technical field. Something to that effect is suggested by the diverging terminology, e.g., their "imitate-the-best" corresponds to "follow-the-leader"

[1]The idea of designing meta-prediction methods that perform almost as well as the best predictor in some group, no matter what data are observed, has even been reinvented several times in different fields, including all the fields mentioned in the following quote from the introduction of Cesa-Bianchi et al. (1997): "We take the extreme position, as advocated by Dawid and Vovk in the theory of prequential probability, Rissanen in his theory of minimum description length and stochastic complexity, and Cover, Lempel and Ziv, Feder and others in the theory of universal prediction and data compression of individual sequences, that no assumptions whatsoever can be made about the actual sequence of outcomes that is observed".

in machine learning (see Kalai and Vempala, 2005). The lack of reference is unfortunate because the theorems of the article by Thorn and Schurz are matched by results in machine learning that both need less assumptions, and yield much better convergence rates. In particular, we need not suppose that the number of predictors is finite, and for the standard algorithms in machine learning resembling theorem 2, the divergence from the best predictor only grows by the *logarithm* of the square root $\sqrt{n}$ of the number of predictors (Cesa-Bianchi and Lugosi, 2006; Vovk, 1990). The novel and most valuable aspect of the general project that is instigated by Thorn and Schurz is, to our mind, that they transfer these insights to the philosophical domain. In similar vein, Schurz (2008) confronts the machine learning literature with the problem of induction. We believe that such cross fertilizations are praiseworthy by themselves.

The considerations of Thorn and Schurz pertain primarily to the comparative performance or optimality of the meta-inductive methods, which is indeed uncontested.[2] We would like to emphasize, though, that nothing about this comparative superiority of meta-induction will help towards the absolute reliability of its predictions. These predictions could be systematically off, for instance because group members share a systematic bias. Thorn and Schurz are careful not to conflate optimality and reliability. However, if such results are introduced in the context of the wise crowd, the distinction may be easily missed. The wisdom of the crowd seems to derive most of its appeal from illustrations that do not merely show that the crowd as a whole does better than most of its members. The salient finding is that the

[2] This is not to say that it is trivial. The meta-inductive method of "imitate-the-best", for instance, might grievously fail in simple settings, and it is not immediately clear that meta-induction in general will not suffer a similar fate. Assume, in a binary setting, that there are two rather simple-minded predictors, the first always predicting a 0 and the other always a 1, while the actual data sequence is 010101 . . .. The follow-the-leader strategy instructs us to always predict according to the predictor that is best so far, and to simply guess if there is none. Now what happens is this: for every odd digit we will choose 0 or 1 at random because both predictors have performed equally well at that point, and for every even digit we will choose – erroneously – digit 0 following the first predictor that outperformed the other by one instance. In total, we will only predict 25 percent of the digits correctly – much worse than random guessing! But note that Schurz's imitate-the-best meta-inductivist won't fall into this particular trap because it is defined in such a way that it will only switch predictors if strictly outperformed.

crowd gets it right. Moreover, as we indicate below, the relation between the good listener and the wise crowd is only a paradoxical one if the latter is indeed concerned with the reliability of the crowd opinion. We therefore consider it appropriate to devote some attention to the absolute reliability of meta-induction.

The most immediate problem is simply that the predictors might all be wrong, and wrong in such a way that no clever combination of their predictions is going to perform much better. Consider a group of predictors each of whom consecutively predicts the next bit in a sequence by tossing a fair coin that is independent of past outcomes, and independent of the coins of the other predictors. Clearly, no meta-inductive method solely based on the inputs of this group can predict substantially better than random guessing, even if the data in fact contain a very predictable pattern. Thus, in general, even though we can guarantee for some meta-inductive methods that, no matter what data arrives, the method will perform about as well as the best predictor for that data, this offers no absolute guarantees. If there is no good predictor around, then the predictions made by the meta-inductive method may still be very bad in an absolute sense.

It may be thought that problems of this kind only show up for clumsily selected sets of predictors. But no dice. Similar problems can be constructed for any predictor, and for any meta-inductive method as well. The reason is that we might be unfortunate enough to be the victim of the inductive equivalent of Descartes' demon. Or, in the word of Thorn and Schurz, we might be the victim of a deceiver. A particularly clear exposition of such deceivers is provided by Putnam (1963a), and in simplified form in Putnam (1963b). Imagine that the world, or more specifically the process that is supposed to be predicted, conspires against the way the predictor is set up: any time the predictor is gaining some confidence in its predictions, for example because it has been right a fair number of times, the data generating process is such that it deviates from whatever it is the predictor is predicting next, thereby significantly reducing absolute performance.[3] Notably, low

[3]An amusing situation of this kind is described in Dostojewski's *Notes from the Underground*, in which the predictability of human beings is considered: " . . .even if man really were nothing but a piano key, even if this were proved to him by natural science and mathematics, even then [. . .] he would purposely do something perverse out of simple

performance will be no different for any meta-inductive method. In general, relative to any set of predictors coupled to a meta-inductive method, there will always be a set of conspiring worlds that are in principle unpredictable, in the same way as that there will always be such a set relative to a single predictor.

These considerations point to the crucial importance of how the set of predictors is composed. In order for the meta-inductive method to have any absolute reliability, its set of predictors needs to contain an element that can pick up the patterns in the actual data. In terms more familiar to statisticians (cf. Dawid, 1982), the meta-inductive method needs to be well-calibrated. But of course, considering the wild variety of patterns that might emerge in the data, there is no guarantee that the patterns in the data will indeed be picked up on by one of the predictors in the set. Therefore, while meta-induction is backed by some striking results on comparative success, we have not achieved anything in absolute sense. Hume's problem of induction still looms over our predictive systems, and as Howson (2000) argues this may well be an inevitable problem.

The natural response to the problem of calibration may be to accommodate all possible patterns in the set of predictors, but this would lead to an unmanageable set: there is in principle an uncountable infinity of patterns. An interesting approach to this strive for maximal generality is again found in algorithmic information theory, in particular the theory of prediction by Solomonoff (1964, 1978). From the assumption that the true data is generated in a computable way, it can be shown that there exist universal predictors that will always quickly converge to the true generating distribution. The constraint to computability allows us to find formal grip, while keeping things highly unrestricted. However, computability still serves as a limitation: the supposition that everything in the world must be calculable is not incontestable.

As a way of addressing this concern, we could opt for an anti-realist notion of truth to accompany this constraint, and stipulate that the truth is by

ingratitude, simply to gain his point . . . [T]he whole work of man really seems to consist in nothing but proving to himself every minute that he is a man and not a piano key!" Human beings, in other words, are exactly the sort of data generating processes that will conspire against any predictive system accessible to themselves.

definition something that can be calculated (cf. Douven et al., 2010). Alternatively, we could relocate the constraint of computability to the predictors, and interpret the above result as saying that there exist universal predictors that always quickly converge to the most successful predictor. From the perfectly natural requirement that every prediction method must be computable, we obtain the result that there exist universally optimal predictors. Of course, by that reading we do lose again the link to the truth, and absolute reliability. The affinity with the project in Schurz (2008) project should be clear. Indeed, one could look at this interpretation of Solomonoff's theory as an idealized limit case of Schurz's approach, where the class of predictors is maximal. To further accentuate the kinship of all of these ideas, it is worth noticing that some of the originators of the machine learning branch of prediction with expert advice, specifically Rissanen and Vovk, were also originally inspired by algorithmic information theory.

In the foregoing we have seen that groups might be reliable guides to the truth, in a relative sense but in an absolute sense too. For this we must presume some relation between the truth and the set of predictors. One possibility is that we were lucky when we selected the set of predictors, and that the correct predictor is already among the chosen set. Another possibility, of which Solomonoff's approach is an example, is that we include very many predictors in our set, by imposing only a very weak constraint on them, and that we further stipulate that the truth must be such that it can be picked up by one of these predictors. Yet another possibility, akin to the first one, is that we presume some inclination towards the truth in each predictor separately. The next section discusses this link to truth in more detail.

3 Good listener or parasitic expert?

Recall the paradox given in the introduction, which suggested a tension at the point where crowd wisdom and meta-induction intersect. Should a group of experts really welcome a member who merely compiles their own opinions on the matter at hand? Notice that the paradox becomes pressing when the aim of the group is to get things right in an absolute sense: it is counterintuitive that such a parasitic expert, or good listener, should bring us any closer to a reliable answer. If, on the other hand, we take a comparative reading of the wise crowd, there seems to be no paradox left. There is little or no surprise in what will happen to the comparative quality of the group opinion, obtained by applying an aggregation method to the group, if the group is enlarged by some people whose opinions are determined by applying that same aggregation method to the original group members. The group opinion will be comparatively better than most or all opinions held by the original group members, and exactly as good as the opinions of the good listeners.

The addition of a good listener to a wise crowd is, by our lights, a conceptual challenge in the context of reliability, not optimality. In what follows we will indicate that, as before, this issue hinges on the way in which truth is factored into the dynamics that governs group opinion. The focus in this part of our commentary lies on the dynamics that drives the wisdom of the crowd, and on how it interacts with the insights from meta-induction.

To clarify matters, it will be instructive to start with the case in which the crowd consists of meta-inductive learners only, and in which the group seems liable to mass hysteria. Thorn and Schurz consider this case in their simulations, but they hardly discuss that there is also an analytic theory that describes the formation of group opinions under such circumstances, to wit, DeGroot's opinion pooling (DeGroot, 1974; Lorenz, 2007). The theory of opinion pooling describes how group members may adapt their opinions on the basis of what other members think about an issue. In every new round of revision, each member takes some weighted average of the opinions of members of the group, perhaps giving a high weight to their own opinion but also giving some weight to other members that they trust. Under certain constraints on the trusts that group members give to each other at each revision round, the result is a process in which the opinions of group members converge to a middling point. In philosophical circles, this process has become known as Lehrer-Wagner consensus formation (Lehrer and Wagner, 1981), as the middling point is arguably a fair consensus for all members of the group.

It seems straightforward to model a group consisting of meta-inductive learners as a process of opinion pooling. Each group member simply gives all their trust to other members and nothing to their own opinion, accommodating that a meta-inductive learner only takes opinions of other members of the group as input. Furthermore, rather than having group members taking a weighted average of the opinions of other members, we have them take some functional combination of these opinions, which might vary from one round of updates to the next. This accommodates the fact that meta-inductive methods cover many more procedures for aggregating the opinions of group members than just taking the same weighted average. More or less in tune with the conclusions that Thorn and Schurz seem to derive from the simulation studies, an opinion pooling process of this kind need certainly not always lead to chaos. Indeed, depending on the meta-inductive methods that the members employ, the dynamics of group opinion may indeed end up in a form of mass hysteria, with group members effectively chasing their own tails without being anchored by any inclination towards the correct opinion. But under suitable conditions, we will find that the opinions of a group of meta-inductive learners will converge, and perhaps even converge to an opinion close to the correct one.

Our aim here is to sketch these conditions on the basis of insights from opinion pooling, and to draw some general lessons about the inclusion of meta-inductive learners in a group. Thorn and Schurz discuss this issue in some detail, and focus on a tension between two epistemic virtues: on the one hand groups should value dissenting and diverse opinions, but on the other hand they should be aware that the correct opinion most likely resides somewhere in between the opinions of the members, putting a premium on concordant opinions. We suggest that we can get a handle on this tension by involving the inclination towards the truth of the group members: if some group members have this inclination, then adding meta-inductive learners to the group need not lead to a wild goose chase through the space of opinions. Eventually, in virtue of group members who are sensitive to the truth, the group as a whole will approach this truth. The speed of convergence, however, will be influenced by the addition of meta-inductive learners.

In the literature on opinion pooling and consensus formation, there have been some attempts to come to grips with these aspects of group dynamics. The Hegselmann and Krause (2002) model of belief formation considers agents who determine and adapt their opinion in the light of their own previous opinions, the opinions of others in the group and, crucially, the results of their independent efforts to collect evidence. This latter component in the opinion dynamics of group members makes sure that each of them has her own inclination towards the truth, because the evidence is assumed to provide some inclination towards the true opinion. The model of Douven and Riegler (2009) generalizes the model of Hegselmann and Krause in various ways, among other things by considering that group members might temporarily disregard or fail to incorporate the opinions of some other group members. Interestingly, Douven and Riegler find that this may have beneficial effects on the quality of the opinions in the group, and hence on the group opinion. Because group members can determine things in relative isolation, some of them will end up reasonably close to the truth quickly. On the other hand, a more accurate opinion can be reached by involving the opinions of other group members, but this higher accuracy will take longer to achieve, making it less attractive to engage in information sharing and beneficial to maintain a level of diversity. Similarly positive results on the advantages of a relative isolation of group members are reached in Zollman (2010), who discusses the benefits of what he terms the transient diversity of groups. Kitcher (1993), finally, argues for an approach to science policy and institution design that acknowledges the value of such diversity.

We do not intend here to review all the insights from this intriguing literature in social epistemology. Instead, we would like to point to the relevance of the above insights for the tension between diversity and concordance that is identified by Thorn and Schurz. The key observation is that, in the studies just cited, the diversity of opinions originates in the fact that group members have their own independent means of getting to the truth. The members of the group all have their own epistemic access to the world, in the form of evidence that they collect independently. Because evidence comes with considerable noise, and because this noise is distributed evenly around the signal, the opinions of group members will show an informative sort of diversity, being scattered around the correct opinion. It is therefore not so much the diversity or independence as such, but more specifically the independence of evidence gathering that improves the group opinion, and hence merits stimulation.

Where does this leave the group of experts who are wondering whether or not to welcome a meta-inductive learner among them? We submit that the source of diversity will be an important consideration in determining whether this will be beneficial or detrimental to the opinion of a group. If the diversity is informative in the sense sketched above, then the addition of a meta-inductive learner will speed up the process of consensus formation, by bringing in the voice of a good listener. If, on the other hand, the diversity is relatively uninformative, perhaps because the distribution of noise is skewed, then the addition of a meta-inductive learner, or parasitic expert, is going to act as a dead weight on the process of convergence to the truth. It may even cause the group to reach consensus long before evidence gathering has led the group opinion to be anywhere near the truth.[4] In sum, if we want to determine how the addition of a meta-learner to a group affects the absolute reliability of the group opinion, we need to know how the opinions of the group members relate to the truth.

We should note that there may be other reasons for stimulating diversity and independence in a group, quite separate from the independence of information gathering. Dissent can serve as a magnifying glass for the topic under consideration, because it will invite the group members to engage with each other, spell out their opinions, and argue for them in detail. In other words, diversity of opinion has a kind of pragmatic value. Seen from that perspective, the addition of a meta-inductive learner to a group will have the detrimental effect that group members will see the reasonableness of a middling position too quickly, without quizzing each other on their opinions. However, in all fairness, there is a great discrepancy between the pragmatic and psychological reality of group dynamics and the abstract and normative models considered in social epistemology. While we think that this discrepancy should be overcome, we also believe that a meaningful normative discussion can be carried out without involving the descriptive side.

[4]In the case of science, we can think of the general public as consisting of meta-inductive learners, who base their scientific opinions on leading experts that appear in the media. Our considerations on the merits and defects of including good listeners in the group might therefore have some bearing on science policy, in particular on when to give the general public a say in the research agenda. However, to our mind current models of the social dimension of science are far too abstract to fulfill such a role in practice.

4 Conclusion

Much of the contribution of Thorn and Schurz concerns the relative optimality of group opinions: they are better than most if not all opinions of group members. In the foregoing, we have indicated two ways in which the truth comes into the equation after all. First, regarding meta-induction, we have emphasized that its absolute reliability hinges on there being some link between the set of predictors and the truth. The intention with this was not to criticize Thorn and Schurz, who focus on relative optimality, but rather to bring truth back into view. Second, we argued that the role of truth is crucial in resolving one of the central tensions in the article by Thorn and Schurz, namely between diversity and concordance. The focus should not lie on diversity of opinion itself, but rather on the independence of how opinions are geared towards the truth. The latter determines the quality of group opinion and the effects of the addition of a meta-inductive learner to the group. The proper analysis of the role of a good listener in a wise crowd might well require us to shift our attention from relative optimality to reliability and truth.

References

Bloor, D. (1976). *Knowledge and Social Imagery.* Routledge, London.

Cesa-Bianchi, N., Freund, Y., Helmbold, D., Haussler, D., Schapire, R., and Warmuth, M. (1997). How to use expert advice. *Journal of the ACM*, 44(3):427–485.

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning and Games.* Cambridge University Press, Cambridge.

Dawid, A. P. (1982). The well-calibrated bayesian. *Journal of the American Statistical Association*, 77(379):605–610.

DeGroot, M. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69:118–121.

Douven, I., Horsten, L., and Romeijn, J. (2010). Probabilist antirealism. *Pacific Philosophical Quarterly*, 91:38–63.

Douven, I. and Riegler, A. (2009). Extending the hegselmann-krause model iii: From single beliefs to complex belief states. *Episteme*, 6:145–163.

Galison, P. (1987). *How experiments end.* University of Chicago Press, Chicago.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded condence: Models, analysis, and simulations. *Journal of Articial Societies and Social Simulation*, 5. http://jasss.soc.surrey.ac.uk/5/3/2.html.

Howson, C. (2000). *Hume's Problem: Induction and the Justification of Belief.* Clarendon Press, Oxford.

Kalai, A. and Vempala, S. (2005). Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307.

Kitcher, P. (1993). *The Advancement of Science.* Oxford University Press, Oxford.

Latour, B. (1987). *Science in action.* Harvard University Press, Cambridge, Massachusetts.

Lehrer, K. and Wagner, C. (1981). *Rational Consensus in Science and Society: A Philosophical and Mathematical Study.* Reidel, London.

Lorenz, J. (2007). Continuous opinion dynamics under bounded condence: A survey. *International Journal of Modern Physics C*, 18:1819–38.

Putnam, H. (1963a). 'degree of conrmation' and inductive logic. In Schilpp, A., editor, *The Philosophy of Rudolph Carnap*, pages 371–383. Open Court, LaSalle.

Putnam, H. (1963b). Probability and confirmation. *The Voice of America, Forum Philosophy of Science 10, U.S. Information Agency*, 10:605–610. Reprinted in Hilary Putnam (1979), Mathematics, Matter and Method. 2nd ed. Cambridge: Cambridge University Press, pp. 293–304.

Schurz, G. (2008). The meta-inductivist's winning strategy in the prediction game: A new approach to Hume's problem. *Philosophy of Science*, 75:278–305.

Solomonoff, R. (1964). A formal theory of inductive inference. Parts I and.II. *Information and Control*, 7:1–22, 224–254.

Solomonoff, R. (1978). Complexity-based induction systems: Comparisons and convergence theorems. *IEEE Transactions on Information Theory*, IT–24(4):422–432.

Vovk, V. (1990). Aggregating strategies. In *Proceedings Third ACM Workshop on Computational Learning Theory (COLT 1990)*, pages 371–383.

Zollman, K. J. S. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72(1):17–35.