Analyse & Kritik 02/2012 (© Lucius & Lucius, Stuttgart) S. 339–365

Paul D. Thorn and Gerhard Schurz

Meta-Induction and the Wisdom of Crowds*

Abstract: Meta-induction, in its various forms, is an *imitative* prediction method, where the prediction methods and the predictions of other agents are imitated to the extent that those methods or agents have proven successful in the past. In past work, Schurz demonstrated the optimality of meta-induction as a method for predicting unknown events and quantities. However, much recent discussion, along with formal and empirical work, on the Wisdom of Crowds has extolled the virtue of *diverse* and *independent* judgment as essential to maintenance of 'wise crowds'. This suggests that meta-inductive prediction methods could undermine the wisdom of the crowd inasmuch these methods recommend that agents imitate the predictions of other agents. In this article, we evaluate meta-inductive methods with a focus on the impact on a group's performance that may result from including meta-inductivists among its members. In addition to considering cases of *global* accessibility (i.e., cases where the judgments of all members of the group are available to all of the group's members), we consider cases where agents only have access to the judgments of other agents within their own *local* neighborhoods.

1. Introduction

In a number of recent papers (2008; 2009c), Gerhard Schurz proposed a new solution to Hume's problem of induction, which emphasizes the importance of meta-induction. In its various forms, meta-induction proceeds by considering the past track record of other agents (and their prediction methods), and makes predictions of future events by reasoning that the agents (and prediction methods) that have been successful in the past will be successful in the future. Schurz demonstrated that under natural conditions various forms of meta-induction are guaranteed to yield optimal results, in the sense of having predictive success rates that converge to the success rate of the meta-inductivist's most successful competitor. These results (which we describe in the following section) do not entail that meta-induction is guaranteed to be a successful prediction strategy. Rather the results guarantee that if any method is successful, then prediction via meta-induction will closely approximate the success of that method.

* Work on this paper was supported by the DFG financed EuroCores LogiCCC project *The Logic of Causal and Probabilistic Reasoning in Uncertain Environments*, and the DFG project *The Role of Meta-Induction in Human Reasoning* (SPP 1516).

The optimality of meta-induction appears to provide a strong prescription for would-be predictors. But the matter is, perhaps, not so simple. The core injunction of meta-induction is to copy the strategies and predictions of those individuals who have proven most reliable. The prescription of meta-induction to copy others is prima facie in tension with prescriptions implicit in recent formal and empirical work on the Wisdom of Crowds. This work emphasizes the importance of agents making their predictions (and, more generally, their judgments) independently of the predictions of other agents.

Francis Galton's account of a contest that occurred at the 1906 West England Fat Stock and Chicken Exhibition is a popular touchstone for discussions of the Wisdom of Crowds, and serves as compelling, if anecdotal, illustration of the wise crowd effect. In the contest recounted by Galton, attendees at a local livestock exhibition could observe a large mature ox, and had the opportunity to guess the weight of the ox (in particular, the weight of the ox's remains after slaughter and dressing). Seven hundred and eighty seven persons entered the contest, and offered wide-ranging guesses. The remarkable fact about these guesses resided in their average (the aggregate judgment of the crowd, as it were). The crowd guessed that the ox would weight 1,197 pounds, and was almost exactly right, as the ox weighted 1,198 pounds.

Empirical studies have illustrated that the judgments of crowds (i.e., the average value of the judgments of a group's members) are remarkably reliable in the face of certain types of query. It is also straightforward to construct formal models of individual judgment wherein the average value of the judgments of a group of individuals tends to be very accurate (and much more accurate than the average accuracy of the group's individual members). Recent empirical studies also show that the accuracy of a crowd's judgment can be severely compromised when agents within the group are privy to the judgments made by other group members (and are thus able to imitate the judgments of other group members) (Lorenza et al. 2011). Similarly, well known formal models of 'wise crowds' require that the judgments of a group's members be stochastically independent of the judgments of other members of the group. So select empirical and formal results suggest that imitating the judgments of other group members is, contra meta-induction, a bad thing. Whether and to what extent this prima facie tension is genuine is the major question addressed in this paper.

2. The Optimality of Global Meta-Induction

While induction is central to scientific method, the problem of justifying induction is notoriously difficult, and was thereby described by C. D. Broad as "the glory of science and the scandal of philosophy". It was David Hume who showed that all standard methods of justification fail when applied to the task of justifying induction. In particular, the reliability of the inductive method cannot be justified by induction, by arguing that induction has been successful in the past, whence—by induction—it will be successful in the future, for this argument is circular, and circular arguments are without justificatory value (cf. Salmon 1957, 46; Schurz 2009b, §3.2).

Today most epistemologists are skeptical concerning the possibility of a direct solution to Hume's problem. In Schurz (2008; 2009c) an approach to Hume's problem is developed that is based on meta-induction. This approach is compatible with Hume's diagnosis that no non-circular argument can establish the reliability of induction. What the approach claims to show is that meta-induction is an epistemically optimal prediction method, in the sense that its predictive success is maximal among all methods of prediction that are accessible to any given agent. Epistemic optimality arguments are a game-theoretical generalization of Reichenbach's best-alternative argument for induction (Reichenbach 1949, §91). Even in radically skeptical scenarios where meta-induction is unreliable, meta-induction can still be optimal in the sense of being 'the best of a bad lot'. Moreover, the mathematical (or *a priori*) justification of meta-induction proposed in Schurz (2008) establishes a non-circular a posteriori justification of object-induction (i.e., induction applied to objects or events), by the following argument: We know from experience that object-inductive prediction methods have so far been more successful than non-inductive ones, whence it is meta-inductively justified to favor object-inductive methods in the future.

Many philosophers and cognitive scientists are skeptical concerning the possibility of universal optimality arguments (cf. Norton 2003; Gigerenzer et al. 1999). However, the arguments of these authors are concerned with methods at the object-level, not at the meta-level. The crucial difference is that meta-methods take the results of other methods as their input and may learn from them. The optimality of meta-induction is restricted to those methods whose output is accessible to the given agent(s). Of course, there might be methods (or forecasters) that do not reveal their predictions, but keep them secret. What the optimality theorems show is that among all accessible prediction methods, meta-inductivist strategies are the best choice. The restriction to accessible methods is not a genuine drawback, because methods whose output is not accessible are irrelevant to the epistemic decision problem.

To demonstrate the optimality of meta-induction, Schurz (2008; 2009c) introduced the notion of a prediction game. It consists of:

(1) An infinite sequence (e) := (e_1, e_2,...) of events (e.g., daily weather conditions or stock values) whose valued are drawn from the unit interval, so that $e_n \in [0, 1]$, for each round, n, of the game.

(2) A finite set of players Π, whose task in each round is to predict the value of the next event. '$p_n(P)$' denotes the prediction of player P at time n, which is delivered at time $n - 1$. It is assumed that admissible predictions are elements of [0, 1]. The players in Π include: (i) one or several meta-inductivists, denoted 'xMI', of a certain type x (see below), and (ii) a finite set of other players P_1, ..., P_m (so-called non-MI-players) who may either be ordinary inductivists (scientists), or alternative players (e.g., God-guided fortune-tellers). In para-normal worlds, the alternative players may have any success rate one wishes.

We assume cognitively finite agents who can simultaneously access only finitely many methods, so the player set is thus assumed to be finite. The predictions of all non-MI-players (or their respective methods) are assumed to be accessible to the meta-inductivists of the considered type xMI. This means that the meta-inductivists are permitted to make their predictions after all the other (non-MI) players have made theirs.

Further notation: The deviation of a prediction p_n from the event e_n is measured by a normalized loss function $l(p_n, e_n) \in [0, 1]$. The natural loss-function is the absolute difference $|p_n - e_n|$. However, the optimality theorems described below are not restricted to natural loss functions: Theorem 1 holds for all monotonic loss-functions, and theorem 2 holds for all convex loss-functions. The score obtained in round n is defined as $s(p_n, e_n) := 1 - l(p_n, e_n)$. The absolute success $a_n(P)$ achieved by player P until time n is P's sum of scores until time n ($\Sigma_{1 \le i \le n}\, s(p_i(P), e_i)$). The success rate $suc_n(P)$ of player P at time n is defined as $suc_n(P) := a_n(P)/n$. Finally, $maxsuc_n$ is the maximal success rate of the non-MI-players at time n. Binary prediction games consist in a special case where events and predictions are elements of $\{0, 1\}$. Here, $suc_n(P)$ coincides with the relative frequency of P's correct predictions until time n.

The simplest type of meta-induction is called 'imitate-the-best'. A player who employs this method is called a 'bMI'. In each round, bMIs imitate the prediction of the non-MI-player with the so-far highest success rate. bMIs change their favorite player as soon as another player becomes strictly better. If there are several best players, bMIs choose the first best player by an assumed ordering of the set of all players. The central result about the imitate-the-best prediction method is as follows:

Theorem 1: For each prediction game ((e), $\{P_1, \ldots, P_m, bMI\}$) that contains a best non-MI-player, B, after winning time n_B (i.e., $suc_n(B) > suc_n(P_i)$ for all $n \ge n_B$ and $P_i \ne B$), the following holds:

(1.1) Short run: For all rounds n, $suc_n(bMI) \ge maxsuc_n - (n_B/n)$, i.e., the bMI's short-run loss compared to the best non-MI-player is never greater than n_B/n.

(1.2) Long run: The bMI's success rate approximates the maximal success of the non-MI-players as n approaches ∞.

The size of the possible short run loss for a bMI derives from the fact that whenever a bMI recognizes that her present favorite P has earned a loss compared to some new best player P*, the bMI also receives this loss, before switching from P to P*. These losses may accumulate. The assumption of theorem 1 that $P_1, \ldots, P_m$, bMI includes a best non-MI-player B, with a winning time n_B, excludes the possibility that bMI can have more than finitely many losses due to switching favorites, since when the winning time n_B is reached, bMI sticks to the best player B forever. This assumption is violated whenever the success rates of two or more leading non-MI-players oscillate endlessly around each other. The worst case of endlessly oscillating success rates is produced by so-called systematic deceivers, who are clairvoyant, and (by definition) deliver a maximally wrong prediction whenever the meta-inductivist chooses them as their favorite, and deliver a maximally correct prediction otherwise. Systematic deceivers not only cause a breakdown in the optimality of simple imitate-the-best meta-induction, but of all kinds of so-called one-favorite meta-induction (such as ε-meta-induction; cf. Schurz 2008, 288), that base their predictions in each round on only one favored player or method. When a one-favorite method plays against systematic deceivers, her success rate converges to zero, while that of the deceivers is never smaller than 1/2.[1]

The optimality of imitate-the-best meta-induction and other one-favorite meta-inductive methods is very general, but not universal: Their long-run optimality is restricted to prediction games with a best player (or finite set of 'ε-best' players; cf. Schurz 2008, th. 2), and their short-run behavior is good only if the winning time of the best player occurs sufficiently early. A complete solution to Hume's problem calls for a meta-inductive strategy whose predictions are optimal in the case of systematic deceivers. Weighted meta-induction fills this role. A player who employs this method is called a "wMI".[2] A wMI predicts a weighted average of the predictions of the so-far 'most attractive' players. The attractivity $\mathrm{at}_n(\mathrm{P})$ of player P at time n is P's surplus success rate compared to the wMI's success: $\mathrm{at}_n(\mathrm{P}) = \mathrm{suc}_n(\mathrm{P}) - \mathrm{suc}_n(\mathrm{wMI})$, provided $\mathrm{suc}_n(\mathrm{P}) > \mathrm{suc}_n(\mathrm{wMI})$, otherwise $\mathrm{at}_n(\mathrm{P}) = 0$. A wMI's predictions are defined as $\mathrm{p}_{n+1}(\mathrm{wMI}) = \sum_i(\mathrm{at}_n(\mathrm{P}_i)\cdot \mathrm{p}_{n+1}(\mathrm{P}_i))\ /\ \sum_i(\mathrm{at}_n(\mathrm{P}_i))$, where P_i ranges over all accessible players. In words: a wMI's prediction for the next round is the attractiveness-weighted average of the attractive players' predictions for the next round. (If no player has positive attractivity, the wMI makes a random guess.) The following establishes weighted meta-induction's universal long-run optimality:

Theorem 2: For every real-valued prediction game ((e), $\{\mathrm{P}_1, \ldots, \mathrm{P}_m, \mathrm{wMI}\}$) whose loss-function $l(\mathrm{p}_n, \mathrm{e}_n)$ is convex in the argument p_n, the following holds:

(2.1) Short run: $\forall n \geq 1$: $\mathrm{suc}_n(\mathrm{wMI}) \geq \mathrm{maxsuc}_n - \sqrt{m/n}$.

(2.2) Long-run: $\mathrm{suc}_n(\mathrm{wMI})$ approximates the non-MI-players' maximal success as n approaches ∞.[3]

Theorem 2 does not directly apply to binary prediction games (and games with discrete-valued events), because a wMI's predictions—being weighted averages—are not binary, but are real values between 0 and 1. Theorem 2 can nevertheless be generalized to binary (and discrete) valued predictions, by assuming a population of sufficiently many, say k, meta-inductivists, who imitate the predictions of each attractive non-MI-player, P, with a population share that is approximately

[1] Theorems and descriptions of relevant computer simulations are found in Schurz 2008, §§4–6; 2009a, §§4–6.

[2] This method is based on results in machine learning theory (cf. Cesa-Bianchi/Lugosi 2006).

[3] In the application of weighted meta-induction, predictively equivalent non-MI players can be identified, without undermining the validity of theorem 2.

equal to P's attractivity (at the given time). Schurz calls this strategy 'collective weighted meta-induction'. The mean success rate of a population of collective weighted meta-inductivists will approximate the maximal success rate of the most attractive non-meta-inductivist players, with an additional maximal short-run loss of $1/(2 \cdot k)$.[4] Under the assumption that a group of collective weighted meta-inductivists share their success, collective weighted meta-induction is also a universally optimal strategy for each individual agent. Without this cooperation assumption, theorem 2 can be applied to a single agent in a binary or discrete prediction game by assuming that single collective weighted meta-inductivists imitate the predictions of each attractive non-MI-player P with a probability that is equal to P's attractivity (at the given time). However, this probabilistic variation is optimal only under the restriction that the environment is not 'demonic', i.e., under the condition that the future event e_{n+1} does not react to the wMI's imitation-choice for the given round n.

All the explained results about prediction games apply equally to action games. In that case, the partition of possible predictions is replaced by a partition of possible actions: At each time, n, each action, a, has a certain payoff $p_n(a)$ (that may change over time), and the score of an action in round n is computed by a loss function $l(a_n(P), max_n)$, where 'max_n' is the maximal payoff of the accessible actions at time n.[5]

3. The Wise Crowd

Following in the footsteps of some recent monographs (Surowiecki 2004; Page 2007, 179), we will say that the judgment of a crowd with respect to a true/false query is the (rounded) average response of its members, treating affirmation as one and disaffirmation as zero (thereby identifying the crowd's judgment with the majority response of its members, rounding to zero, disaffirmation, in the case of ties). Similarly, we say that the judgment of a crowd with respect to a scalar multiple choice query is the average response of its members (rounded as necessary).[6] In line with the preceding conventions, we say that a crowd is wise to the extent that its judgments are accurate. The question of how accurate a crowd's judgments must be in order to be considered wise is left open. In adjudicating claims to the effect that a given crowd is wise, it is often useful to make comparisons of the crowd's accuracy to the accuracy of its individual members, or to how accurate the crowd's judgments would have been within a salient counterfactual, such as one where some or all of the crowd's members make their judgments after having the opportunity to imitate some other members of the crowd.

Anecdotes such at Francis Galton's are relatively widespread, and it is clear that the judgment of a crowd with respect to some kinds of query frequently

---

[4] For theorems and computer simulations, see Schurz 2008, §§7–8; 2009a, §7; 2009c, §5.

[5] Details of this generalization are found in Schurz 2012a.

[6] We here omit consideration of non-scalar multiple choice queries, which may require subtle treatment.

exhibits uncanny accuracy (what we call “the wise crowd effect”). In the same vein as the task Galton described, Jack Treynor and others have illustrated the wise crowd effect by having groups of students guess the number of jelly beans contained in a large jar (Surowiecki 2004, 5). Invariably, the judgment of the crowd is very accurate, and exceeds the accuracy of all but a small handful of its members. Various markets, from stock exchanges to professional football betting lines (Surowiecki 2004, 12–15), and prediction markets such as the Iowa Electronic Markets and the Hollywood Stock Exchange (that were self-consciously designed to harness the wise crowd effect) have demonstrated the accuracy of groups of independently acting individuals in making various kinds of prediction (Surowiecki 2004, 17–22; Page 2007, 178). Recent empirical work has also demonstrated the tendency of crowds to make accurate judgments, as well as demonstrating the fact that even modest information about the judgments of other group members can undermine the wise crowd effect (Lorenza et al. 2011).

An early mathematical model that exhibits a sufficient condition for a wise crowd is described by the Condorcet Jury Theorem. The theorem considers a group of individuals, where for each member of the group, the probability is r that that member of the group will make a correct judgment regarding the truth value of some proposition α. It is further assumed that each individual's likelihood of making a correct judgment is stochastically independent of whether other members of the group make correct judgments. Under these conditions, the theorem tells us that the likelihood that the group is wise (i.e., that the majority response of its members is correct) converges to one as the size of the group n approaches ∞.

The Weak Law of Large Numbers suggests an obvious means of modeling wise crowds in the case of scalar multiple choice queries. Where ϵ is any number greater than 0, and X_n is a sample of n stochastically independent identically distributed random variables with mean μ, the Law of Large Numbers tells us that the probability that the mean value of the elements of X_n differs from μ by more than ϵ converges to zero as n approaches ∞. We may thus conceive the sample X_n as a set of predictions made by a group of n individuals about the value some unknown quantity μ, where the elements of X_n are independently and identically distributed around μ. In that case, the Law of Large Numbers tells us that, for any ϵ greater than zero, the probability that the group's judgment about the value of μ differs from μ by more than ϵ goes to zero as the size of the group n approaches ∞.

The Condorcet Jury Theorem and the Law of Large Numbers provide models describing how the average judgment of a group's members can be extremely accurate, provided the group is large and its members have some truth-bias (i.e., under the condition that there is a better chance than not that each group member makes a true judgment in the case of true/false queries, and under the condition that each group member's judgment is distributed around the true value with a mean value that is identical to the true value, in the case of scalar multiple choice queries). We will say that such a truth bias is the smaller, the larger the variance of the judgements around the true value. Observe that the assumption of even a ‘small’ truth bias is a ‘strong’ assumption inasmuch as it excludes systematic errors, in which case the mean value of the judgements of a group would not agree with the true value.[7]

Generally speaking, the assumptions under which the two theorems apply (such as the assumption that the accuracy of the judgments of agents are stochastically independent, for example) are rather unrealistic. A more realistic formal model of wise crowds is found in Page (2007), drawing on work by Krogh and Vedelsby (1995). The model of Krogh and Vedelsby dispenses with all assumptions about the character and interrelations of the probability distributions governing the judgments of a group's members. Their model simply assumes a group of agents who make judgments about the value of some unknown event (or object) x. Each agent, α, produces a judgment, $V^{\alpha}(x)$, corresponding to each event x. The judgment of the group, V(x), is defined as the weighted average of its member's individual judgments: $V(x) = \sum_{\alpha} V^{\alpha}(x)\cdot(1/n)$, where n is the size of the group.[8] The ambiguity of the input x for an agent α is defined as $D^{\alpha}(x) = (V^{\alpha}(x) - V(x))^2$, and the average ambiguity for the group as $D(x) = \sum_{\alpha} D^{\alpha}(x)\cdot(1/n) = \sum_{\alpha}(V^{\alpha}(x) - V(x))^2\cdot(1/\mathrm{n})$. Where $f(x)$ is the true value of x, the quadratic error of an individual, α, is $E^{\alpha}(x) = (f(x) - V^{\alpha}(x))^2$, and the quadratic error of the group's judgment is $E^{G}(x) = (f(x) - V(x))^2$. The average quadratic error for the members of the group is $E(x) = \sum_{\alpha} E^{\alpha}(x)\cdot(1/n) = \sum_{\alpha} (f(x) - V^{\alpha}(x))^2\cdot(1/n)$. Krogh and Vedelsby's simple, but important, observation is as follows:

Theorem 3: $E^{G}(x) = E(x) - D(x)$.

An immediate consequence of Theorem 3 is that $E^{G}(x) \leq E(x)$, i.e., the quadratic error of the group is no greater than the average of the quadratic error of the group's members. The difference between the two quantities is described by the average variance of the judgments of the group's members, as measured by D(x). In other words, the more diverse the judgments of a group, the more the accuracy of the group's judgment exceeds the accuracy of its average member (as measured by quadratic error).

Krogh and Vedelsby result holds for a more realistic model of judgments among groups. The result also represents a more balanced assessment of the value of diverse judgment. If we could hold E(x) fixed and vary D(x) (a measure of the diversity of a group's judgment), then we maximize the degree to which a crowd is wise by maximizing diversity. But diversity is, literally, one half of the equation: While increasing (or maintaining diversity) benefits the wisdom of the crowd, so does decreasing the average of the quadratic error of the group's members. If there were a means decreasing the value of the latter quantity while incurring relatively smaller decreases in diversity, then we could decrease diversity while at the same time increasing the wisdom of the crowd. In many

[7] Rather than assume that the judgements of each member of the population (or crowd) is symmetrically distributed around the truth, we could also produce a more realistic model of the wise crowd effect by assuming that the judgment of each agent has a systematic bias and that these systematic biases are themselves symmetrically distributed around the truth. The exploration of this model is left as future work.

[8] We here assume that each agent's judgment receives weight $1/n$. Krogh and Vedelsby's results hold generally where the weights assigned to the members of the group sum to one.

cases, introducing meta-inductivists into a group is a prime means of reducing the average error-rate among the group. As we will see (in the following section), there are also cases where the decrease in the average error rate among a crowd's members outstrips the 'damage' to the crowds' wisdom resulting from a loss in diversity.

Before proceeding we would also like to suggest, in contradiction to Page (2007, 208), that collective diversity (as measured by D(x)) is not as important as individual ability (as measured by E(x)) to the wisdom of a crowd (as measured by $E^G(x)$). To begin, we note that increasing diversity, D(x), is relatively easy, while reducing mean individual error, E(x), may be practically impossible. More importantly, decreasing E(x) to zero is sufficient (independently of the resulting effect on diversity, D(x)) for decreasing $E^G(x)$ to zero, while increasing diversity (independently of the resulting effect on E(x)) is not sufficient for decreasing $E^G(x)$. The first point is obvious, and recommends the application of imitate-the-best meta-induction as a means to achieving a wise crowd in any case where one member of the group is correctly identified as a perfect predictor. To see the second point, suppose we have a group of two agents attempting to predict the value of an unknown quantity μ. Suppose that $\mu = 0$, and the first of the two agents has judged that $\mu = 5$. If the second agent has judged that μ is certainly greater than 5, then he can increase diversity by guessing that μ is quite high (e.g., $\mu = 100$, $\mu = 1{,}000$, or $\mu = 1{,}000{,}000$). But the more our agent strives to increase diversity, D(x), the more the resulting disproportionate increase in E(x) and in $E^G(x)$.

4. The Relations between Meta-induction and the Wise Crowd

While the wise crowd effect recommends that forecasters (or players of a game) make their predictions independently from each other, meta-inductive learning suggests that forecasters should imitate the most successful forcasters whose predictions are accessible, and if they do this, their predictions are no longer independent of each other. So there is a tension between the injunctions of meta-induction and the preconditions for wise crowds. In the face of this tension, we evaluate meta-inductivist methods by considering the success of meta-inductivists within groups, and the impact on group performance that may result from including meta-inductivists among its members. In line with those persons that extol the virtues of diverse and independent judgment (e.g., Surowiecki 2004; Page 2007, and Lorenza et al. 2011), we describe a variety of conditions where replacing non-imitative players by meta-inductivists of certain sorts does indeed compromise the wisdom of the crowd, by reducing the accuracy of the average value of the judgments of that group of individuals. Note, however, that the investigated cases don't represent a shortfall in the individual performance of meta-inductivists, but simply situations where the precondition of independence does not hold, whence the crowd is less wise than it might otherwise have been, i.e., the 'judgments of the crowd' are less accurate than they might have been. This interpretation of our results is underscored by three points:

First: None of our simulations are in contradiction to the results on the optimality of meta-induction described in *section 2*.

Second: It makes a big difference whether the meta-inductivists replace or enrich existing strategies. As emphasized in Schurz (2008a; 2009c), meta-induction is parasitic on already existing independent prediction strategies (which are also called 'object-strategies' in Schurz, in distinction to 'meta-strategies' that imitate other strategies). The optimality theorems described above imply that whatever strategies are accessible to a given agent, it is always wise to apply meta-induction to them. In other words, adding meta-inductive strategies can only improve and will never diminish the maximal success rate, though replacing independent strategies by meta-inductive ones may reduce the maximal success rate, because such replacement may result in the exclusion of successful independent strategies.[9] For this reason, it is not generally recommendable to replace independent strategies by meta-inductive ones, but only to enrich them. If the independent strategies are not replaced but complemented by meta-inductivists, then the strategy WC, for 'wisdom of the crowd', i.e., the strategy of predicting the average of values of the independent predictions of a group's members, could be imitated by meta-inductive players. So the wisdom of the crowd, as defined by the average of the independent predictions of the group's members, would not be lost in such a population, while the use of meta-inductive strategies would improve the success rates of the individuals.

Third: The judgment of a crowd is not assumed by us to be a strategy that is accessible or applied by any individual player. Indeed, if we do assume that at least one member of a population plays the strategy WC (by accessing the independent predictions of all forecasters and predicting their mean value), then the theorems about meta-induction explained in *section 2* can be applied. It follows from these theorems that in all scenarios in which the strategy WC is optimal, the meta-inductivist strategies will approximate the success of WC (i.e., the mean success of meta-inductivists having access to the strategy WC will be as good as the strategy WC). As this is a consequence of already existing results, we focus here on the situation in which no individual plays the strategy WC, and

[9] The result that social learning is parasitic on individual learning was recently confirmed by Rendell et al. 2010. This paper investigates the result of a tournament of computer-programmed strategies within a scenario called "multi-armed bandit". Within such scenarios, every player (or strategy) may choose, in each round, one of many possible actions with unknown payoffs. Each player/strategy had the possibility to choose, each round, between three options: (a) individual learning (observe the payoff of one of many possible actions that you perform by yourself), (b) social learning (observe the payoffs of the actions of several other population members), or (c) exploit (perform a strategy whose payoffs had been learned by simple induction). The payoffs were assumed to be constant; so simple observation of observed payoffs was guaranteed to be relatively successful. Within the tournament, it turned out that social learning strategies (strategies that applied social learning (b) much more often than individual learning (a)) were the most successful. However, in a second tournament in which the strategies had to play within homogeneous populations, where all players adopt the same strategy, the effect was reversed: The social strategies were the worst and the individual strategies the best. This is also called "Roger's paradox" (Rendell et al. 2010, 72).

ask what happens to the wisdom of the crowd when meta-inductivists or other social learners replace independent forecasters. The effect of such replacements on the wisdom of the crowd is neither trivial nor entailed by already existing results on meta-induction and other forms of social learning (e.g. Schurz 2012b; Hegselmann/Krause 2005; 2009; Hartmann et al. 2009).

We also acknowledge that the most difficult epistemological situations for meta-inductive strategies are ones in which neither the event frequencies nor the success rates converge to limits in the long run ($n \to \infty$), but are endlessly oscillating. However, in this paper we will make the simplifying assumption of converging event frequencies and converging success rates. This is justified because our focus is on a different theme, namely what happens with the mean success rate of a crowd when social learning strategies replace independent strategies. Investigation of scenarios with permanently changing event-frequencies and success-rates is left for future work.

In contrast to the single shot prediction games described by Galton (guessing the weight of an ox) and Treynor (guessing the number of jelly beans), we envision an iterated prediction game where players may obeserve the accuracy of the predictions made by other players in previous rounds of the game. We furthermore envision a game where players have the opportunity to observe the opinions of other players at given moments, and thereby imitate the opinions and predictions of other players. Given these features, we may regard our simulations as modelling a variation of the prediction games described by Galton and Treynor, where the game is iterated, and players have the opportunity to observe the predictions and performance of (at least some) other players, and thereby imitate other players on the basis of those obervations. The simulations are also adequate to modeling similar predictions about, for example, future stock prices, such as predictions about whether a given stock price will be up at the close of trading on the following day (a binary-valued prediction), and what the price of a given stock will be at the close of trading on the following day (a real-valued prediction). The precise details are described in the following section.

5. The Formal Setup

To assess the impact of including meta-inductivists in a crowd (or of having members of a crowd adopt meta-inductive methods), we used computer simulations (programmed in Visual Basic.NET). Departing slightly from the sort of prediction games described in section 3, our simulations were based on the following ingredients:

(1) A quadratic grid consisting of $100 \times 100 = 10{,}000$ cells. Each cell corresponds to an individual player.

(2) For some simulations, we assumed that each agent in the grid has access to the success rates and the present judgment of every other player. In other simulations, we assumed that each player only has access to information concerning the players in her Moore-neighborhood, i.e., to the

neighborhood consisting of the player herself plus the eight immediately surrounding players

(3) The event sequence is either: (i) a random sequence of values chosen according to a uniform probability distribution on the unit-interval [0, 1], or (ii) a binary event sequence generated by rounding the elements of a sequence of the sort described in (i), where values greater than 0.5 are treated as 1, and the remaining values are treated as 0. In the case of a binary event sequence, players are required to predict that the true value of any event is 0 or 1. In the case of the real-valued event sequence, players may predict any real number. Our assumption is that players are not aware that the values of the actual event sequence are drawn from the unit interval.

(4) In the case of a binary event sequence, each player has a predefined independent reliability, r, which is the player's hypothetical success rate in the long run, assuming she bases her predictions solely on her own abilities (and does not imitate other players). (A player's indepedendent unreliability, u, is 1–r.) In the case of a real-valued event sequence, we distinguish two kinds of players: truth-biased players and ignorant players. Each truth-biased player has a predefined independent unreliability u, where u is the limiting average absolute deviation of the agent's predictions from the true values, assuming she bases her predictions solely on her own abilities. We also assumed that a truth-biased agent's independent judgments in the real-valued case are normally distributed with a mean identical to the true event-value (so that the standard deviation of an agent's independent guess is $u \cdot \sqrt{2/\pi}$). ($\sqrt{2/\pi}$ is the ratio of the mean absolute deviation to the standard deviation in the case of normal distributions.) In contrast to truth-basised players, the predictions of ignorant players are chosen by a uniform distribution on the unit interval. (In the case of a binary event sequence, we would treat ignorant players as blind guessers, who predict that the next value in the event sequence is 0 or 1 with equal probability.)

(5) Again the game consists of rounds, but now in addition, each round consists of successive cycles, in which predictions may be updated by imitating the predictions of other accessible players.

(6) In addition to their independent prediction abilities, some players may apply one of the following imitative prediction methods to other accessible members within the grid:

   (a) Imitate-the-best meta-induction: Such players (bMIs) imitate the player whose present success rate is greatest among all accessible players. In the first round, and in the first cycle of any round bMIs predict by independent means. Subsequent to the first round, bMIs select a favorite whose success rate is highest among all accessible players, and switch favorites only when another player's success rate exceeds the success rate of her current favorite. (In the second round when mutiple players have the highest success rate, or in a subseqeuent rounds where multiple players have success rates which exceed the success rate of her current favorite, bMIs choose the nearest player among the most successful players to be her new favorite. In the case where nearness does not break all ties, the new favorite is chosen by a predefined ordering of all players.)

(b) Weighted meta-induction wMI: In the face of a real-valued event sequence, wMIs predict the attractivity weighted average of the predictions of those players accessible to the wMI. In the case of binary event-sequences, we required the 'weighted' meta-inductivist to deliver binary-valued predictions. So in the case of a binary-valued event sequence, wMIs predict the rounded attractivity weighted average of the predictions of those players accessible to the wMI, where values greater than 0.5 are rounded to 1, and values of 0.5 or less are rounded to 0. Thus for binary event types a wMI predicts what the majority of the attractive forecatsers predict. In face of both real-valued and binary event sequences, wMIs predict by independent means in the first round, in the first cycle of any round, and whenever they themselves have the highest success rate (with the result rounded in the binary case).

(c) Peer-imitation: Peer-imitators predict an unweighted average of the predictions of those players accessible to the peer-imitator (with the result rounded in the binary case).

Compared to the sort of prediction games described in *section 2*, the present setup allows for mutual imitation between imitative players of various sorts. Since a player can imitate another player only after that player has made a prediction, the imitation process is now modeled via successive update cycles, in which every player imitates the predictions that her favorite(s) delivered in the previous cycle. In the first cycle of each round, each player delivers a prediction based on her independent method. In all following cycles, independent players repeat their initial prediction, while imitating players apply their imitative prediction method to the predictions made by accessible players in the previous cycle. This continues until the process of cycling reaches an equilibrium state (i.e., a state that is not changed by further updating), or until a preselected deadline, i.e., a maximum number of cycles, is reached. The predictions in the last cycle of a round are called the final predictions of that round. After the final predictions for a round are determined, the actual success rate for each player is updated, and a new round (with a new sequence of prediction cycles) begins, until the final round of the game is reached.

6. The Meta-Inductivist in the Crowd with Universal Accessibility

In the present section, we illustrate some effects of including meta-inductivists in a crowd, in cases where the predictions and success rates of all agents are accessible to all agents. As a baseline, we consider populations composed wholly of independent predictors (i.e., players who always and only predict according to their independent prediction abilities). The results are presented in *table 1.* Note that we present the results of the simulations in terms of the predefined unreliability of the independent predictions of players, and in terms of the average size of the error of the predictions made by all players within a game and the average size of the error of the group's judgment (which we call the "average global error".)

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% independent (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4703	0.4801	0.4899	0.4990	0.5000	0.5500	0.5000	0.7499	1.0005	4.9974	19.9979	0.3291
average global group error	0.0000	0.0000	0.0186	0.4297	0.4943	1.0000	0.0050	0.0075	0.0102	0.0516	0.1904	0.2422

Table 1

Our simulations reflect the fact that the mean individual error rates of non-imitators converge to their independent unreliabilities. We also observe (as predicted by the Condorcet Jury Theorem and the Law of Large Numbers) that the average global group error (i.e., the mean group error over the 1000 rounds considered for each simulation) for non-imitators is quite low, with some exceptions. In the binary case, setting individual unreliability, u, to be somewhat less than 0.5 is sufficient to make it a practical certainty that the crowd is very wise. On the other hand, setting u to be somewhat higher than 0.5 is sufficient to ensure that the crowd is very unwise. In the real-valued case, even a modest truth bias, such as u = 20 (i.e., the limiting mean absolute deviation for an agent's prediction from the true value is 20), is sufficient to achieve a relatively low average global group error. An interesting comparison, in this case, is to ignorant non-imitators who predict values selected via a uniform distribution on the unit interval. Since the mean difference between two values selected via a uniform distribution on the unit interval is $1/3$, we see that the average individual error converges to $1/3$. The average group error converges to $1/4$, since average group prediction is $1/2$, and the average distance between $1/2$ and a value selected via a uniform distribution on the unit interval is $1/4$. This is an instantiation of the effect of additive variance illustrated by the model of Krogh and Vedelsby (and theorem 3), which assumes neither truth-bias nor independence. It is perhaps surprising that the mean group error for non-imitators with very low independent unreliability, such as u = 20, is less than that of ignorants (who are not so ignorant inasmuch as their predictions fall within the same range as the true values, namely [0, 1]). The result illustrates the remarkable effect of assuming that the predicted values for agents have a mean value identical to the true value. Indeed, the present simulations provide a stochastic model of the wise crowd effect, which is exhibited by the anecdotes of Galton (guessing the weight of an ox), and Treynor (guessing the number of jelly beans), and by various prediction markets.

We next consider populations composed wholly of bMIs, which replace the independent predictors; so only the bMI's predictions and their success rates are accessible to the members of the population. In comparison with populations of non-imitators just considered, we may regard the data concerning these bMIs as representing a counterfactual wherein the non-imitators had instead been bMIs (i.e., a population of independent predictors with a fixed independent unreliability, u, is replaced by a population of bMIs whose independent prediction unreliability is also u). The results are presented in *table 2*.

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% bMI (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4780	0.4767	0.4906	0.4959	0.5017	0.5664	0.4952	0.7436	1.0081	5.1173	19.9985	0.3379
average global group error	0.4768	0.4763	0.4895	0.4952	0.5009	0.5658	0.4947	0.7429	1.0071	5.1123	19.9784	0.3377

Table 2

Within the described populations of bMIs, the average group error converges to the average individual error. This occurs since one of the bMIs eventually achieves a success rate higher than all of her peers, and thereafter all of the bMIs imitate that bMI (since that imitation pattern is an equilibrium state). This means that there is no wise crowd effect within any group composed wholly of bMIs, and that the mean individual error of such groups converges more slowly to the independent unreliability, u, of the group's members, as compared to a population of non-imitators. The fact that the average group error tends to be slightly smaller than the mean individual error is due to a wise crowd effect that occurs in the first round (in the binary case) or in the early rounds (in the real-valued case), prior to the point at which a single bMI distinguishing herself as the 'most reliable'.

The absence of the wise crowd effect, in the present case, may be explained as follows: If a bMI player plays very well in the early rounds of a prediction game, sufficient to achieve a success rate higher than all of the bMIs for a single round, then the other bMIs will imitate the predictions of that player from that point on, regardless of how inaccurate the predictions of the player become. So for the bMIs, minimal random fluctuations in the error rate of individual group members are enough to deadlock the population in a trap where each bMI imitates a single player. To address this problem we introduced a variety of bMI, εbMI, who switches to a new favorite only if the success of the new favorite is greater than that of the old favorite by at least ε. However, this variation did not significantly change our result, the disappearance of the wise crowd effect. This is shown in the following tables:

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% εbMI, ε = 0.05 (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4733	0.4619	0.5004	0.5366	0.4896	0.5482	0.5042	0.7672	1.0419	5.1299	19.3489	0.3342
average global group error	0.4690	0.4600	0.4980	0.5380	0.4880	0.5500	0.5036	0.7664	1.0408	5.1246	19.3291	0.3340

Table 3

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% εbMI, ε = 0.25 (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0,499)	100 (0.4999)	100 (0.55)	100 (0..5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4717	0.4766	0.4875	0.5085	0.4786	0.5595	0.5041	0.7587	1.0004	4.8509	19.9751	0.3337
average global group error	0.4720	0.4660	0.4590	0.5140	0.4760	0.5730	0.3775	0.7124	0.9851	4.8457	19.9552	0.2702

Table 4

The explanation of this result for εbMIs is as follows: In each round, only one of at most a small number of εbMIs will have maximal success rates, and each individual εbMI player imitates one of those players, so again, the diversity of the individual predictions made by the εbMIs will be extremely reduced (to one or a few), whence averaging over these few predictions cannot produce a significant wise crowd effect.

We next consider populations composed wholly of wMIs. The results are presented in *table 5*.

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% wMI (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4620	0.4660	0.5030	0.5120	0.5120	0.9853	0.0149	0.0217	0.0287	0.1455	0.5832	0.2497
average global group error	0.4611	0.4646	0.5020	0.5120	0.5110	0.9991	0.0114	0.0167	0.0220	0.1131	0.4438	0.2494

Table 5

Considering *table 5*, we notice some remarkable patterns that characterize groups composed wholly of wMIs. In the case of binary prediction games (e.g., when the task is to predict whether a given stock price will be up at the close of trading on the following day), where the independent unreliability, u, of the wMIs is less than 0.5, we observe no wise crowd effect. This is due to the fact that for binary events wMIs predict according to the majority of attractive players. In these cases, the populations of wMIs tend to reach an equilibrium state where a single wMI is distinguished as the most reliable, while all of the less reliable wMIs have identical reliability. These cases are in contrast to the binary case where u is greater than 0.5. In that case, the wMIs did not stratify themselves as they did in the cases where u was less than 0.5. Instead we observe a sort of 'reverse wise crowd effect', similar to the effect observed in the case of non-imitators (with the additional effect of surging individual error rates).

In contrast to the binary case, the wMIs exhibit a wise crowd effect in the case of real-valued event sequences (e.g., when the task is to predict the price of a given stock at the close of trading on the following day), so long as the wMI are not ignorants. The effect is slightly weaker than in the case of independent predictors, since some diversity is lost through imitation. The reward for the loss in average group success rate (in comparison to non-imitators) is a large increase in average individual success rates.

Recall that all of the results presented here depend on our assumption that the wise crowd prediction strategy (WC) is itself not accessible to the meta-inductivists (whether bMIs or wMIs). In any case where this strategy is accessible, meta-inductivists would start to imitate it, and thereby make individual predictions that are as accurate as those of the wise crowd.

The scenarios investigated so far are unrealistic in assuming that all of the predictors in the group are equally reliable. This unrealistic assumption is also biased against meta-inductivists, whose forte consists in imitating the predictions of those predictors whose predictions are the most accurate. In the next three tables, we present data from simulations for groups in which meta-inductivists replace independent strategies. These simulations differed from the simulations describes in *tables 1, 2, and 5*, by including a subpopulation of highly reliable independent predictors.

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% independent A (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent B (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.4331	0.4419	0.4509	0.4591	0.4598	0.5051	0.4601	0.6851	0.9098	4.5104	18.0132	0.3087
average global group error	0.0000	0.0000	0.0000	0.0000	0.0000	0.8620	0.0047	0.0071	0.0095	0.0483	0.1881	0.2236

Table 7

Once again, we see that non-imitators have individual error rates that converge to their independent unreliabilities. Thus we see only a small (but significant) decrease in individual unreliability resulting from the inclusion of subpopulation of very reliable independent predictors ($u = 0.1$). The crowd's wisdom also increases, but the effect is minimal.

The impact on the average individual error rate of players is much greater when we replace the less reliable subgroup on non-imitators with bMIs or wMIs (*tables 8 and 9*). The prediction strategy of the bMIs yields the result that the individual error rates of the bMIs approximates the error rate, $u = 0.1$, of members of the highly reliable subgroup. The bMIs also achieve lower group error rates, in addition to lower individual error rates (as compared to non-imitators), in the binary case where u is greater than 0.5, and in the real-valued case where u is very high (20) or reflects the absence of a truth-bias (i).

In the case where the less reliable subgroup is composed of wMIs, the average group error rate is usually lower than that of the bMIs. This difference derives from the fact that wMIs tend to imitate (via weighting) multiple players, thereby by preserving greater diversity, which has a positive impact on accuracy in the case where the imitated players are truth-biased. For similar reasons, wMIs achieve lower individual error rates, in most cases.

rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% bMI (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent B (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.0974	0.0908	0.1144	0.0991	0.1017	0.1063	0.0961	0.1041	0.1000	0.1065	0.1165	0.1016
average global group error	0.0969	0.0890	0.1150	0.0980	0.1010	0.1070	0.0856	0.0934	0.0891	0.0921	0.0895	0.0915

Table 8

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% wMIs (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent B (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.0793	0.0711	0.0784	0.0775	0.0703	0.0775	0.0235	0.0303	0.0378	0.0966	0.2085	0.1554
average global group error	0.0758	0.0670	0.0750	0.0740	0.0659	0.0750	0.0101	0.0155	0.0212	0.0762	0.1694	0.1453

Table 9

Finally, we see that applying imitate-the-best or weighted meta-induction results in higher individual success rates (as compared to making independent predictions), so long as we assume that the meta-inductivist has the opportunity to imitate players whose independent reliability exceeds her own. We maintain that this is a realistic assumption. Under the realistic assumption of heterogeneous success rates, the application of meta-inductive prediction methods may also achieve higher group success rates. This tends to occur when many of a group's members have independent reliabilities that are not truth-biased.

We can summarize the results so far by considering how groups composed of the three types of players (non-imitators, bMIs, and wMIs) would perform on an iterated prediction game. For purposes of illustration, we consider two cases: (1) the players are predicting whether a given stock price will be up at the close of trading on the following day (a binary-valued prediction), and (2) the players are predicting what the price of a given stock will be at the close of trading on the following day (a real-valued prediction). The performance of individual non-imitators in either sort of prediction game isn't very interesting: The individual performance of such players (as measured by average error) precisely tracks their individual unreliability. The group performance of such players is more interesting: a small truth bias is sufficient to translate into group judgments that are extremely accurate. Unlike non-imitators, bMIs and wMIs track the performance of other players and imitate the ones that seem to be doing well. In the case of both binary-valued predictions (predictions about whether a given stock price will be up) and real-valued predictions (predictions about what the price of a given stock will be), bMIs gravitate to imitating the predictions a single player. The individual accuracy of bMIs is thereby determined by the accuracy of the group's most reliable member. Similarly, if bMIs represent the vast majority of a group's members, then the accuracy of that group will be identical to the accuracy of the group's most reliable member, since all of the bMIs will imitate that individual. Unlike an bMI, a wMI will imitate all (accessible) players that are performing better than himself. This feature of wMIs serves to recapitulate some of the diversity that is present in the independent predictions made by individuals within groups whose membership is dominated by wMIs, save in the case binary-valued predictions (e.g., predictions about whether a given stock price will be up) among a population of wMIs with homogeneous independent reliability. Outside of such cases, the tendency of wMIs to recapitulate the diversity found in the independent predictions of group members yields a modest wise crowd effect, along with reduced individual error rates, so long as there is a truth bias in the independent predictions of the wMIs.

In the present section, we considered situations where players had the opportunity to monitor and imitate the predictions all of the members of the group. In the following section, we consider a more realistic species of situation where the opportunity to monitor other players is limited.

7. The Meta-Inductivist in the Crowd with Restricted Accessibility

In the present section, we illustrate the effect of including meta-inductivists in a crowd, in cases where access to the predictions and success rates of agents is limited to other agents in their Moorean-neighbourhood. Within these scenarios we consider peer imitation, in addition to the prediction methods considered in *section 6*. We did not run additional simulations with non-imitators, since the performance of these players is invariant when accessibility is restricted, since these players never imitate other players.

In all of the simulations considered in this section, we set the maximum number of update cycles for each round to ten. This means, in effect, that information about the predictions and success-rates of one agent can travel nine cells, at most, to reach another agent. To understand why this is the case, consider *figure 1*, in which a single perfectly reliable expert is surrounded by perfectly unreliable bMIs. In the first cycle of the first round, the expert in the center makes a perfect prediction (with maximal score 1) and the bMIs surrounding make him a fully imperfect prediction (with score, s, less than 1). In the next cycle nothing changes, so round 1 immediately converges to equilibrium. The first cycle of the second round begins as before. But in the second cycle of the second round the expert will have a slightly higher success rate than the bMIs, so the bMIs in the first neighborhood layer will imitate the expert's prediction of the first cycle. In the third cycle of the 2nd round nothing further changes. In the third round, the first-layer of bMIs around the expert have an increased success-rate (compared to the bMIs more distant from the expert), and so, the bMIs in the 2nd neighborhood layer imitate the bMIs in the 1st neighborhood layer, whence in the 3rd cycle of the 3rd round, the expert's prediction is imitated by the bMIs in the 2nd neighborhood layer (etc.). In conclusion, it takes n rounds with n cycles for an experts' prediction to spread to $n-1$ iterated layers of a neighborhood.

g g g g g	g g g g g	**p p p p p**
g g g g g	g **p p p** g	**p p p p p**
g g **p** g g	g **p p p** g	**p p p p p**
g g g g g	g **p p p** g	**p p p p p**
g g g g g	g g g g g	**p p p p p**
1st round 1st cycle 2nd round 1st cycle 3rd round 1st cycle etc.	2nd round 2nd cycle 3rd round 2nd cycle etc.	3rd round 3rd cycle etc.

Figure 1: Spread of an expert prediction through cycles and rounds. Center cell = expert, p = the expert's prediction, neighboring cells = initially ignorant bMIs, g = bMI's predictions.

The propagation of predictive success by local meta-induction in social networks in which the accessibility is restricted to Moore-neighborhoods, and the advantage of local meta-induction over success-independent social learning methods such as peer-imitation or authority-imitation, has been investigated in detail in Schurz (2012b). In that paper, wise crowd effects were not studied; it was rather assumed that laymen were ignorants (and not truth-biased), so the wise crowd effect could not arise. In this section, we study the wise crowd effect under the condition that social learning methods are restricted, due to limitations in what other players are accessible.

The results of our simulations for populations composed wholly of peer-imitators, bMIs, and wMIs are described in *tables 10, 11, and 12*, below. In addition to presenting the mean group error rate for these simulations (reflecting the judgment of the crowd as a whole), we now include data concerning the mean error rate of individual Moorean-neighborhoods (reflecting the 'group judgments' of individual Moorean-neighborhoods). We here use the expression 'average local group error' to denote the mean error rate of individual Moorean-neighborhoods. For all simulations, we estimated the value of the average local group error rate by randomly sampling ten Moorean-neighborhoods in each round of each simulation, and compiling for each sampled neighborhood, N, the average distance from the true event value from the average of values predicted by the members of N (with rounding in the binary case).

Peer imitation is a very simple imitative prediction method. The method does not consider the success rates of its neighbors, but proceeds as if the predictions of each neighbor is equally credible. The strategy inherent in peer-imitation has some connection to the wise crowd phenomena inasmuch as the judgment strategy of peer imitation is similar to the procedure of treating the unweighted average judgments of a group's members as the group's judgment. Moreover, in the case where a peer-imitator has access to the judgments of all members of a group, her judgments will be identical to the judgment of the group. The added effect of peer-imitation over non-imitation, in the latter case of universal access, is that accurate (or inaccurate) judgments on the part of the group translates into accurate (or inaccurate) judgments on the part of the peer-imitator. In cases where access is limited (as described in *table 10*) the connection between the accuracy of the group and the accuracy of the peer-imitator is weakened, but we still observe a tendency of peer-imitators to emulate the judgment of the group, resulting in improved individual accuracy in cases where the group's accuracy is high, and poor individual accuracy when the group's accuracy is poor. We also see that peer-imitation has a small effect in decreasing the diversity of the group, and thereby on the accuracy of the judgments of the group (compared to non-imitation), in cases where the predicted events are binary (e.g., predictions about whether a given stock price will be up) and there is a truth-bias in individual judgment.

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% Peer Imitator (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.3599	0.4034	0.4529	0.4952	0.4998	0.7271	0.0590	0.0885	0.1183	0.5905	2.3659	0.2521
average global group error	0.0000	0.0010	0.0700	0.4340	0.4930	1.0000	0.0050	0.0073	0.0101	0.0493	0.2026	0.2510
average local group error	0.3499	0.3995	0.4474	0.491	0.4914	0.7338	0.0560	0.0829	0.1110	0.5525	2.2152	0.2516

Table 10

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% bMI (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4708	0.4797	0.4884	0.4978	0.5014	0.5492	0.4990	0.7505	1.0110	3.8788	14.1567	0.3351
average global group error	0.2570	0.3267	0.4162	0.4710	0.5160	0.8320	0.0578	0.0798	0.1130	0.5746	1.6877	0.2548
average local group error	0.4793	0.4798	0.4968	0.4963	0.5014	0.5485	0.4388	0.6618	0.8742	3.3768	12.5065	0.3212

Table 11

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% wMI (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.4665	0.4821	0.4880	0.4979	0.5005	0.5501	0.4268	0.6380	0.8470	4.2110	16.9375	0.3116
average global group error	0.2630	0.3760	0.4220	0.4850	0.4800	0.8350	0.0437	0.0637	0.0813	0.4149	1.6728	0.2509
average local group error	0.4712	0.4840	0.4904	0.4979	0.4977	0.5482	0.3906	0.5812	0.7855	3.8140	15.5949	0.3088

Table 12

In cases where the independent unreliability, u, of all players is identical (such as described in *tables 10, 11, and 12*), peer-imitators perform at least is well as bMIs and wMIs with respect to individual and group success rates, so long as u is truth-biased. In other words, in the case of homogeneous populations, peer-imitators will be at least as accurate as bMIs and wMIs at predicting whether a given stock price will be up at the close of trading on the following day (a binary-valued prediction), and at predicting what the price of a given stock will be at the close of trading on the following day (a real-valued prediction), so long as the individual reliability of the group's members is truth-biased. The situation is somewhat different in the case where the bMIs and wMIs have the opportunity to imitate players with higher independent reliabilities. Simulations of such situations are described in the following tables:

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% peer imitator (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.1660	0.1881	0.2225	0.2478	0.2599	0.4117	0.0508	0.0637	0.0783	0.3257	1.2435	0.1318
average global group error	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0032	0.0042	0.0055	0.0258	0.1038	0.1182
average local group error	0.1540	0.1783	0.2151	0.2311	0.2611	0.4128	0.0425	0.0553	0.0683	0.2966	1.1413	0.1225

Table 16

rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% bMIs (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.1019	0.1021	0.1030	0.1024	0.1018	0.1030	0.1008	0.1010	0.1017	0.1067	0.1255	0.1008
average global group error	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.0059	0.0058	0.0051	0.0045	0.0052	0.0067
average local group error	0.0844	0.0878	0.0932	0.0886	0.0867	0.0872	0.0755	0.0758	0.0756	0.0791	0.0927	0.0760

Table 17

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% wMIs (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.0969	0.0973	0.0980	0.0983	0.0973	0.0976	0.0977	0.0993	0.1006	0.1153	0.1257	0.0988
average global group error	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.0093	0.0094	0.0090	0.0096	0.0123	0.0224
average local group error	0.0850	0.0832	0.0835	0.0872	0.0899	0.0909	0.0816	0.0832	0.0824	0.0930	0.0947	0.0839

Table 18

In the cases described in *tables 16, 17, and 18*, where we include a minority subpopulation of highly reliable independent predictors, we observe that the performance of bMIs and wMIs exceeds that of peer-imitators in most situations. Peer-imitators performed better than meta-inductivists only in the case of predicting real-valued events (e.g., predicting the price of a given stock at the close of trading on the following day), where the truth bias of the predictions of the imitator majority was quite strong.

8. Cautious Weighted Meta-Induction: The Best of Both Worlds?

The data generated by our simulations illustrate that a group containing bMIs or wMIs, in the place of non-imitators, normally produces lower average individual error rates. There are, however, many situations where a group containing bMIs or wMIs, in place of non-imitators, generates higher group error rates. We also observed that, in select cases, peer imitators of comparable unreliability can achieve lower average individual error rates and lower average group error rates. This pattern is clearest when we compare a population composed wholly of truth-biased peer imitators with independent unreliability, u, with a population of bMIs with identical independent unreliability. In this case, the bMIs will imitate other bMIs whose independent unreliability is no less than their own, resulting in no decrease in individual unreliability but a precipitous decrease in diversity. The same effect occurs with wMIs, but to a lesser extent. One obvious modification of weighted meta-induction that would improve its performance in a variety of cases would be to increase the range of players to which the wMI assigns weight in determining her predictions.

Call a player a 'cautious weighted meta-inductivist', if she assigns weight to each accessible player provided, roughly, that it is 'not likely' that that player is less reliable than the cautious weighted meta-inductivist. There are several ways to make precise what *not likely* amounts to in the present context. Schematically we want a cautious weighted meta-inductivist to assign weight to another player's prediction provided that player's score is at least $s_c - f(n)$, where s_c is the relevant cautious weighted meta-inductivist's success rate, n is the round number, and f is a function, such that $f(n)$ goes to zero as n goes to infinity. (The attractivity of an accessible player P would similarly be redefined as: $\text{at}_n(\text{P}) = \text{suc}_n(\text{P}) + f(n) - \text{suc}_n(\text{wMI})$, provided $\text{suc}_n(\text{P}) > \text{suc}_n(\text{wMI}) - f(n)$, otherwise $\text{at}_n(\text{P}) = 0$.) For the sake of simplicity, we assume that $f(n) = 1/(\log_2(n) \cdot 3)$. This assumption is somewhat arbitrary, but delivers the reasonable result that $f(2) = 1/3$, $f(10) \approx 0.10$, $f(100) \approx 0.05$, and $f(1000) \approx 0.03$. We denote this species of meta-induction "w_cMI", and note that under the conditions described in Theorem 2 (above), $\text{suc}_n(\text{w}_\text{c}\text{MI})$ approximates the maximal success rate of non-MI-players as n approaches ∞. Under a broad range of conditions, w_cMIs also achieves better individual performance than non-imitating players and peer imitators, while at the same time contributing to high group scores. The following tables summarizes some simulations that we ran (which bear out the preceding claims).

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% w_cMI (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.0005	0.0005	0.0240	0.4040	0.4955	0.9996	0.0056	0.0084	0.0109	0.0598	0.2412	0.2507
average global group error	0.0000	0.0000	0.0240	0.4040	0.4900	1.0000	0.0051	0.0076	0.0099	0.0545	0.2194	0.2506

Table 19

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	2	2	2	2	2	2	2	2	2	2	2	2
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% w_cMI (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent B (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.0105	0.0104	0.0271	0.0793	0.0793	0.1092	0.0152	0.0177	0.0205	0.0628	0.1951	0.1546
average global group error	0.0000	0.0000	0.0180	0.0760	0.0760	0.1250	0.0048	0.0071	0.0096	0.0482	0.1622	0.1445

Table 20

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% w_cMI (unreliability)	100 (0.47)	100 (0.48)	100 (0.49)	100 (0.499)	100 (0.4999)	100 (0.55)	100 (0.5)	100 (0.75)	100 (1)	100 (5)	100 (20)	100 (i)
average individual error	0.3702	0.4163	0.4570	0.4950	0.5006	0.6431	0.0597	0.0901	0.1212	0.6867	3.8292	0.2528
average group error	0.0020	0.0140	0.0800	0.4490	0.5080	0.9950	0.0051	0.0078	0.0102	0.0593	0.3416	0.2516
average local group error	0.3559	0.4139	0.4558	0.4966	0.4987	0.6499	0.0560	0.0849	0.1127	0.6440	3.5151	0.2530

Table 21

Comparing the performance of w_cMIs to that of wMIs, we see that w_cMIs approximate the success rates of wMIs, in those cases where the performance of weighted meta-induction is the best (i.e., cases where the group contains a sub-population of highly reliable independent players). When accessibility is universal, w_cMIs approximate success rates of non-imitators, in those cases where the performance of non-imitation is the best (i.e., cases where the group does not contain a sub-population of highly reliable independent players). When accessibility is restricted, w_cMIs approximate the success rates of peer-imitation, in those cases where the performance of peer-imitation is the best (i.e., cases where the group does not contain a sub-population of highly reliable independent players).

# of rounds	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
# of cycles	10	10	10	10	10	10	10	10	10	10	10	10
event type	binary	binary	binary	binary	binary	binary	real	real	real	real	real	real
% w_cMI (unreliability)	90 (0.47)	90 (0.48)	90 (0.49)	90 (0.499)	90 (0.4999)	90 (0.55)	90 (0.5)	90 (0.75)	90 (1)	90 (5)	90 (20)	90 (i)
% independent (unreliability)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)	10 (0.1)
average individual error	0.1100	0.1128	0.1160	0.1198	0.1171	0.1290	0.0660	0.0797	0.0880	0.1312	0.1720	0.1045
average group error	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.0045	0.0058	0.0064	0.0090	0.0108	0.0890
average local group error	0.0947	0.0991	0.1034	0.1064	0.1077	0.1194	0.0546	0.0668	0.0752	0.1109	0.1413	0.0956

Table 22

The simulations just described characterize the performance of w_cMIs. The feature of a w_cMI that distinguishes her from a wMIs is her tendency to attach some credence to the judgment of every agent who has not distinguished himself as being obviously less reliable than the w_cMI. This feature of w_cMIs inhibits the tendency of meta-inductive prediction strategies to eliminate diversity within the judgments of a group's members (and thereby the wisdom of the group) without the benefit of a compensating increase in the reliability of the meta-inductivist. Note that the good performance of w_cMIs is independent of the kind of predictions considered, whether binary-valued (e.g., predicting whether a given stock price will be up on the following day), or real-valued (e.g., predicting what the price of a given stock will be at the close of trading on the following day).

Despite being an attractive prediction method, there are some conditions under which w_cMI fails to perform as well as non-imitating players and peer imitators in making a contribution to the accuracy of a group. This will occur when a group of truth-biased w_cMIs with independent unreliability, u, find themselves within a population that contains a small number of truth-biased agents with independent unreliability, s, where s is somewhat greater than u.

9. Conclusion

Much recent formal and empirical work on the Wisdom of Crowds has extolled the virtue of independent and diverse judgment as essential to the maintenance of 'wise crowds'. In contrast, recent work by Schurz demonstrates the optimality of meta-induction as a method for predicting unknown events and quantities. Inasmuch as meta-induction is an imitative prediction method whose application reduces diversity among the predictions of a group, the application of meta-induction may have a negative effect on the accuracy of the average of a crowd's judgment, in cases where the average of the crowd's judgment is not accessible to any of the individual members of the crowd.

Both diversity and individual accuracy contribute to the accuracy of a crowd's judgment, and decreases in diversity as a result of meta-induction can result in simultaneous increases in average individual accuracy and in the accuracy of the crowd's judgment. To pursue the issue further, we simulated a variety of situations, and observed the costs and benefits of applying meta-inductive predictive methods. In all of the simulations that we ran, we implemented the assumption that the accuracy of the independent judgments of each player was stochastically independent of the accuracy of the other players' judgments. This assumption is quite unrealistic, and is favorable to the performance of non-imitation in the generation of accurate group judgments, and to the performance of peer-imitation in the generation of accurate individual and group judgments.

In the case of a group whose members have identical independent reliabilities (and the discussed stochastic independence assumption is made), non-imitation is best at producing a wise crowd, though cautious weighted meta-inductivists are approximately as good. Even in a group whose members have identical independent unreliabilities, meta-inductive prediction methods, such as weighted meta-induction and cautious weighted meta-induction, generate lower individual error rates, since such meta-inductivists tend to predict an average of the values predicted by at least some of their peers (and thereby harness the wise crowd effect).

In the case of groups containing agents with varied independent unreliabilities, all meta-inductive methods are highly useful to the individual. From the crowd's view, imitate-the-best meta-induction is the least attractive meta-inductive method, since it does not harness any form of wise crowd effect. Weighted meta-induction and cautious weighted meta-induction do harness such an effect, and thereby generate lower individual and groups error rates, in some cases. Across the range of cases that we considered, cautious weighted meta-induction reigns supreme. This method simultaneously approximates the strengths of non-imitation, peer-imitation, and weighted meta-induction.

Bibliography

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

Douven, I./A. Rieger (2010), Extending the Hegselmann-Krause Model I, in *Logic Journal of the IGPL* 18(2), 323–335

Gigerenzer, G. et al. (1999), *Simple Heuristic That Make Us Smart*, New York

Hartmann, S./C. Martini/J. Sprenger (2009), Consensual Decision-Making among Epistemic Peers, in *Episteme* 6, 110–129

Krogh, A./J. Vedelsby (1995), Neural Network Ensembles, Cross Validation, and Active Learning, in: Tesauro, G./D. Touretzky/T. Leen (eds.), *Advances in Neural Information Processing 7*, Cambridge

Lorenza, J./K. Rauhutb/F. Schweitzera/D. Helbing (2011), How Social Influence Can Undermine the Wisdom of Crowd Effect, in: *Proc Natl Acad Sci USA* 108, 9020–9025

Norton, J. (2003), A Material Theory of Induction, in: *Philosophy of Science* 70, 647–670

Page, S. (2007), *The Difference—How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton

Reichenbach, H. (1949), *The Theory of Probability*, Berkeley

Rendell, L. et al. (2010), Why Copy Others? Insights from the Social Learning Strategies Tournament, in: *Science* 328, 208–213

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

— (2009a), Meta-Induction and Social Epistemology: Computer Simulations of Prediction Games, in: *Episteme* 6, 201–220

— (2009b), Meliorative Reliabilist Epistemology: Where Externalism and Internalism Meet, in: *Grazer Philosophische Studien* 79, 41–62

— (2009c), Meta-Induction. A Game-Theoretical Approach, in: Glymour, C. et al. (eds.), *Logic, Methodology and Philosophy of Science*, London, 241–266

— (2012a), Meta-Induction and the Problem of Fundamental Disagreement, to appear in: Jäger, C./W. Löffler (eds.), *Epistemology: Contexts, Values, Disagreement*, Frankfurt

— (2012b), Meta-Induction in Epistemic Networks and Social Spread of Knowledge, in: *Episteme* 9(2), 151–170

Surowiecki, J. (2004), *The Wisdom of Crowds*, New York