Wise Crowds, Clever Meta-Inductivists

Paul D. Thorn

Abstract Formal and empirical work on the Wisdom of Crowds has extolled the virtue of *diverse* and *independent* judgment as essential to the maintenance of 'wise crowds'. In other words, communication and imitation among members of a group may have the negative effect of decreasing the aggregate wisdom of the group. In contrast, it is demonstrable that certain meta-inductive methods provide optimal means for predicting unknown events. Such meta-inductive methods are essentially imitative, where the predictions of other agents are imitated to the extent that those agents have proven successful in the past. Despite the (self-serving) optimality of meta-inductive methods, their imitative nature may undermine the 'wisdom of the crowd', since these methods recommend that agents imitate the predictions of other agents. In this paper, I present a replication of selected results of Thorn and Schurz, illustrating the effect on a group's performance that may result from having members of a group adopt meta-inductive methods. I then expand on the work of Thorn and Schurz by considering three simple measures by which meta-inductive prediction methods may improve their own performance, while simultaneously mitigating their negative impact on group performance. The effects of adopting these maneuvers are investigated using computer simulations.

Keywords meta-induction, the wisdom of crowds, judgment aggregation, social epistemology

1 Introduction

In several recent papers (2008, 2009b), Gerhard Schurz proposed a response to Hume's problem of induction, based on *meta-induction*. In its various forms, meta-induction proceeds by considering the past track record of other agents (and/or prediction methods), and makes predictions of future events by reasoning that agents (and/or prediction methods) that have been successful in the past will be successful in the future. Schurz demonstrated that, under plausible 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.

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 prescriptions of meta-induction are thus in tension with prescriptions implicit in recent formal and empirical work on the Wisdom of Crowds. Such work emphasizes the importance of agents making their predictions (and 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 livestock exhibition could observe a mature ox, and had the opportunity to guess its

weight. 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 'judgment of the crowd'. The crowd guessed that the ox would weigh 1,197 pounds, while the ox weighed 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 (Surowiecki 2004, 5-22; Page 2007, 178). It is also straightforward to construct formal models of individual judgment wherein the average value of the judgments of a group tends to be very accurate. Recent empirical studies also show that the accuracy of a crowd's judgment can be compromised when agents within the group are aware of the judgments made by other group members (and are thus able to *imitate* 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 (conditional on the value of the predicted event). So select empirical and formal results *suggest* that imitating the judgments of other group members is, *contra* meta-induction, a bad thing.

In a recent paper, Thorn and Schurz (2012) presented results concerning the impact on a group's performance that may result from having members of a group adopt meta-inductive methods (cf. Schurz 2012). In this paper, I replicate a selection of those results, illustrating that, in a variety of circumstances, the adoption of meta-inductive methods can decrease the accuracy of the aggregate judgment of the group. I then expand on previous work by considering three simple measures by which meta-inductive prediction methods may improve their own performance, while simultaneously mitigating their negative impact on the aggregate judgment of the group.

2 The Optimality of Global Meta-Induction

To demonstrate the optimality of meta-induction, Schurz (2008, 2009b) introduced the notion of a *prediction game*, consisting of:

(1) An infinite sequence (e) = $(e_1,e_2,\ldots)$ of events, whose values are drawn from the unit interval, i.e., $e_n \in [0,1]$, for each round, n, of the game (and from {0,1} in the case of *binary* prediction games).

(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 $n-1$. The players in Π include: one or several meta-inductivist players of various kinds (see below), and a finite set of non-MI-players players $P_1,\ldots,P_m$. It is assumed that the MI-players make their predictions after the non-MI-players, and may thus imitate the predictions of the non-MI-players.

Within prediction games, the deviation of a prediction p_n from the event e_n is measured by a normalized loss function $l(p_n,e_n) \in [0,1]$. A prominent loss-function measures the absolute difference between event and prediction, $|e_n - p_n|$, but the optimality theorems described below are not restricted to this loss function: Theorem 1 holds for *monotonic* loss-functions, and theorem 2 holds for *convex* loss-functions. The *score*, $s(p_n,e_n)$, obtained in round n is defined as $1-l(p_n,e_n)$. The *success rate*, $suc_n(P)$, of player P, at time n, is $(\Sigma_{1\leq i\leq n}\, s(p_i(P),e_i)/n$. Finally, $maxsuc_n$ is the maximal success rate of the non-MI-players at time n.

The simplest type of meta-induction is called "imitate-the-best". In each round, bMIs (players who employ imitate-the-best meta-induction) 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 achieves a higher success-rate. If there are several best players, bMIs chooses her favorite by a predefined

ordering of the non-MI-players. The central result concerning imitate-the-best prediction method is as follows:

Theorem 1 (Schurz 2008): For every prediction game ((e),{$P_1,\ldots,P_m$,bMI}) that contains a *best* non-MI-player, B, after some round n_B (i.e., $suc_n(B) > suc_n(P_i)$ for all $n \geq n_B$ and $P_i \neq B$), the following holds:
(1.1) *Short run*: For all rounds n, $suc_n(bMI) \geq maxsuc_n - (n_B/n)$.
(1.2) *Long run*: As n approaches ∞, $suc_n(bMI)$ converges to $maxsuc_n$.

The assumption of Theorem 1 that there is a *best* non-MI-player, B, after some finite number of rounds is rather strong. A satisfactory solution to Hume's problem calls for a meta-inductive strategy whose performance is optimal when this assumption does not hold. *Weighted meta-induction* fills this role. wMIs (players who employ weighted meta-induction) predict a weighted average of the predictions of the so-far 'most attractive' players. The attractivity $at_n(P)$ of player P at time n is P's surplus success-rate compared to the wMI's success: $at_n(P) = suc_n(P) - suc_n(wMI)$, provided $suc_n(P) > suc_n(wMI)$, otherwise $at_n(P) = 0$. A wMI's predictions are defined as $p_{n+1}(wMI) = \sum_P(at_n(P)\cdot p_{n+1}(P))/\sum_P(at_n(P))$, where P ranges over all accessible players. (If no player has positive attractivity, the wMI makes a random guess.) The following establishes weighted meta-induction's long-run optimality:

Theorem 2 (Schurz 2008, cf. Cesa-Bianchi and Lugosi 2006): For every real-valued prediction game ((e),{$P_1,\ldots,P_m$,wMI}) whose loss-function $l(p_n,e_n)$ is *convex* in the argument p_n, the following holds:
(2.1) *Short run*: $\forall n \geq 1$: $suc_n(wMI) \geq maxsuc_n - \sqrt{(m/n)}$.
(2.2) *Long-run*: As n approaches ∞, $suc_n(wMI)$ converges to $maxsuc_n$.

Theorem 2 does not apply directly to binary prediction games, because a wMI's predictions are real-valued. However, theorem 2 can be generalized to binary valued predictions, by positing 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 equal to P's attractivity. The *mean success rate* of such populations approximates the maximal success rate of the most attractive non-MI-players, with an additional maximal short-run loss of $1/(2\cdot k)$ (Schurz 2008, 2009a, 2009b). Similar convergence results hold for the *expected* success-rate a meta-inductivist who predicts respective outcomes with probability equal to the population shares represented among such groups.

3 The Wise Crowd

Following in the footsteps of some recent monographs (Surowiecki 2004; Page 2007, 179), I will say that the *judgment of a crowd* with respect to *query* is the average response of its members (rounded if necessary), treating affirmation as *one* and disaffirmation as *zero*, in the case of binary predictions. Along with the preceding convention, I say that *a crowd is wise* to the extent that its judgments are *accurate*.

Anecdotes such at Francis Galton's (section 1) are relatively widespread, and it is clear that the judgment of a crowd with respect to some kinds of query frequently exhibits uncanny accuracy (i.e., the wise crowd effect). In this same vein, Jack Treynor illustrated the wise crowd effect by having groups of students guess the number of jelly beans contained in a large jar (Surowiecki 2004, 5). 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 have demonstrated the accuracy of groups of independently acting individuals in making various kinds of prediction (Surowiecki 2004, 17-22; Page 2007, 178).

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 ($r > 0.5$) that that member of the group will make a correct judgment regarding the truth value of some proposition. 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 majority response of its members is correct converges to *one* as the size of the group approaches ∞.

The Weak Law of Large Numbers suggests an obvious means of modeling wise crowds in the case of real-valued events. Where ε is any number greater than 0, and X_n is a sample of n 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 of 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, for all $\varepsilon>0$, that the probability that the group's judgment about the value of μ differs from μ by more than ε goes to zero as 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 real-valued queries.[1]

While the wise crowd effect recommends that forecasters make their predictions independently, the optimality of meta-induction suggests that forecasters should imitate the most successful forecasters whose predictions are accessible. So there is a tension between the preconditions for wise crowds, and the injunctions of meta-induction. In the face of this tension, Thorn and Schurz (2012) evaluated the impact on group performance that may result from having members of a group adopt meta-inductive methods. Their results illustrate a variety of conditions under which *replacing* non-imitative players by meta-inductivists reduces the accuracy of the aggregate judgment of the group. After introducing the formal framework of (Thorn and Schurz 2012), I summarize some of their results. I then consider three simple measures by which meta-inductive prediction methods may improve their own performance, while simultaneously mitigating their negative impact on group performance.

4 The Formal Setup

Departing slightly from the prediction games described in section 2, the simulations described here include the following elements:

(1) A quadratic grid consisting of 100×100 = 10,000 cells. Each cell corresponds to an individual player.

[1] The independence assumptions under which the two theorems apply limit the applicability of the corresponding models. A more realistic formal model is found in (Page 2007), drawing on work by Krogh and Vedelsby (1995).

(2) For some simulations, each agent has *access* to the success rates and the present judgment of every other player. In other simulations, each player only has access to information concerning the players in her *Moore-neighborhood*, i.e., to herself and the eight immediately surrounding players.

(3) The event sequence is either: a random sequence of values chosen according to a uniform probability distribution on the unit-interval [0,1], or a binary event sequence generated by rounding the elements of a sequence of the preceding sort, where values greater than 0.5 are rounded to 1. 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 (thereby permitting the possibility of arbitrarily large errors).

(4) In the case of a binary event sequence, each player has a predefined *independent reliability*, r, which is the player's probability of making a correct prediction in any given round (as determined by Bernoulli trials), assuming she bases her predictions solely on her own abilities, and independently of other players. Each player's *independent unreliability*, u, is $1-r$. In the case of a real-valued sequence, each player's prediction is assumed to be normally distributed with a mean identical to the true event-value, where the mean absolute deviation is the player's *independent unreliability*, u.[2]

(5) 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 apply one of the following imitative prediction methods to other *accessible* players:

(a) Weighted meta-induction wMI[3]: 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, wMIs predict the *rounded* attractivity weighted average of the predictions of those players accessible to the wMI. In face of both real-valued and binary event sequences, wMIs predict by independent means in the first round, in the first cycle of each round, and whenever they themselves have the highest success rate.

(b) Peer-imitation: Peer-imitators predict an *unweighted* average of the predictions of those players accessible to the peer-imitator.

In contrast to the sort of prediction games described in section 2, the present setup allows for mutual imitation between imitative players. 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 players may imitate 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 abilities. In all following cycles, independent players repeat their initial prediction, while imitative players apply their imitative prediction method to the predictions made by accessible players in the previous cycle. This continues until a preselected maximum number of cycles is reached. 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.

---

[2] So the standard deviation of an agent's independent guess is $u\cdot\sqrt{(2/\pi)}$, since $\sqrt{(2/\pi)}$ is the ratio of the mean absolute deviation to the standard deviation in the case of normal distributions.

[3] As space is limited, I focus exclusively of weighted meta-induction, which performed better than imitate-the-best meta-induction in the simulations studied in (Thorn and Schurz 2012).

5 Groups with Universal Accessibility

In the present section, I replicate results from (Thorn and Schurz 2012) in order to illustrate some effects which ensue when members of a group adopt weighted meta-induction, in cases where the predictions and success rates of all agents are accessible to all agents. Figure 1 presents results for binary predictions, comparing populations composed wholly of independent predictors to ones composed wholly of wMIs, in games that lasted 1000 rounds, with 2 cycles per round. The independent *unreliability* of the agents is the independent variable. The respective mean *error rate* (i.e., the mean linear distance between predicted and actual values) is the dependent variable.

Fig. 1
Binary Events,
Universal Access,
No Experts

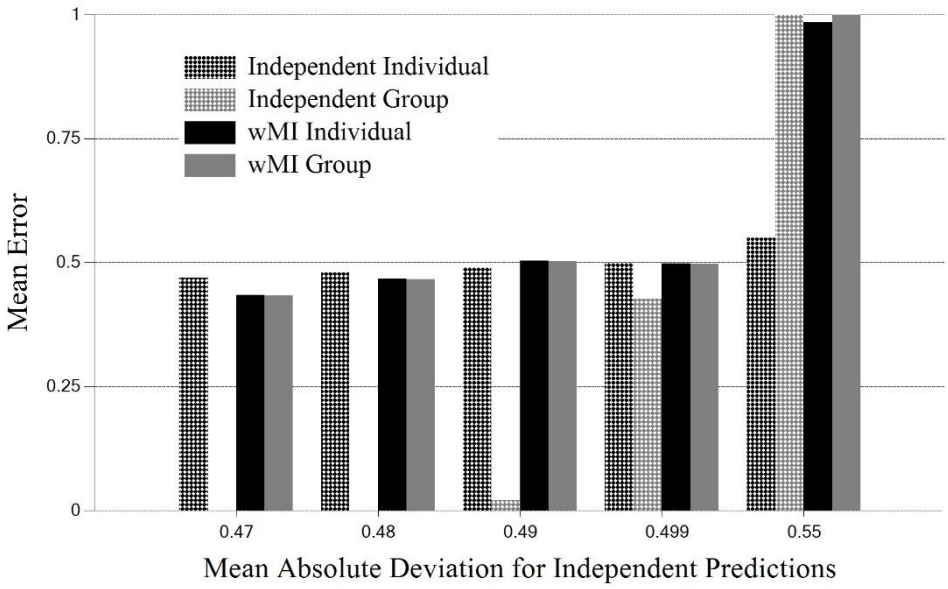
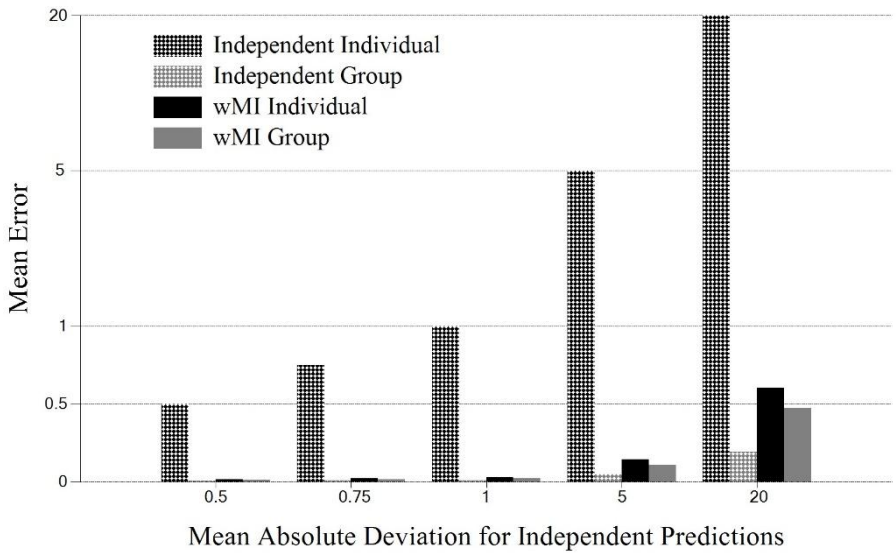


Figure 2 presents results analogous to figure 1 for real-valued predictions. Note that for all the figures concerning real-valued predictions, the scale for the intervals [0,1], [1,5], and [5,20] differs in order to magnify small differences in the values of the dependent variable that occur particularly in the interval [0,1].

Fig. 2
Real-Valued Events,
Universal Access,
No Experts



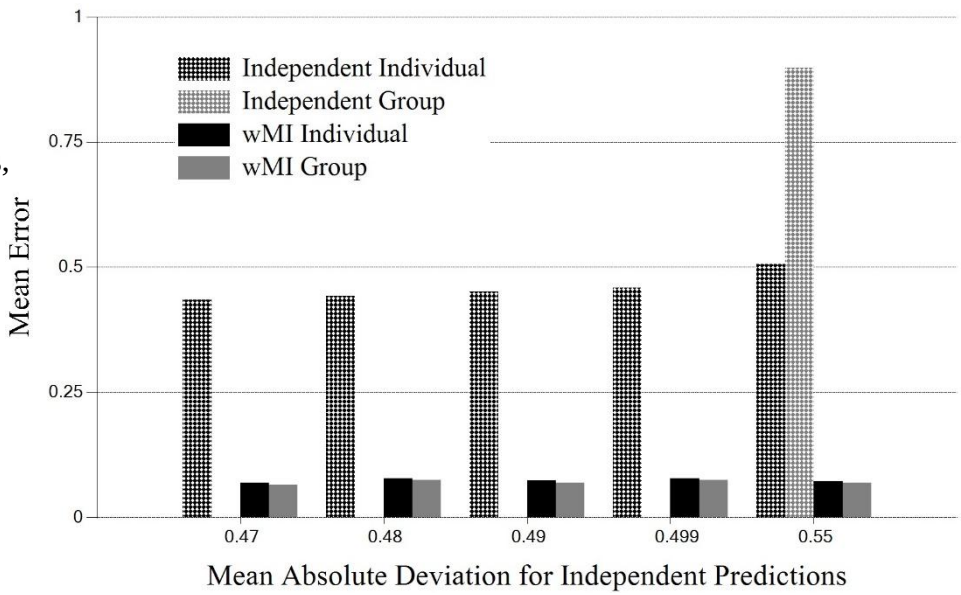
The results represented in figures 1 and 2 reflect the fact that the mean individual error rates of non-imitators converge to their independent unreliabilities. We also observe (as predicted by the Condorcet Jury Theorem and the Law of Large Numbers) that the mean group error for non-imitators is usually quite low. 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 (while 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 $u = 20$, is sufficient to achieve a relatively low mean group error.

Noteworthy patterns characterize groups composed wholly of wMIs. In the case of binary prediction games, where the independent unreliability, *u*, is less than 0.5, we observe no wise crowd effect. In these cases, the populations of wMIs tend to reach an equilibrium state where a single wMI is distinguished as 'most successful', while the remaining wMIs have *identical* success rates, with the result that each prediction made by members of the group is identical to the prediction of the wMI who is most successful. These cases are in contrast to the binary case where *u* is greater than 0.5, where we observe a '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, wMIs exhibit a wise crowd effect in the case of real-valued event sequences. The effect is slightly weaker than in the case of independent predictors, since some diversity is lost through imitation. The reward for the increase in mean group error is a large decrease in mean individual error.

The scenarios represented in figures 1 and 2 are unrealistic in assuming that all of the predictors in the group are equally reliable. This unrealistic assumption is biased against wMIs, whose *strength* consists in imitating the predictions of those predictors whose predictions are the most accurate. The simulations represented by the following figures differ from the ones represented by figures 1 and 2, by including a 10% subpopulation of (expert) independent predictors, with an independent unreliability of 0.1. Note that the independent variable is the independent unreliability of the *non-expert* members of the population, while the mean error rates are derived from the predictions of both experts and non-experts.

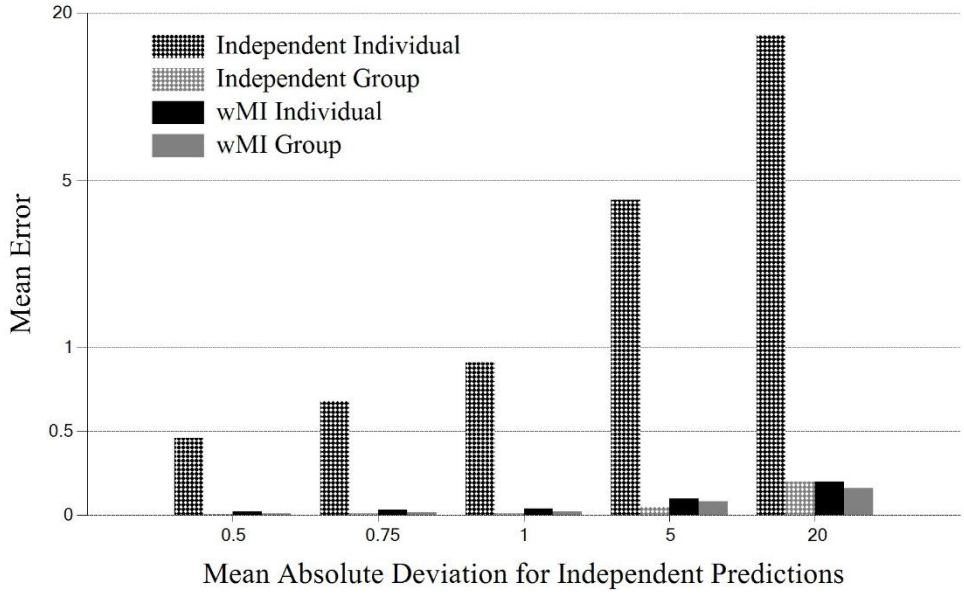
Fig. 3
Binary Events, Universal Access, 10% Experts



Once again, non-imitators have mean individual error rates that converge to their mean independent unreliabilities (which is higher due to the inclusion of subpopulation of highly reliable experts). The impact on the mean individual error rate is much greater when we replace the less reliable subgroup of non-imitators with wMIs: The prediction strategy of the wMIs yields the result that the individual error rates of the wMIs approximates the independent unreliability, $u = 0.1$, of members of the highly reliable subgroup (which translates into low group error rates). More generally, applying weighted meta-induction results in lower individual error rates (as compared to independent predictors), so long as we *assume* the wMIs have the opportunity to imitate truth-biased

predictors whose independent reliability exceeds their own. The present assumption is typically plausible.

Fig. 4
Real-Valued Events, Universal Access, 10% Experts



6 Groups with Restricted Accessibility

In the present section, I replicate results from (Thorn and Schurz 2012) in order to illustrate some effects that ensue when members of a group adopt weighted meta-induction, in cases where players only have access to the predictions and success rates of agents in their Moore-neighborhood. In all of the simulations considered in this section, each game lasted 1000 rounds, and had 10 update cycles per round. Within these simulations, the performance of peer imitation is compared with weighted meta-induction. Peer-imitation has some connection to the wise crowd phenomena inasmuch as the predictions of a peer-imitator will be identical to the judgment of the group, in the case where the peer-imitator has access to the judgments of all members of a group. The added effect of peer-imitation over non-imitation, in the 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 figures 5-8) 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 where the group's accuracy is poor. Peer-imitation also has a small effect in decreasing the diversity of the group, and thereby on the accuracy of the judgments of the group (in comparison to non-imitation).

Figure 5 presents results for binary prediction games, comparing populations composed wholly of peer imitators to ones composed of wMIs. Figure 6 presents results analogous to figure 5 for real-valued predictions.

In cases where the independent unreliability, u, of all players is identical and truth-biased (i.e., in all the cases described in figures 5 and 6, save the case where u=0.55), peer-imitators perform at least as well as wMIs, with respect to individual and group error rates. As it turns out, peer-imitators (within groups of peer-imitators) are incredibly adept in pooling the independent predictions of players with whom they do not have direct access, by taking the average of the predictions neighbors, who took the average of the predictions neighbors, etc. The performance of the wMIs is considerably improved in the case where the wMIs have the opportunity to imitate players with high independent reliabilities. Figures 7 and 8 presents results analogous to figure 5 and 6, save that 10% of the population consists of (expert) independent predictors, with u=0.1.

Fig. 5
Binary Events,
Limited Access,
No Experts

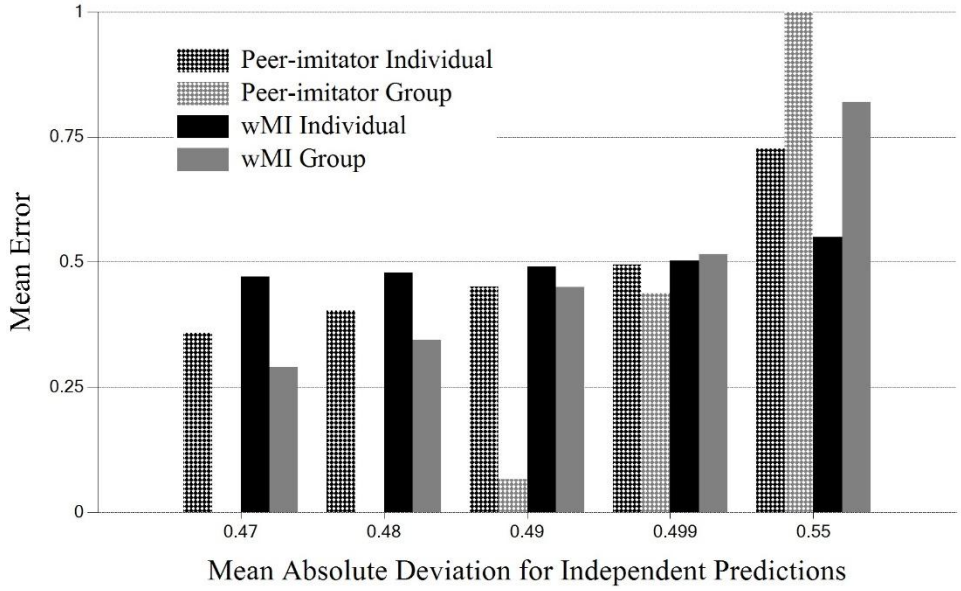


Fig. 6
Real-Valued Events,
Limited Access,
No Experts

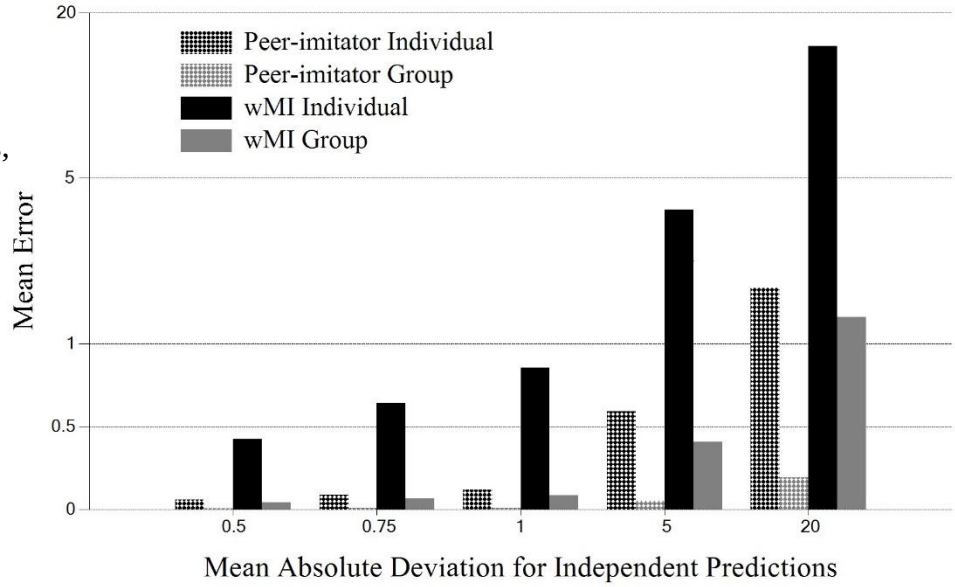


Fig. 7
Binary Events,
Limited Access,
10% Experts

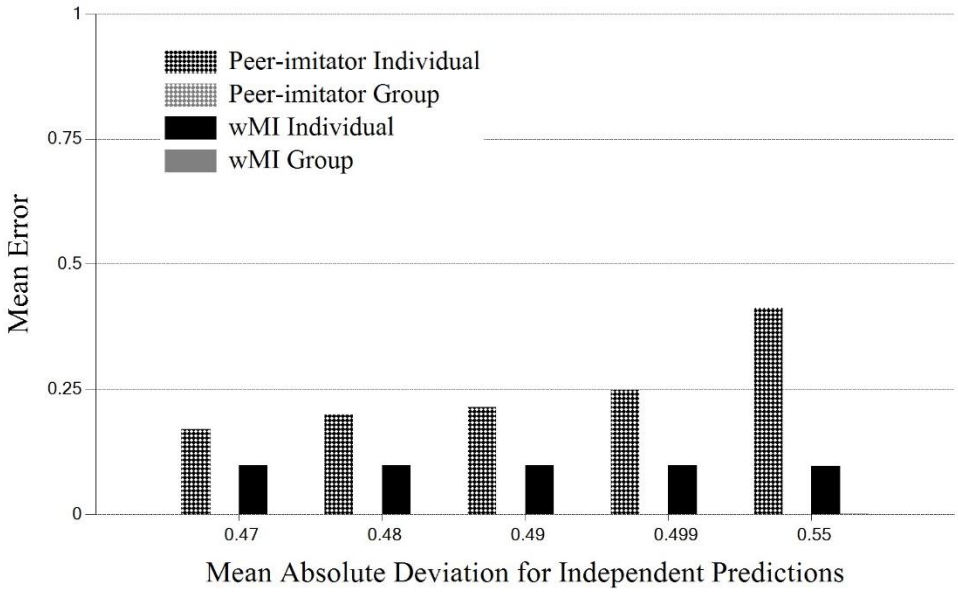
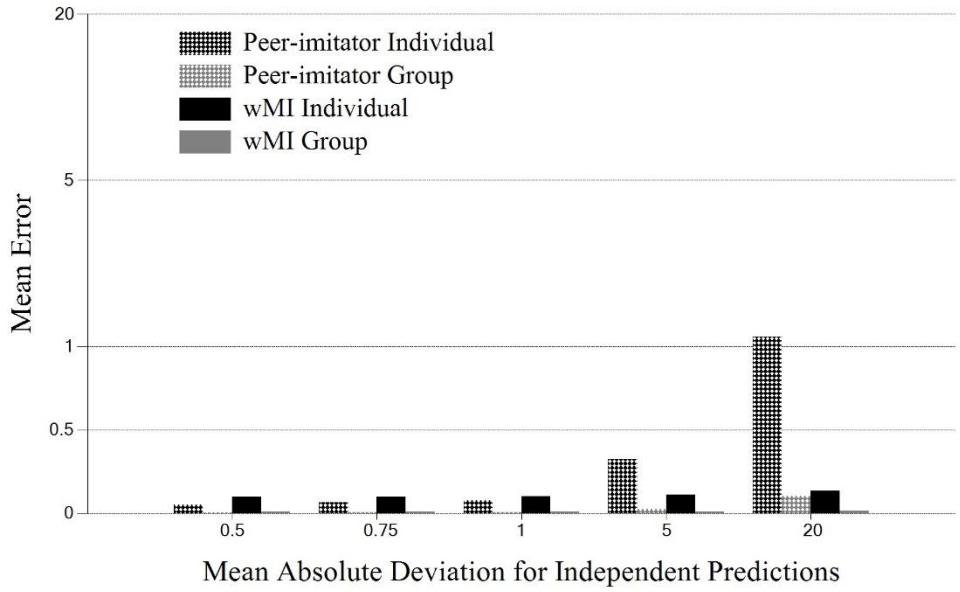


Fig. 8
Real-Valued Events,
Limited Access,
10% Experts



In all the cases described in figures 7 and 8, where we include a subpopulation of expert predictors, we observe that the performance of wMIs matches or exceeds that of peer-imitators.

7 Meta-Induction with Safeguards

While wMI performed well in many of the simulations conducted in (Thorn and Schurz 2012), it performed poorly in others: In binary games with universal accessibility, the mean group error for wMIs was significantly larger than that of independent predictors, in cases where the independent unreliability of the wMIs was low. The mean individual error rates for wMIs were also greater, in the case of binary prediction games where the independent unreliability of the wMIs was high, in the absence of experts. When accessibility was limited, the mean individual and group error rates for wMIs were greater than those of peer-imitators, in all cases where experts were absent and the independent unreliability of the wMIs was low.

I here propose three measures in an attempt to improve the performance of wMIs.[4] The first two measures involve the inclusion of 'virtual' players within the player set available to wMIs as a basis for imitation. First, the MI-players studied in this section, wMI*s, include, as an imitable player, a player that predicts the unweighted average of the predictions of all non-virtual players accessible to the respective wMI*. Second, in the case of binary prediction games, wMI*s consider, as imitable players, the 'inverse' player of each non-virtual players that is accessible to the respective wMI*, where an inverse player always predicts of the opposite of her respective 'non-inverse'. While the former maneuver reflects a self-conscious awareness of the wise crowd phenomena (and attempt to harness it), the latter maneuver aims to make a virtue of systematic error (in the case where $u > 0.5$). The third measure employed by wMI*s is to base their attractivity weights for accessible players on the predictions made in the second to last cycle of each round, which accords with the fact that wMIs are not actually able to imitate the predictions made in the final cycle. The following figures are analogous to figures 1, 2, 5, and 6, and illustrate the effect of the three measures, within those simulations that were the most difficult for regular wMIs, i.e., those

[4] An alternate variant of weighted meta-induction is considered in (Thorn and Schurz 2012). The variant considered here performs significantly better in several of the situations considered in (Thorn and Schurz 2012). It is also possible to construct situations where the variant from (Thorn and Schurz 2012) performs significantly worse than regular weighted meta-induction, which is not the case for the variant considered here.

Fig. 9
Binary Events,
Universal Access,
No Experts

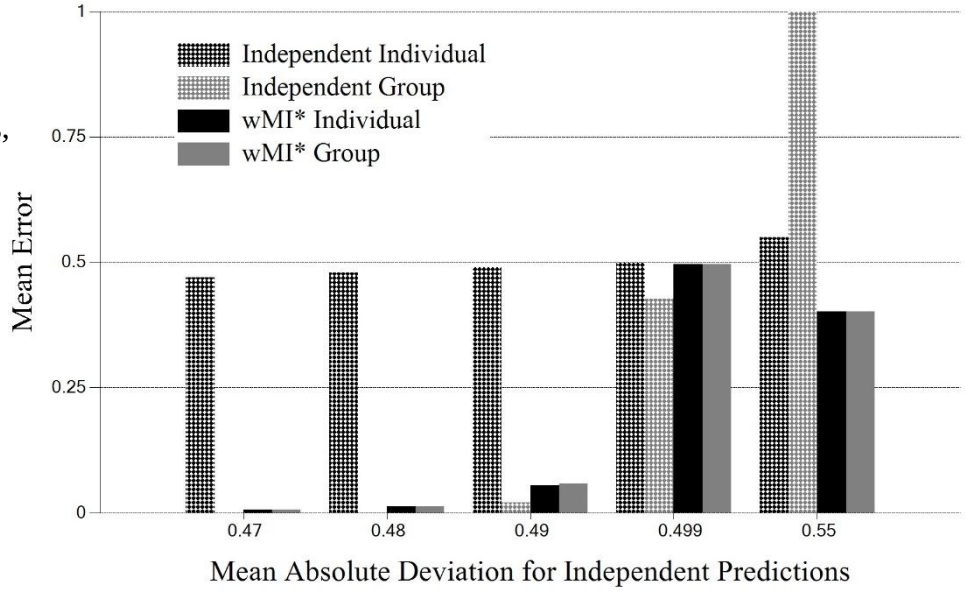


Fig. 10
Real-Valued Events,
Universal Access,
No Experts

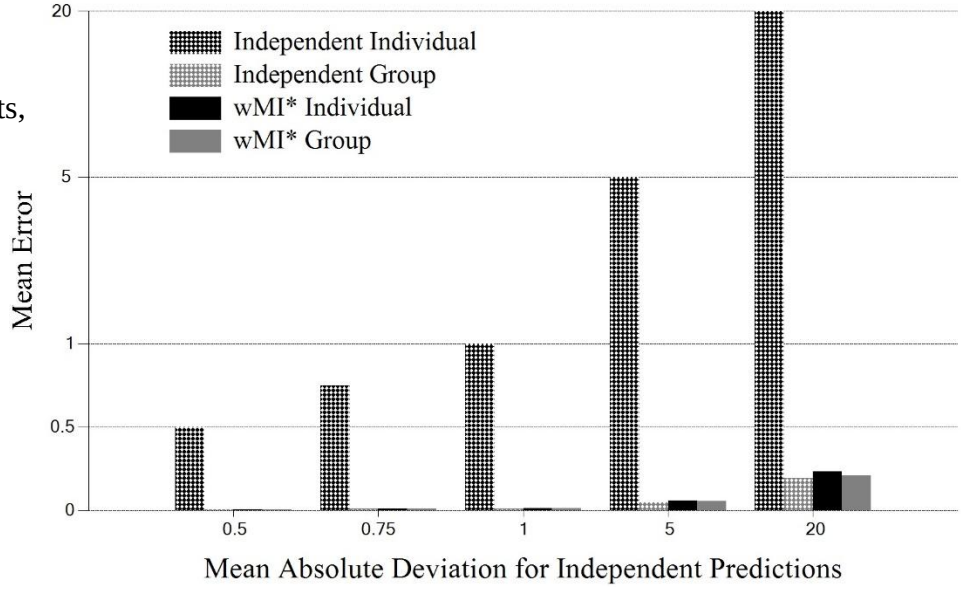
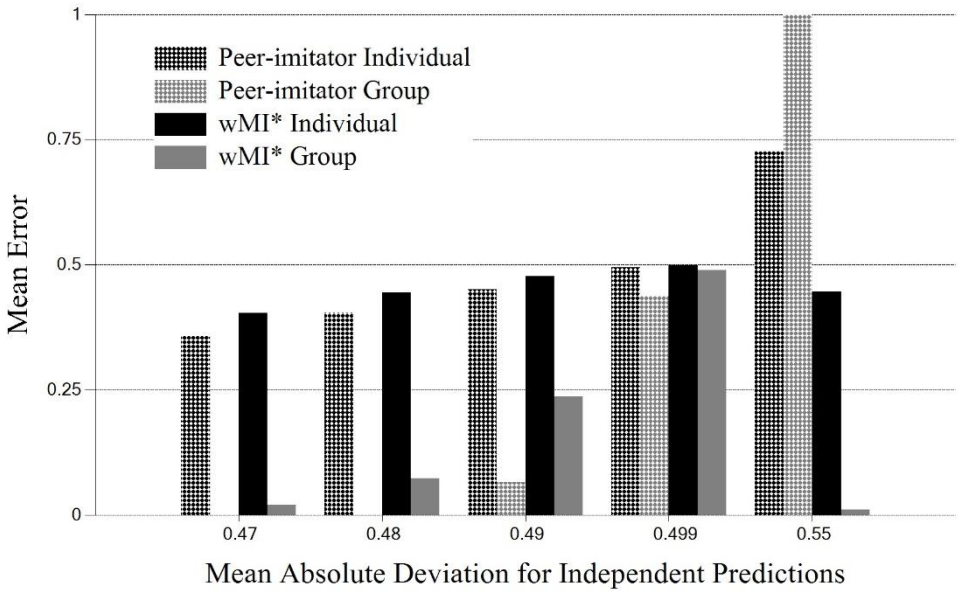
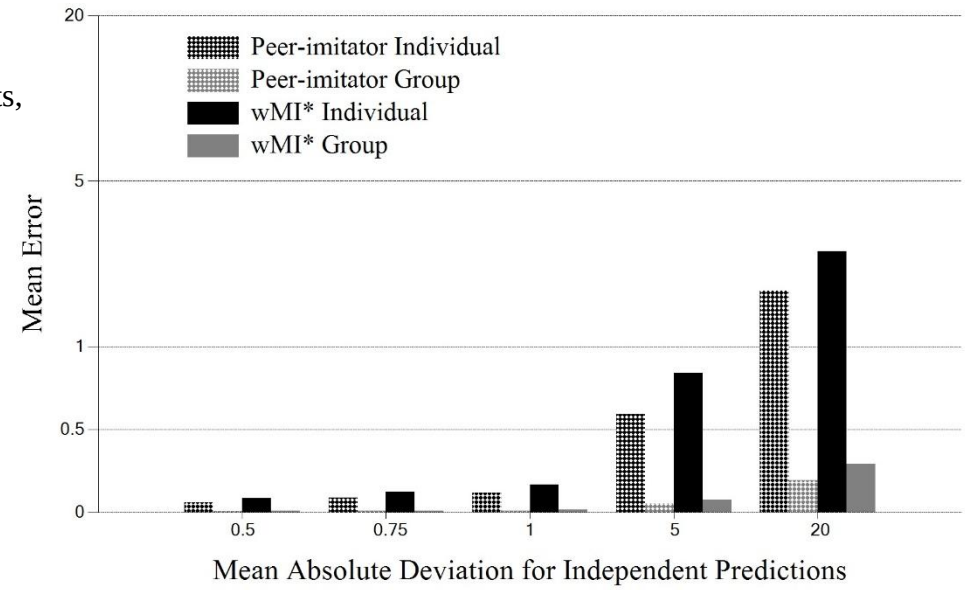


Fig. 11
Binary Events,
Limited Access,
No Experts



situations where the population contained no highly reliable experts. The performance in simulations that include experts is similar to regular weighted meta-induction.

Fig. 12
Real-Valued Events,
Limited Access,
No Experts



In all of the situations considered here, wMI*s performed better or (almost) as well as independent predictors, and peer-imitators, respectively. Cases where the wMI*s trailed behind their competitors are largely the result of losses earned in the early rounds of the game. While better performance could be achieved within the situations considered here, the required measures would be complicated, and would yield only marginal improvements. The performance of wMI*s will also be relatively good within variations of the situations considered here, so long as the (un)reliabilities of the independent predictions of the participating players converge to limits, at a pace commensurate to the length of the respective games.

8 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 (2008, 2009b) 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. However, as we saw in the preceding section, it is possible to safeguard meta-inductive methods by simple measures which allow meta-inductive prediction methods to improve their own performance, while simultaneously mitigating their negative impact on group performance.

References

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

Krogh, A., Vedelsby, J. (1995). Neural Network Ensembles, Cross Validation, and Active Learning. In G. Tesauro et al (Eds.) *Advances in Neural Information Processing 7*, MIT Press, Cambridge.

Lorenza, J. et al (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Science*s, 108, 9020-9025.

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

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.

Schurz, G. (2009a). Meta-Induction and Social Epistemology: Computer Simulations of Prediction Games. *Episteme*, 6, 201-220.

Schurz, G. (2009b). Meta-Induction. A Game-Theoretical Approach. In C. Glymour et al. (Eds.) *Logic, Methodology and Philosophy of Science*. College Publications, London, 241-266.

Schurz, G. (2012). Meta-Induction in Epistemic Networks and Social Spread of Knowledge. *Episteme*, 9, 151-170.

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

Thorn, P., Schurz, G. (2012). Meta-Induction and the Wisdom of Crowds. *Analyse & Kritik*, 34, 339-366.