Research Article

Optimising collective accuracy among rational individuals in sequential decision-making with competition

Richard P Mann [ID]

Department of Statistics, School of Mathematics, University of Leeds, Leeds, UK

Abstract

Theoretical results underpinning the wisdom of the crowd, such as the Condorcet Jury Theorem, point to substantial accuracy gains through aggregation of decisions or opinions, but the foundations of this theorem are routinely undermined in circumstances where individuals are able to adapt their own choices based after observing what other agents have chosen. In sequential decision-making, rational agents use the choices of others as a source of information about the correct decision, creating powerful correlations between different agents' choices that violate the assumptions of independence on which the Condorcet Jury Theorem depends. In this paper, I show how such correlations emerge when agents are rewarded solely based on their individual accuracy, and the impact of this on collective accuracy. I then demonstrate how a simple competitive reward scheme, where agents' rewards are greater if they correctly choose options that few have already chosen, can induce rational agents to make independent choices, returning the group to optimal levels of collective accuracy. I further show that this reward scheme is robust, offering improvements to collective accuracy across wide range of competition strengths, suggesting that such schemes could be effectively implemented in real-world contexts to improve collective wisdom.

Keywords

agent-based model, collective behaviour, condorcet, rational choice, wisdom of the crowd

Significance statement

The wisdom of the crowd relies on the aggregated opinions of many individuals rather than that of a single expert. This underpins the widespread use of review systems that pool the judgements of many consumers, is invoked as a justification for democratic governance, and motivates harvesting the opinions of users of social media and other online forums. However, the accuracy of collective judgements is undermined by social interactions, as individuals copy what others think. I propose a way of rewarding individuals for offering accurate opinions that removes the incentive to copy, using competition between those expressing the same opinion. If individuals maximise their rewards, this makes collective judgements much more accurate, increasing the possibilities for extracting wisdom from the crowd.

Corresponding author:

Richard P Mann, Department of Statistics, School of Mathematics, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK.
Email: r.p.mann@leeds.ac.uk

Introduction

It has long been recognised that aggregating the opinions, estimates or decisions of many individuals can give superior results compared to relying on a single individual alone (De Condorcet, 1785; Galton, 1907; Surowiecki, 2005). Such aggregation is a simple but potentially powerful example of collective intelligence, and one that acts as both a justification for democratic decision-making institutions (Landemore, 2012; List and Goodin, 2001) and a motivation for utilising fora such as social media to harness the potential of global collective knowledge (Klein, 2011).

The Condorcet Jury Theorem (CJT) (Boland, 1989; De Condorcet, 1785) demonstrates that collective accuracy, in the form of a majority vote, can far exceed individual accuracy under an idealised assumption that agents choose independently. Indeed, where two options are equally likely a priori, and signals in favour of each are equally reliable, independent majority voting is the optimal method for aggregating anonymous, dichotomous votes (Austen-Smith and Banks, 1996). While the CJT has motivated many appeals to the wisdom of the crowd (e.g. (King and Cowlishaw, 2007; List and Goodin, 2001; List, 2004)), in reality, this independence assumption is routinely violated in collective decision-making scenarios where agents are able to observe each other and use social information to motivate their own choices (Surowiecki, 2005). The wisdom of the crowd requires that a group must effectively aggregate the private information held by its members, but information cascades can result from social learning, such that within a group a large proportion of individuals simply follow the decisions made by others, without reference to any private information they may have (Bikhchandani et al., 1992). Empirical studies have demonstrated how readily humans copy the actions of others (Faria et al., 2010; Gallup et al., 2012; Mann et al., 2013), in common with other animals (Sumpter and Pratt, 2009), when those actions are readily observable. The tendency of agents to follow the decisions of others can be rational from an individual perspective (Anderson and Holt, 1997; Bikhchandani et al., 1992; Mann, 2018; Tump et al., 2020), but such self-reinforcing cascades of social information can cause very large scale errors in collective judgement, as illustrated anecdotally in the historical examples given by Mackay in 'Extraordinary Popular Delusions and the Madness of Crowds' (Mackay, 1841). Scientific study also suggests that, under controlled conditions, allowing individuals to update their own beliefs in the light of observing others tends to reduce the accuracy of collective estimations (Lorenz et al., 2011), even as it increases the average accuracy of individual agents (Tump et al., 2020).

The dangers of relying on social information are highly pertinent since sequential decision-making is common across a wide variety of domains. On an individual level, we often choose what to buy, where to eat, or even how to vote based on the choices or expressed opinions of others before us.

Decisions of great importance are frequently taken after substantial discussion between those voting, such as in a jury trial, in company hiring boards and in the monetary policy committees of central banks deciding on interest rate adjustments. In each case, those voting on the decision may be influenced by others who express opinions before them. Sequential decisions may be present even when a system is designed to elicit individuals' independent decisions. Consider, for example, the case of formalised peer-review of scientific publications or grant proposals. Here, reviewers apparently provide their reviews independently, but this ignores the effect of author status, which provides a proxy for the decisions of past reviewers of the same author. Likewise, characteristics such as the fame of an individual, or the market share of a product, may serve to indicate a preponderance of past choices made, even if these are not directly observed. Advertising that points to the number of users or consumers of a product is suggestive of the influence such past decisions can have on future purchases.

Given the prevalence of sequential decision-making across many areas in which we may wish to access collective knowledge, how might we overcome its deleterious effects upon collective wisdom? One potential solution is the introduction of competition between agents who make the same choice (Hong et al., 2012; Mann and Helbing, 2017), thus penalising agents who follow others. Previous work on sequential decision-making (Arganda et al., 2012; Mann, 2018, 2020; Pérez-Escudero and De Polavieja, 2011) has assumed that rewards are independent of which choices other agents make, with such choices being useful only as a source of information about the rewards available in the environment.

In this paper, I extend this framework to allow for rewards that depend intrinsically upon the choices made by others, such that an option may become more or less rewarding based on how many other agents have also chosen it. Using this model, I show how social information in the absence of competition can reduce the collective accuracy of a group, and how introducing competition in the form of diminishing rewards for options already chosen by other agents can eliminate correlations between agents' choices, and return the group to a maximum level of collective accuracy.

Model

I consider a binary choice scenario with potential options labelled as A and B. In any given choice, one option is 'correct' and the other is 'incorrect'. This scenario is similar to that in (Mann, 2018), and the model described below largely follows the framework developed in that paper. Parameter definitions are summarised in Table 1, and R code to reproduce the analysis is included as a supplementary file.

Agents sequentially choose either A or B, and are able to observe all choices made before their own, such that these constitute common knowledge (Aumann, 1976). The choice

Table 1. Definitions of model parameters.

Parameter	Definition
x	Identity of the correct choice: $x = 1$ for A. $x = -1$ for B
Δ_i	Private signal received by agent i
ϵ	Scale of noise in private signals
S	Social information in the form of a sequence of previous choices
C_k	The decision made by the kth agent in a sequence (1 for A, -1 for B)
Δ^*	The threshold applied to private signals by an agent, such that the agent chooses A if $\Delta_i > \Delta^*$
n_A, n_B	The number of agents that have chosen A and B
q	The probability that a single agent will choose the correct side: $q = \Phi(1/\epsilon)$
Q	The ratio of correct to incorrect decisions by a single agent $Q = q/(1 - q)$
β	A competition factor

made by individual i can be labelled as $C_i = 1$ if A is chosen, or $C_i = -1$ if B is chosen, and a sequence of k decisions S is an ordered series $C_1, C_2, \ldots C_k$. The collective decision is defined as the majority choice when all n agents have decided, and for simplicity I consider only cases where n is odd so there are no tied collective decisions.

Agents choose between the two options based on reward criteria and their own inferences about the probability that each option is correct, so as to maximise their expected reward. The true state of the world can be given by a variable x, which takes the value $x = 1$ if option A is correct, and $x = -1$ if option B is correct. All agents are assumed to share a common, symmetric and uninformative prior about the value of x:

$$P(x = 1) = P(x = -1) = \frac{1}{2}. \tag{1}$$

Agents are informed by two sources of information. The first is a noisy private signal Δ_i received independently by each agent i, with variance ϵ:

$$p(\Delta_i | x) = \frac{1}{\epsilon} \phi\left(\frac{\Delta_i - x}{\epsilon}\right), \tag{2}$$

where $\phi(\cdot)$ is the standardised normal probability distribution function. The second source of information is the social information provided by the sequence of previous decisions S. Agents update their knowledge of x by performing Bayesian inference:

$$P(x | \Delta_i, S) = \frac{P(x)P(S | x)p(\Delta_i | x)}{\sum_{x \in \{-1,1\}} P(x)P(S | x)p(\Delta_i | x)}, \tag{3}$$

where the equation above makes use of the assumed independence of private signals, and thus the independence of S and Δ_i conditioned on x.

The structure of this model is assumed to be common knowledge amongst all agents. That is, agents are assumed to know that they all share a common, symmetric prior belief about x, and to be aware of the distribution of private signals and to know that this is the same for all agents.

Furthermore, agents know that all other agents have this same information. The rewards for making correct choices introduced in the next section are also assumed to be common knowledge in the same manner. Finally, agents are assumed to act rationally in the sense that they seek to maximise their expected reward and this rationality is once again common knowledge amongst all agents.

Rewards

Agents are motivated to make accurate choices by a retrospective reward policy that assigns rewards once the true correct choice is known. A simple and intuitive policy is to reward agents if they made the correct choice, thus motivating each individual to be as accurate as possible. This can be labelled as 'binary' rewards, in common with previous models on simultaneous decision-making (Hong et al., 2012; Mann and Helbing, 2017), since agents receive a reward of either zero or one (in some standardised reward units) for each choice. This reward policy can be defined mathematically via a reward function $r(C_i, S, x)$ that depends on the choice, C_i, made by individual i, the sequence of past decisions S and the true state of the world x, with binary rewards being defined as:

$$r_{\text{binary}}(C_i, S, x) = \delta_{C_i, x}, \tag{4}$$

where $\delta_{l,k}$ is the Kronecker delta function.

A binary reward function attributes rewards based solely on the accuracy of an individual's choice, and is independent of the decisions made by others. More generally we can consider a reward scheme that depends on past decisions that the agent can observe. A simple way to do this is to modulate the reward with a function that depends on the individual choice and S:

$$r_{\text{general}}(C_i, S, x) = f(C_i, S)\delta_{C_i, x}. \tag{5}$$

This continues to reward (and thus incentivise) accuracy through the $\delta_{C_i, x}$ term, but can also directly reward

or penalise choosing the same option as others, thus incentivising either conformity or diversity of choices. In the case of the first agent S is by definition the empty set $\varnothing$, and I define $f(C_i, \varnothing) = 1$, meaning that for the first agent the general reward scheme and the binary reward scheme are identical.

Rational individual choice

Given a reward function $r(C_i, S, x) = f(C_i, S)\delta_{C_i, x}$, an agent can evaluate the expected reward $\mathbb{E}(r_A | S, \Delta_i)$ from choosing A, conditioned on the available private and social information:

$$\begin{aligned}
\mathbb{E}(r_A | S, \Delta_i) &= \sum_{x \in \{-1,1\}} r(C_i = 1, S, x)P(x | \Delta_i, S) \\
&= f(C_i = 1, S)P(x = 1 | \Delta_i, S)
\end{aligned} \tag{6}$$

and similarly for choosing B:

$$\mathbb{E}(r_B | S, \Delta_i) = f(C_i = -1, S)P(x = -1 | \Delta_i, S), \tag{7}$$

According to the principle of expected reward maximisation, a rational agent will then select A if and only if $\mathbb{E}(r_A | S, \Delta_i) > \mathbb{E}(r_B | S, \Delta_i)$ (since Δ_i is real-valued, a tied expectation has zero probability mass). Using the general reward function above, this condition simplifies to:

$$\mathbb{E}(r_A | S, \Delta_i) > \mathbb{E}(r_B | S, \Delta_i)$$
$$\Rightarrow \quad \frac{P(x = 1 | \Delta_i, S)}{P(x = -1 | \Delta_i, S)} > \frac{f(C_i = -1, S)}{f(C_i = 1, S)}. \tag{8}$$

That is, for an agent to choose A, its assessment of the difference in probability for A to be correct rather than B must outweigh any penalty it receives for choosing A over B based on the past decisions.

A feature of the above decision-making procedure is that there exists some critical value of an agent's private information, Δ_i^*, which would make the expected reward of choosing A or B equal:

$$\mathbb{E}(r_A | S, \Delta_i^*) = \mathbb{E}(r_B | S, \Delta_i^*) \tag{9}$$

This implies that agent i will choose A if and only if $\Delta_i > \Delta_i^*$. Substituting the definition of the expected reward and the conditional probability $P(x | \Delta_i, S)$, this gives:

$$\frac{P(S | x = 1)\phi((\Delta_i^* - 1)/\epsilon)f(C_i = 1, S)}{P(S | x = -1)\phi((\Delta_i^* + 1)/\epsilon)f(C_i = -1, S)} = 1. \tag{10}$$

Simplifying the term $\phi((\Delta_i^* - 1)/\epsilon)/\phi((\Delta_i^* + 1)/\epsilon) = \exp(2\Delta_i^*/\epsilon^2)$, the expression above can be rearranged to give:

$$\Delta_i^* = \frac{\epsilon^2}{2}\left(\log\left(\frac{P(S | x = -1)}{P(S | x = 1)}\right) + \log\left(\frac{f(C_i = -1, S)}{f(C_i = 1, S)}\right)\right) \tag{11}$$

Since subsequent agents are able to observe the sequence S that agent i was responding to, they can calculate the corresponding value of Δ_i^*. Combined with observing the decision agent i makes, this enables them to infer whether agent i's private information was greater than or less than this threshold value. The probability of a sequence S, conditioned on x, can therefore be evaluated with reference to each of the thresholds calculated for the previous agents:

$$\begin{aligned}
P(S | x) &= \prod_{j \in \text{choosingA}} P\left(\Delta_j > \Delta_j^*\right) \times \prod_{j \in \text{choosingB}} P\left(\Delta_j < \Delta_j^*\right) \\
&= \prod_{j < i} \Phi\left(C_j\left(x - \Delta_j^*\right)/\epsilon\right)
\end{aligned} \tag{12}$$

where $\Phi(\cdot)$ is the cumulative probability function of the standard normal distribution. Since the thresholds depend themselves on the past sequence of decisions, the probability of a sequence can be evaluated recursively by calculating the threshold for each sub-sequence.

In the above description, I have assumed that rewards are deterministic functions of the choice the agent makes, the state of the world and the choices of other agents. However, from a mathematical perspective what determines the rational choice for a focal agent is the expectation of these rewards. Therefore a stochastic reward that preserves the same expectation will give the same rational strategy, but may provide less reliable feedback to the agents.

Results

Social response under binary rewards

The influence of social information on decision-making can be characterised qby observing its effect on both individual decisions and on the aggregate outcomes in groups. A simple way to visualise the influence of previous choices on an individual's decision is via the probability that a focal individual will choose the correct option, arbitrarily taken to be option A, conditioned on there having previously been n_A and n_B agents choosing A and B, respectively. This probability is shown in Figure 1(a), assuming that agents are responding to binary rewards ($f(C_i = 1, S) = f(C_i = -1, S) = 1$). Because the decision to choose A or B depends in theory on both the full sequence of previous choices and the agent's private information, the probability shown in this figure is a weighted average over all sequences consistent with
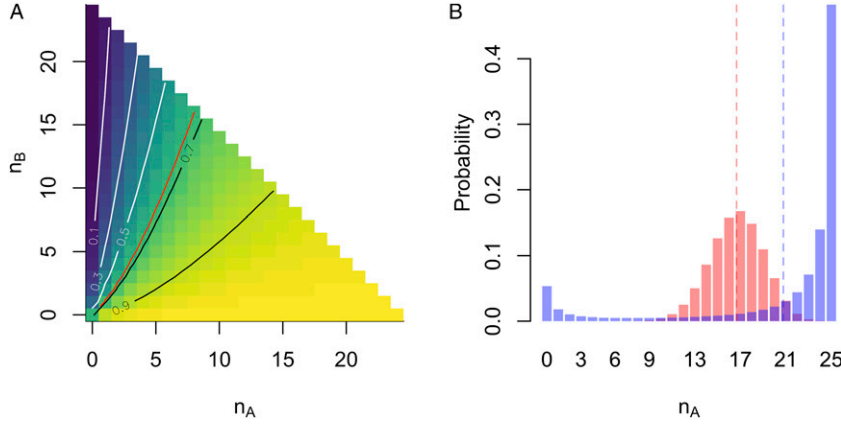
Figure 1. Characterising the response to social information in sequential decisions under binary rewards. (a) The probability that an agent will choose option A when that is the correct choice, conditioned on the number of previous decisions for options A and B, averaging over all sequences consistent with those aggregate number of decisions. In this example, the environmental noise is $\epsilon = 2.32$, giving an individual choice accuracy of $q = 2/3$; the red contour line indicates this probability. (b) The probability for n_A agents to select option A when that is the correct choice, averaged over a full sequence of decisions in a group of $n = 25$ agents. The blue bars indicate the probability when rational agents are subject to binary rewards, red bars indicate the probability if all agents select independently. The dashed lines indicate the mean of each probability distribution. Agents responding rationally to binary rewards have a higher average number of individually successful decisions, but a lower probability of a correct majority decision.

specified values of n_A and n_B, and all possible values of the focal agent's private information:

$$P(C_i = 1 \mid n_A, n_B) = \sum_{S \in n_A, n_B} P(\Delta_i > \Delta_S^*) P(S \mid x = 1), \quad (13)$$

where the summation is over the set of all sequences with n_A and n_B individuals choosing A and B.

This figure shows that agents respond strongly to the decisions made by others, such that the probability to choose A is highly dependent on the values of n_A and n_B. In particular, in most cases where $n_B > n_A$ the focal agent is less likely to choose the correct option than they would be if they chose independently; the red contour lines indicates this independent choice probability. This implies that incorrect decisions by agents at the beginning of the sequence can lead to a cascade of later agents also making incorrect choices, as has been demonstrated in previous models of sequential decision-making (Bikhchandani et al., 1992). In the model considered here, cascades are neither inevitable nor irreversible, because agents always have a finite probability to choose the less popular option. Nonetheless, the strong probability to follow what others have chosen is reflected in distribution of aggregate outcomes at the group level, characterised by the probability that n_A agents will select option A in total. This is plotted in Figure 1(b) for both the case of independent decisions (red bars) and for agents using social information with binary rewards (blue bars). This plot shows the dramatic difference in aggregate outcomes that results from social information use. When agents choose independently the aggregate outcomes are

clustered in a binomial distribution that peaks at the mean value of $n_A = n\Phi(1/\epsilon)$, with a very low probability that fewer than half the agents choose A. Under social information the aggregate outcomes become bimodal, with a large peak at $n_A = n$ and a secondary peak at $n_A = 0$. The result of this is that the mean number of correct decisions increases (compare the blue and red dashed lines), but there is a much greater probability that a majority of agents will choose the incorrect option (B). As such, each individual is more likely to choose the correct option, but the majority choice of the group is less likely to be correct. Under social influence, the typical choice majorities also become much larger, whether in favour of the correct or incorrect option, whereas without social influence unanimity is unlikely.

Condorcet-retrieving reward function

Under binary rewards, agents tend to follow past decisions with increasing strength over the course of a sequence of choices (Figure 1(a)). Since this breaks the assumption of independence in the CJT, it also reduces the accuracy of collective decisions as defined by the majority choice, as shown in Figure 1(b). To improve collective accuracy, it is therefore necessary to reduce the correlation between decisions. If agents make choices independently, this implies that the threshold value of Δ_S^* should be independent of the value of S. Since agents begin with a symmetric prior $P(x = 1) = 1/2$, it further implies that this threshold must be zero – that is, agents will choose A or B based solely on the direction of their private information. One can therefore

retrieve independent choices, and thus the accuracy implied by the CJT, by seeking a reward function $r_{\text{condorcet}}(C_i, S, x)$ such that:

$$\mathbb{E}(r_A|\, S, \Delta_i = 0) = \mathbb{E}(r_B|\, S, \Delta_i = 0) \forall S. \qquad (14)$$

Expanding the definition for the expected reward, this gives:

$$\frac{P(x = 1|\, \Delta_i = 0, S) f(C_i = 1, S)}{P(x = -1|\, \Delta_i = 0, S) f(C_i = -1, S)} = 1. \qquad (15)$$

By substituting Bayes rule for the conditional probability of x, we get:

$$\frac{p(\Delta_i = 0|\, x = 1) P(S|\, x = 1)}{p(\Delta_i = 0|\, x = -1) P(S|\, x = -1)} = \frac{f(C_i = -1, S)}{f(C_i = 1, S)} \qquad (16)$$

By construction, under this Condorcet reward scheme, all thresholds for private information are zero. The probability $P(S|x)$ thus simplifies to a product of independent choices:

$$P(S|\, x) = \Phi\left(\frac{x}{\epsilon}\right)^{n_A} \Phi\left(\frac{-x}{\epsilon}\right)^{n_B}, \qquad (17)$$

where n_A is the number of agents who have previously chosen A and n_B the number who have chosen B. Substituting this expression and recognising that $p(\Delta_i = 0|x = 1) = p(\Delta_i = 0|x = -1)$, we therefore get:

$$\frac{\Phi(1/\epsilon)^{n_A} \Phi(-1/\epsilon)^{n_B}}{\Phi(-1/\epsilon)^{n_A} \Phi(1/\epsilon)^{n_B}} = \frac{f(C_i = -1, S)}{f(C_i = 1, S)} \qquad (18)$$

This expression can be simplified by defining $q = \Phi(1/\epsilon)$ as the probability that a single agent will independently choose the correct option. This then reduces to:

$$\begin{aligned}\frac{f(C_i = -1, S)}{f(C_i = 1, S)} &= \frac{q^{n_A}(1 - q)^{n_B}}{(1 - q)^{n_A} q^{n_B}} \\ &= Q^{n_A - n_B},\end{aligned} \qquad (19)$$

where $Q = q/(1 - q)$ is the odds ratio for a single agent to choose correctly. This expression can be satisfied by a reward scheme:

$$r_{\text{condorcet}}(C_i, S, x) = \begin{cases} \delta_{C_i, x} Q^{-n_A} & \text{if } C_i = 1 \\ \delta_{C_i, x} Q^{-n_B} & \text{if } C_i = -1 \end{cases} \qquad (20)$$

This expression shows that rational agents can be motivated to make independent choices if the rewards for each option are reduced geometrically with the number of agents that have already chosen that option. This is a very convenient reward system for several reasons: First, it is symmetric in the way it treats both options, so neither option needs to be arbitrarily favoured or penalised. This is important since we assume that purpose of observing the

collective decision is to determine the correct choice, and thus the reward-setter does not know this in advance. Second, the form of the required penalty for each option depends only on the number of agents that have previous chosen it, so these penalties can be implemented locally without reference to the number choosing the other option, or the order in which those choice were made. Third, it resembles a form of competition, with each agent exhausting a fixed proportion of the potential reward remaining for the option it chooses. The geometric reduction in rewards means that for any group size the total rewards available from each option are bounded by:

$$\text{Maximum reward} = 1 + \frac{1}{Q} + \frac{1}{Q^2} \ldots = \frac{Q}{(Q - 1)}. \qquad (21)$$

Similarly, the expected total reward can be calculated as:

$$\begin{aligned}\mathbb{E}(\text{total reward}) &= \mathbb{E}\left(1 + \frac{1}{Q} + \frac{1}{Q^2} \ldots \frac{1}{Q^k}\right), k\widetilde{\text{Bin}}(n, q) \\ &= \frac{Q - 2^n(1 - q)^n}{Q - 1}\end{aligned}$$
$$(22)$$

Any system that assigns rewards under this scheme can therefore estimate and bound the total rewards it would potentially need to allocate. It is notable that high values of Q indicate problems that are relatively simple for individual decision makers, and these represent the lowest expectation and bound on total rewards; this naturally allows a reward system to allocate the greatest reward budget to the most difficult problems.

Robustness of collective accuracy under varying competition

The reward scheme derived above is constructed so as to maximise the accuracy of the majority choice by making individual rational decisions statistically independent, and it accomplishes this through imposing a specific form of competitive penalty. As discussed above, this form of competitive penalty has many agreeable features for implementation in real-world decisions problems. However, selecting the precise strength of the competitive penalty requires knowing in advance how difficult the decision problem is, that is, knowing the value of Q. In general, it is unlikely that this would be precisely known in advance, although a system designer may have some intuition about whether a given decision is easy or difficult. As such, it is important to assess how robust such reward system is to misspecification of the competition strength. To do this, we can evaluate the collective accuracy under a reward scheme with variable competition strength β:

$$r_{\text{competitive}}(C_i, S, x) = \begin{cases} \delta_{C_i,x}\beta^{-n_A} & \text{if } C_i = 1 \\ \delta_{C_i,x}\beta^{-n_B} & \text{if } C_i = -1, \end{cases} \quad (23)$$

where we know from the above argument that the optimum value of β should be Q. Under this reward scheme, the relation for critical thresholds given by equation (11) can be simplified and evaluated efficiently as:

$$\Delta_i^* = \frac{\epsilon^2}{2}\sum_{j<i}\left[\log\Phi\left(\frac{C_j\left(1-\Delta_j^*\right)}{\epsilon}\right)\right.$$
$$\left. -\log\Phi\left(\frac{C_j\left(-1-\Delta_j^*\right)}{\epsilon}\right)\right] \quad (24)$$
$$+\frac{\epsilon^2}{2}(n_A - n_B)\log\beta,$$

where n_A and n_B are the number of decisions for A and B, respectively, within the sequence S, and β can be either greater than one (competition) or less than one (rewarding conformity).

The expected collective accuracy under this reward scheme can be evaluated by directly calculating the expected proportion of accurate majority decisions as a function of the adjustable competition parameter β. This is done by calculating the probability of every possible sequence of decisions in a group of n agents (hence 2^n possible sequences) for $x = 1$ and summing the probability of that set of sequences where the majority of decisions are for the correct option A:

$$\mathbb{E}(\text{collective accuracy}) = \sum_{S \in n_A > n_B} P(S\,|\, x = 1), \quad (25)$$

where $P(S|x = 1)$ is given by evaluating equation (12).

Figure 2(a) shows the collective accuracy as a function of β for group sizes from $n = 3$ to $n = 25$ with an environmental noise level set of $\epsilon = 2.32$, implying $q = 2/3$ and $Q = 2$ (i.e. individuals will make the correct choice twice as often as the wrong choice when choosing alone). This demonstrates a clear peak in accuracy in each case at the expected value of $\beta = 2$, indicated by the dashed red line. At this optimum point collective accuracy matches that expected from the CJT. While a range of values of $\beta > 1$ induce greater collective accuracy than under binary rewards ($\beta = 1$), value of $\beta < 1$, which reward agents for copying past decisions, dramatically reduce collective accuracy. Collective accuracy is more robust to values of β that are greater than Q (the optimum) than those that are lower, especially in larger groups; Figure 2(b) shows the individual accuracy for the same range of group sizes and competition strengths, demonstrating that increases in collective accuracy induced by competition lead to decreases in individual accuracy – collective accuracy is maximised when individual accuracy

falls to that expected from a single agent without social information, as this is when agents choose independently. Average individual accuracy is maximised at values of β slightly greater than one, since under binary rewards each agent is motivated to maximise its own accuracy, without regard for the value of the social information it provides to those further along the sequence of decision makers (cf. (Torney et al., 2015)).

Figure 2(c) shows the relationship between the collective accuracy achieved by the Condorcet reward scheme and that achieved without competition (binary rewards), showing that the effect is stronger in larger groups, which suffer relatively more from information cascades under binary rewards. Although collective accuracy is maximised when competition is optimised to produce independent decision-making, there is a range of values of β which induce greater collective accuracy than under binary rewards, as seen in Figure 2(a). The size of this range shows how well-tuned competition must be to generate improvements in collective accuracy, and thus is indicative of how plausible effectively implementing such a reward scheme might be in practice. Figure 2(d) shows the maximum value of β that outperforms binary rewards as a function of q, for group sizes from $n = 3$ to $n = 25$ (solid lines), as well as the optimal value of β for comparison (dashed line). Inherently, easier decisions permit a greater range of effective competition strengths, and this range increases very rapidly as q approaches one (note the log scale on the y-axis). Larger groups also permit a wider range of effective competition strengths, even though the optimal competition strength does not depend on n.

Discussion

When agents are rewarded solely for their individual accuracy they tend to follow previous decisions. While this increases the expected proportion of agents that make the correct choice, it reduces the probability that the majority of agents is correct compared to agents who make their decisions independently. Errors in early decisions can make subsequent decision-makers less accurate than they would have been alone. Hence, while on average individually beneficial, social information is deleterious to anyone seeking to use the wisdom of the crowd by relying on the majority opinion.

Social information may potentially be restricted exogenously, by insisting that individuals make their choices without access to the choices made by others. Such a scenario requires tight control of the information individuals have access to, and is unlikely to be plausible when making use of collective wisdom in real-world contexts such as online review systems and social media (Klein, 2011). Alternatively, agents could be motivated to make independent choices by rewarding them
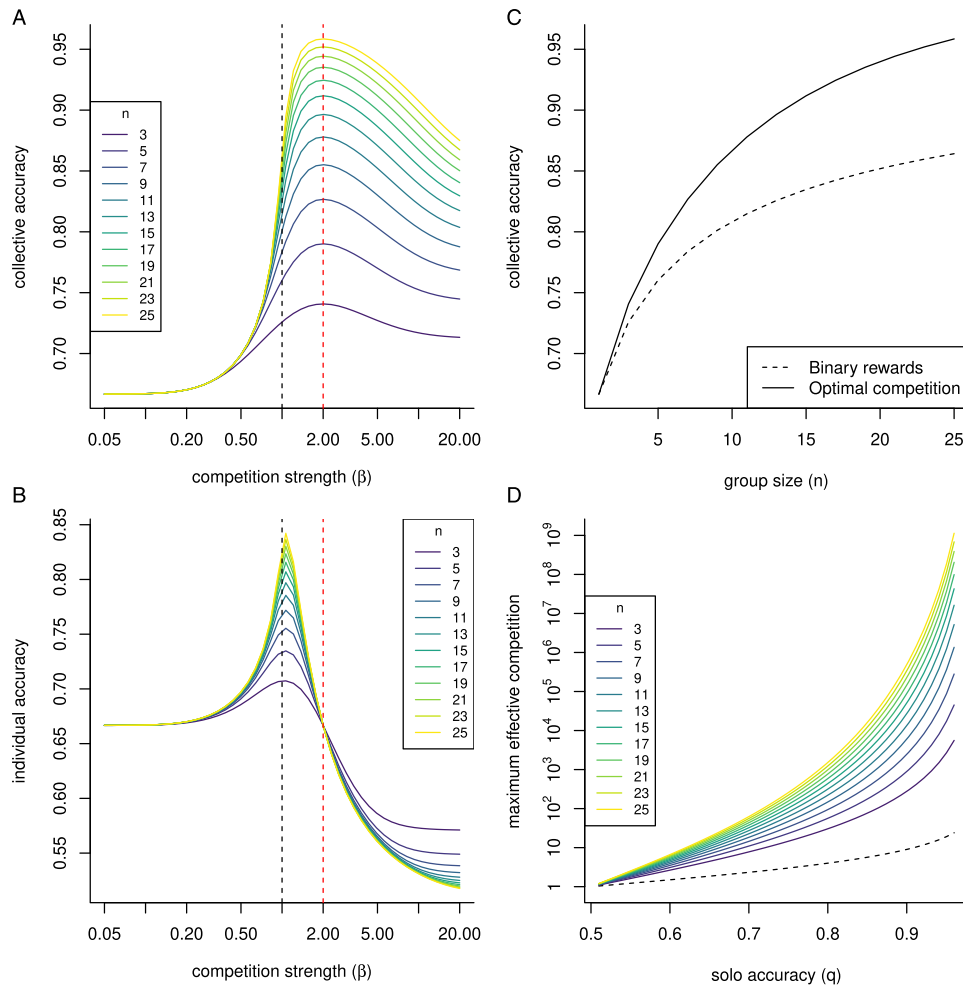
Figure 2. The effect of competition on collective accuracy. (a) With q = 2/3, across different group sizes (n) collective accuracy increases with increasing competition (β) up to an optimal value given by $\beta = Q$ (indicated by the dashed red line), where collective accuracy matches that predicted by the Condorcet Jury Theorem. Higher levels of competition reduce collective accuracy, with sufficiently high values of β leading to lower collective accuracy than under binary rewards (β = 1, indicated by the dashed black line). Values of β less than one, indicating rewards for conformity, always lead to lower collective accuracy. (b) Individual accuracy is maximised at values of β close to one, indicating weak positive competition, and increases with group size. At the optimal competition for collective accuracy, individual accuracy is the same for all group sizes as agents choose independently. (c) The collective accuracy under optimal competition (solid line) compared to that achieved under binary rewards (dashed line) as a function of group size. (d) The maximum value of β for which competitive rewards outperform binary rewards, for varying group size and as a function of q (representing the probability for a solo agent to choose correctly). The dashed line shows the optimal value of $\beta = Q$ for comparison. The range of effective competition values (those that improve on binary rewards) is greater for easier decisions and in larger group sizes. Note the logarithmic scale on the y-axis.

when the collective vote is correct, rather than based on their individual accuracy; when no choice is favoured a priori this would make sincere independent voting the Nash equilibrium strategy (Austen-Smith and Banks, 1996). However, collective rewards have several disadvantages. In any but the smallest groups, agents are highly unlikely to represent the pivotal vote, and thus, the reward they receive will most often not depend on the choice they make. This means that agents will have to experience many collective choices to experience any

feedback that can direct them to the optimal strategy, reducing both learning by reinforcement and the psychological motivation to invest effort in making informed choices. Moreover, if rewards are allocated based on collective performance, agents benefit from being in larger groups, which magnify this problem. Where groups are relatively small and composed of a fixed membership over time, collective rewards could be effective. In other scenarios, where group membership is neither fixed in size or composition over sustained periods of time, it is

desirable to provide individual rewards that dependably vary with the choices those individuals make.

Here, I have demonstrated that, among rational and selfish agents, a simple competitive reward scheme that reduces the rewards available from already-popular choices can, in theory, return a group to the accuracy implied by the Cordorcet Jury Theorem. This result depends on the assumption that the environmental information received by the agents is truly independent and is not systematically wrong, but effectively balances the expected gains of following social information by choosing the more popular option, and so prevents the information cascades that limit the collective accuracy of sequential decision-making. Under such a reward scheme, and within the assumptions of the model used here, agent's decisions become independent, and depend only on their private information. This increases the probability that the majority of the group will make the correct choice, albeit at the cost of making each individual somewhat less accurate on average. This paper has derived the optimal form and magnitude of this competition in the context of a model in which an agent observes the full sequence of previous decisions, but because agents' decisions become independent under the optimal competition it would retain the same form if agents instead observed simplified aggregate statistics regarding how many agents had made each choice (Mann, 2021). As such, it is applicable across a wide range of domains where the nature of social information may vary.

Introducing competition that penalises agents for following popular choices is an established mechanism for motivating agents to make decisions that improve collective accuracy by reducing the correlation between different decision-makers (Hong et al., 2012; Mann and Helbing, 2017), and is an important feature of markets as a forecasting mechanism, whether explicitly prediction markets (Wolfers and Zitzewitz, 2004), betting exchanges or financial markets. In this paper, I have shown that competition can also fulfil this role in a sequential decision-making context where agents can observe the choices made by all those who decide before them and utilise that information in their own decision-making. While the optimal level of competition is unlikely to be known a priori for any given decision or decision-making system, sensitivity analysis shows that introducing a small degree of competition typically improves upon performance from binary rewards alone; in an adaptive system competitive pressure can thus be gradually raised to determine optimal performance. Except in very difficult decisions ($q \simeq 0.5$), competitive rewards are relatively forgiving to miscalibration, providing improvements on binary rewards across a wide scale of competition strengths.

In this paper, I have assumed that individuals make choices strictly according to a predefined sequential ordering, and in a context where the group size is known by all, and each agent knows where in the order it is. However, under the reward scheme, I have proposed an agent need not consider, or even be aware of, agents later in the sequence, since these choices cannot provide any information or affect its own reward. This is beneficial in contexts where groups form organically rather than being fixed. It is notable that the greatest possible rewards are available to agents who decide early in the sequence. Although I have considered the decision sequence to be fixed, relaxing this assumption might incentivise those with the strongest private information to choose first. As the maximum potential reward diminishes later in the sequence, this could also provide a motivation for agents to turn their attention to other problems once enough decisions have been made to make the eventual collective vote effectively known. A further possibility is for agents to abstain from voting when their private information is weak, which could be encouraged by introducing a small cost for participation. An interesting line of future research would be to model such a dynamic environment of agents choosing which problems to contribute to while seeking the greatest individual net returns.

The theoretical efficacy of competitive rewards raises the possibility that such incentives could be used to improve collective accuracy across a range of real-world contexts. For example, the collective judgement of the scientific community (as reflected in majority expressed opinion) on issues where there is significant uncertainty could potentially be improved by systematically assigning greater rewards to those later proved correct when fewer others also expressed that opinion; these rewards might be in the form of promotions, research funding or simply scientific reputation. To some degree, such competitive rewards already feature in many communities, and many scientists, economists and political pundits have made their reputation by advocating for a minority viewpoint that was later proved correct: a notable example is the case of Barry Marshall and Robin Warren, who won the 2005 Nobel Prize in Physiology or Medicine for their discovery of the link between *H. pylori* and stomach ulcers (Marshall et al., 1985). However, other pressures that incentivise social and professional conformity are also common, such as what Irving Janis termed 'Groupthink' (Janis, 1983) – the tendency to excessively value consensus with other group members. Conformity may also be imposed by systemic factors such as needing to convince others that your ideas are plausible before they can be explored (Gross and Bergstrom, 2021).

From an external point of view, the results of this study suggest we should assign greater credibility to the collective wisdom of communities where such competitive rewards are the norm, motivating both accuracy and independence. Conversely, the collective wisdom of communities characterised by strong social norms of conformity (effectively negative competition) should be assigned lower credibility. Notably, although competitive, some communities such as

political punditry rarely demand or reward specific, falsifiable predictions (Tetlock and Gardner, 2016); for competitive rewards to drive collective accuracy there must be penalties (or lack of reward) for inaccurate predictions, otherwise individuals are simply motivated to identify and state a unique opinion without regard for its accuracy. The optimal reward structure identified here requires that rewards are still contingent on accuracy.

Where the collective accuracy of group decisions (at least as expressed via majority voting) is highly desirable, we should seek to reduce pressures that induce conformity and introduce competitive rewards that motivate more independent judgements. However, as well as potentially causing social friction (if norms of social conformity are violated), this also comes at the cost of a likely reduction in individual accuracy. While the group may be more accurate, more individuals will be wrong. This highlights that systems of collective decision-making that aim for collective accuracy must not only seek to be tolerant of conflicting views, but must also tolerate a greater level of individual decision-making failure.

To be maximally effective such competitive rewards need to be predictable. Agents should be able to either rationally adjust their choices in the light of known reward schemes, or such rewards should be consistent enough to allow adaptation by reinforcement learning – a process which may take time to affect behavioural change (Burton-Chellew et al., 2015; Burton-Chellew and West, 2021; Burton-Chellew and Guérin, 2021). Rewards should also be well-tuned to the specific context (particularly the level of individual certainty, but also the size of the community). This suggests that competitive reward structures should be made more explicit, and calibrated through systematic trial and error in a particular community. Notably, the results presented here suggest that excessive competition is likely to be more effective than too little. The model presented here is theoretical and assumes that agents are either well informed about potential rewards and respond rationally, or reliably adapt via reinforcement based on experience. Such assumptions, and the efficacy of innovative collective decision-making systems, are ultimately empirical questions that must be tested experimentally.

Finally, this paper has considered problems in which agents seek to ascertain the answer to a question of external empirical fact, such as whether it will rain tomorrow, or which of two teams will win a sporting event. It should be noted that some attempts to leverage the predictive power of social information instead focus on questions where the answer is endogenous to the community from which that social information is drawn. An example is the use of social media to predict which movies will attract large box office returns (Asur and Huberman, 2010), since presumably the commenters on social media represent a sample of potential movie-goers. In these cases, there is likely to be more value in simply measuring the aggregate opinion of individuals, since the expression of interest in a movie is itself a predictor of attendance, regardless of whether that interest is itself socially driven.

Declaration of conflicting interests

Funding

ORCID iD

Richard P Mann   https://orcid.org/0000-0003-0701-1274

Supplemental Material

Supplemental material for this article is available online.

References

Anderson LR and Holt CA (1997) *Information Cascades in the Laboratory*. The American Economic Review, 87(5). 847–862.

Arganda S, Pérez-Escudero A and De Polavieja GG (2012) A common rule for decision-making in animal collectives across species. *Proceedings of the National Academy of Sciences of the United States of America* 109: 20508–20513.

Asur S and Huberman BA (2010) Predicting the future with social media. In: 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Toronto, ON, Canada, 31 August 2010–03 September 2010, pp. 492–499, 1. IEEE.

Aumann RJ (1976) Agreeing to disagree. *The Annals of Statistics* 4(6): 1236–1239.

Austen-Smith D and Banks J S (1996) Information aggregation, rationality, and the condorcet jury theorem. *American Political Science Review* 90(1): 34–45.

Bikhchandani S, Hirshleifer D and Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* 100(5): 992–1026.

Boland PJ (1989) Majority systems and the condorcet jury theorem. *The Statistician* 38(3): 181–189.

Burton-Chellew MN and Guérin C (2022) Self-interested learning is more important than fair-minded conditional cooperation in public-goods games. *Evolutionary Human Sciences*, 4, E46. DOI: 10.1017/ehs.2022.45.

Burton-Chellew MN and West SA (2021) Payoff-based learning best explains the rate of decline in cooperation across 237 public-goods games. *Nature Human Behaviour* 5: 1330–1338.

Burton-Chellew MN, Nax HH and West SA (2015) Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society of London B: Biological Sciences* 282(1801): 20142678.

De Condorcet N (1785) *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix.* Paris: Imprimerie royale.

Faria J, Krause S and Krause J (2010) Collective behavior in road crossing pedestrians: the role of social information. *Behavioral Ecology* 21(6): 1236–1242.

Gallup AC, Hale JJ, Sumpter DJT, et al. (2012) Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences of the United States of America* 109(19): 7245–7250.

Galton F (1907) Vox populi. *Nature* 75(7): 450–451.

Gross K and Bergstrom CT (2021) Why ex post peer review encourages high-risk research while ex ante review discourages it. *Proceedings of the National Academy of Sciences* 118(51): e2111615118.

Hong L, Page SE and Riolo M (2012) Incentives, information, and emergent collective accuracy. *Managerial and Decision Economics* 33(5–6): 323–334.

Janis I L (1983) *Groupthink.* Boston: Houghton Mifflin.

King AJ and Cowlishaw G (2007) When to use social information: the advantage of large group size in individual decision making. *Biology Letters* 3(2): 137–139.

Klein M (2011) *How to Harvest Collective Wisdom on Complex Problems: An Introduction to the MIT Deliberatorium.* Cambridge: MIT Center for Collective Intelligence working paper.

Landemore H (2012) *Democratic Reason.* Princeton: Princeton University Press.

List C and Goodin RE (2001) Epistemic democracy: generalizing the condorcet jury theorem. *Journal of Political Philosophy* 9(3): 277–306.

List C (2004) Democracy in animal groups: a political science perspective. *Trends in Ecology and Evolution* 19(4): 168–169.

Lorenz J, Rauhut H, Schweitzer F, et al. (2011) How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences of the United States of America* 108(22): 9020–9025.

Mackay C (1841) *Extraordinary Popular Delusions and the Madness of Crowds.* New York: Farrar, Straus and Giroux.

Mann RP and Helbing D (2017) Optimal incentives for collective intelligence. *Proceedings of the National Academy of Sciences of the United States of America* 114(20): 5077–5082.

Mann RP, Faria J, Sumpter DJT, et al. (2013) The dynamics of audience applause. *Journal of the Royal Society, Interface* 10: 20130466.

Mann RP (2018) Collective decision making by rational individuals. *Proceedings of the National Academy of Sciences of the United States of America* 115(44): E10387–E10396.

Mann RP (2020) Collective decision-making by rational agents with differing preferences. *Proceedings of the National Academy of Sciences of the United States of America* 117(19): 10388–10396.

Mann RP (2021) Optimal use of simplified social information in sequential decision-making. *Journal of the Royal Society Interface* 18(179): 20210082.

Marshall BJ, Armstrong JA, McGechie DB, et al. (1985) Attempt to fulfil Koch's postulates for pyloric Campylobacter. *The Medical Journal of Australia* 142(8): 436–439.

Pérez-Escudero A and De Polavieja G G (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLoS Computational Biology* 7(11): e1002282.

Sumpter DJT and Pratt SC (2009) Quorum responses and consensus decision making. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364(1518): 743–753.

Surowiecki J (2005) *The Wisdom of Crowds.* London: Abacus.

Tetlock PE and Gardner D (2016) *Superforecasting: The Art and Science of Prediction.* London: Random House.

Torney CJ, Lorenzi T, Couzin ID, et al. (2015) Social information use and the evolution of unresponsiveness in collective systems. *Journal of the Royal Society Interface* 12(103): 20140893.

Tump AN, Pleskac TJ and Kurvers RH (2020) Wise or mad crowds? The cognitive mechanisms underlying information cascades. *Science Advances* 6(29): eabb0266.

Wolfers J and Zitzewitz E (2004) Prediction markets. *Journal of Economic Perspectives* 18: 107–126.