Debiasing the crowd: how to select social information to improve judgment accuracy?

Short title: How to select social information to improve judgment accuracy?

Bertrand Jayles[1,*], Clément Sire[2], Ralf H.J.M Kurvers[1]

[1]Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany
[2]Laboratoire de Physique Théorique, Centre National de la Recherche Scientifique (CNRS), Université de Toulouse – Paul Sabatier (UPS), Toulouse, France

Abstract

Cognitive biases are widespread in humans and animals alike, and can sometimes be reinforced by social interactions. One prime bias in judgment and decision making is the human tendency to underestimate large quantities. Former research on social influence in estimation tasks has generally focused on the impact of single estimates on individual and collective accuracy, showing that randomly sharing estimates does not reduce the underestimation bias. Here, we test a method of social information sharing that exploits the known relationship between the true value and the level of underestimation, and study if it can counteract the underestimation bias. We performed estimation experiments in which participants had to estimate a series of quantities twice, before and after receiving estimates from one or several group members. Our purpose was threefold: to study (i) whether restructuring the sharing of social information can reduce the underestimation bias, (ii) how the number of estimates received affects sensitivity to social influence and estimation accuracy, and (iii) the mechanisms underlying the integration of multiple estimates. Our restructuring of social interactions successfully countered the underestimation bias. Moreover, we find that sharing more than one estimate also reduces the underestimation bias. Underlying our results are a human tendency to herd, to trust larger estimates than one's own more than smaller estimates, and to follow disparate social information less. Using a computational modeling approach, we demonstrate that these effects are indeed key to explain the experimental results. Overall, our results show that existing knowledge on biases can be used to dampen their negative effects and boost judgment accuracy, paving the way for combating other cognitive biases threatening collective systems.

*Corresponding author – jayles@mpib-berlin.mpg.de

Authors summary

Humans and animals are subject to a variety of cognitive biases that hamper the quality of their judgments. We study the possibility to attenuate such biases, by strategically selecting the pieces of social information to share in human groups. We focus on the underestimation bias, a tendency to underestimate large quantities. In estimation experiments, participants were asked to estimate quantities before and after receiving estimates from other group members. We varied the number of shared estimates, and their selection method. Our results show that it is indeed possible to counter the underestimation bias, by exposing participants to estimates that tend to overestimate the group median. Subjects followed the social information more when (i) it was further away from their own estimate, (ii) the pieces of social information showed a high agreement, and (iii) it was on average higher than their own estimate. We introduce a model highlighting the core role of these effects in explaining the observed patterns of social influence and estimation accuracy. The model is in good agreement with the data. The success of our method paves the way for testing similar interventions in different social systems to impede other cognitive biases.

Introduction

Human and non-human animal judgments and decisions are characterized by a plethora of cognitive biases, i.e., deviations from assumed rationality in judgment [1]. Biases at the individual level can have negative consequences at the collective level. For instance, Mahmoodi et al. showed that the human tendency to give equal weight to the opinions of individuals (equality bias) leads to sub-optimal collective decisions when groups harbor individuals with different competences [2]. Understanding the role of cognitive biases in collective systems is becoming increasingly important in modern digital societies.

The recent advent and soar of information technology has substantially altered human interactions, in particular how social information is shared and processed: people share content and opinions with thousands of contacts on social networks such as Facebook and Twitter [3, 4, 5], and rate and comment on sellers and products on websites like Amazon, TripAdvisor, and AirBnB [6, 7, 8]. While this new age of social information exchange carries vast potential for enhanced collaborative work [9] and collective intelligence [10, 11, 12, 13], it also bears the risks of amplifying existing biases. For instance, the tendency to favor interactions with like-minded people (in-group bias [14]) is reinforced by recommender systems, enhancing the emergence of echo chambers [15] and filter bubbles [16] which, in turn, further increases the risk of opinion polarization. Given the importance of the role of biases in social systems, it is important to develop strategies that can reduce their detrimental impact on judgments and decisions in social information sharing contexts.

One promising, yet hitherto untested, strategy to reduce the detrimental impact of biases is to use prior knowledge on these biases when designing the structure of social interactions. Here, we will test whether such a strategy can be employed to reduce the negative effects of a specific bias on individual and collective judgments in human groups. We use the framework of estimation tasks, which are well-suited to quantitative studies on social interactions [17, 18, 19, 20], and focus on the *underestimation bias*. The underestimation bias is a well-documented human tendency to underestimate large quantities in estimation tasks [20, 21, 22, 23, 24, 25, 26, 27, 28, 29]. The underestimation bias has been reported across various tasks, including in estimations of numerosity, population sizes of cities, pricing, astronomical or geological events, and risk judgment [20, 26, 27, 28, 29]. Previous research—using a dot estimation task—showed that this effect already starts when the actual number of dots exceeds 10 [22]. This study (and others) suggest that the tendency to underestimate large quantities could stem from an internal compression of perceived stimuli [22, 23, 24]. The seminal study by Lorenz et al. (2011) has shown that the effects of the underestimation bias could be amplified after social interactions in human groups, and deteriorate judgment accuracy [19].

We here investigate the effects of different interaction structures, aimed at counteracting the underestimation bias, on individual and collective accuracy (details are given below). Moreover, we investigate how these structures interact with the number of estimates shared in shaping accuracy. Previous research on estimation tasks has largely overlooked both of these factors. Thus far, research on estimation tasks mostly discussed the beneficial or detrimental effects of social influence on group performance [19, 30, 31, 32, 33, 34, 35, 36]. Moreover, most previous studies focused on the impact of a *single* piece of information (one estimate or the average of several estimates), or did not systematically vary their number. In addition, in most studies, subjects received social information from *randomly selected* individuals (either group members, or participants from former experiments) [17, 18, 19, 20, 31, 35, 36, 37, 38, 39]. In contrast to these previous works, in many daily choices under social influence, one generally considers not only one, but several sources of social information, and these sources are rarely chosen randomly [40]. Even when not actively selecting information sources, one routinely experiences recommended content (e.g., books on Amazon, movies on Netflix, or videos on YouTube) generated by algorithms which incorporate our "tastes" (i.e., previous choices) and that of (similar) others [41].

Following these observations, we confronted groups with a series of estimation tasks, in which individuals first estimated miscellaneous (large) quantities, and then re-evaluated their estimates after receiving a varying number of estimates τ ($\tau = 1$, 3, 5, 7, 9, and 11) from other group members. Crucially, the shared estimates were selected in three different manners:

- *Random* treatment: subjects received personal estimates from τ random other group members. Previous research showed that when individuals in groups receive single, randomly selected estimates, individual accuracy improves because estimates converge, but collective accuracy does not [19, 20]. Since several random estimates do not, on average, carry higher information quality than a single random estimate, we did not expect collective accuracy to improve when individuals received multiple random estimates. However, we predicted that increasing the number of estimates shared would lead to a higher imitation rate and thus to increased improvements in individual accuracy.

- *Median* treatment: subjects received as social information the τ estimates from other subjects whose logarithm[1] are closest to the *median* log estimate m of the group (excluding their own). This selection method thus selects central values of the distribution and removes extreme values. Since median estimates in estimation tasks are generally closer to the true value than randomly selected estimates (Wisdom of Crowds) [42, 43, 44], we expected higher

[1] Logarithms are more suitable because humans perceive numbers logarithmically (order of magnitudes) [45].

improvements in collective and individual accuracy than in the Random treatment.

- *Shifted-Median* treatment: as detailed above, humans have a tendency to underestimate large quantities in estimation tasks. Recent works have suggested aggregation measures taking this bias into account, or the possibility to counteract it using artificially generated social information [20, 26]. Building on this, we here test a method that exploits prior knowledge on this underestimation bias, by selecting estimates that are likely to reduce its effects. We define, for each group and each question, a shifted (overestimated) value m' of the median log estimate m that approximates the log of the true value T (thus compensating the underestimation bias), exploiting a relationship between m and $\log(T)$ identified from prior studies using similar tasks (for details see Experimental Design). Individuals received the estimates of which logarithms were closest to $m' > m$ (except their own). This selection method also tends to eliminate extreme values, but additionally favors estimates that are slightly above the center of the distribution. Given the overall underestimation bias, values slightly above the center of the distribution are, on average, closer to the true value than values at the center of the distribution. Therefore, we expected the highest improvements in collective and individual accuracy in this treatment. Note that our method uses prior domain knowledge (to estimate the true value of a quantity) but does not require a priori knowledge of the true value of the quantity at hand. That is, the accuracy of the selected estimates is a priori unknown, and they are only statistically expected to be closer to the truth.

We first describe the distributions of estimates and sensitivities to social influence in all conditions. Doing so, we shed light on the key effects influencing subjects' response to social information, which are: (i) the dispersion of the social information, (ii) the distance between the personal estimate and the social information, and (iii) whether the social information is higher or lower than the personal estimate. We then build a model of social information integration incorporating these findings, and use it to further analyze the impact of the number of shared estimates on social influenceability and estimation accuracy. We find, in accordance with our prediction, that improvements in collective accuracy are indeed highest in the Shifted-Median treatment, demonstrating the success of our method in counteracting the underestimation bias.

Experimental design

Participants were 216 students, distributed over 18 groups of 12 individuals. Each individual was confronted with 36 estimation questions displayed on a tactile tablet. Questions were a mix of general knowledge and numerosity, and involved moderately large to very large quantities (the list of questions and answers is provided in the Supplementary Information). Each question was asked twice: first, subjects were asked to provide their personal estimate E_{p}. Next, they received as social information the estimate(s) of one or several group member(s), and were asked to provide a second estimate E_{s} (see illustration in Supplementary Figure S1). When providing the social information, we varied (i) the number of estimates shown ($\tau = 1$, 3, 5, 7, 9, or 11) and (ii) how they were selected (Random, Median, or Shifted-Median treatments). The subjects were not aware of the three different treatments and were simply told that they would receive τ estimates from the other participants. Each group of 12 individuals experienced each of the 18 unique conditions (i.e., combination of number of estimates shared and their selection method) twice. Across all 18 groups, each of the 36 unique questions was asked once at every unique condition, resulting in $12 \times 36 = 432$ estimates per condition (both before and after social information sharing). Students received course credits for participation and were, additionally, incentivized based on their performance. Full experimental details can be found in the Supplementary Information.

Compensating the underestimation bias

Previous research on estimation tasks has shown that the distributions of raw estimates is generally right skewed, while the distribution of their logarithm is much more symmetric [19, 30, 38, 46]. Indeed, when considering large values, humans tend to think in terms of order of magnitude [45], making the logarithm of estimates a natural quantity to consider in estimation tasks [20].

The mean or median of log estimates is often used to measure the quality of collective judgments in such tasks (Wisdom of Crowds). Since distributions of log estimates for most quantities are closer to Laplace distributions than to Gaussian distributions [47], the median is more reliable than the mean[2] in estimating the Wisdom of Crowds [21].

Figure 1a shows that, within our domain (data are from a previous study [20]), there is a linear relationship between the median log estimate m and the log of the true value T: $m \sim \gamma \ \log(T)$, where $\gamma \approx 0.9$ is the slope of the relationship (the "shifted-median parameter"). Note that $\gamma < 1$ denotes the underestimation bias.

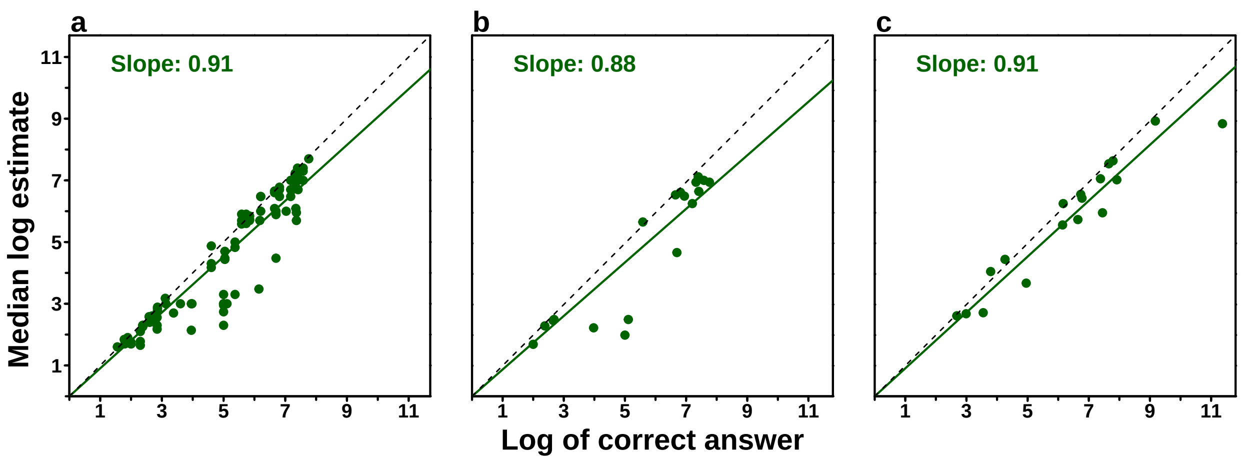


Figure 1: The relationship between the logarithm of the correct answer and the median of the logarithm of estimates for (a) 98 questions (one dot per question) taken from a former study [20] and (b, c) 36 questions from the current experiment. Among the 36 questions, 18 were already asked in the above cited study (b) and 18 were new (c). The slopes of the linear regression lines are 0.91 (a), 0.88 (b) and 0.91 (c), underlining the robustness of this linear trend. Note that slopes lower than 1 reflect the underestimation bias.

We used this relationship to construct, for each group and each question, a value m' (the "shifted-median value") aimed to compensate the underestimation bias, i.e., to approximate the (log of the) truth: $m' = m/\gamma \sim \log(T)$, with $\gamma = 0.9$. m' then served as a reference to select the estimates provided to the subjects in the Shifted-Median treatment.

[2]The median and the mean are the maximum likelihood estimators of the center of Laplace and Gaussian distributions, respectively.

Visual inspection confirms that the previously identified linear relationship not only holds for the same questions as in the previous study (half of our questions; Figure 1b), but also carries over to new questions (other half; Figure 1c), underlining its consistency. Supplementary Figure S2 shows that this relationship is present in general knowledge and numerosity questions, as well as for moderately large and very large quantities. All questions and participants' answers are included as Supplementary Material.

Results

Distribution of estimates

Following previous studies, we use the quantity $X = \log\left(\frac{E}{T}\right)$ to represent estimates, where E is the actual estimate of a quantity and T the corresponding true value [19, 20, 21, 48]. This normalization ensures that estimates of different quantities are comparable, and represents a deviation from the truth in terms of orders of magnitude. In the following, we will, for simplicity, refer to X as "estimates", with X_{p} referring to personal estimates and X_{s} to second estimates. Figure 2 shows the distributions of X_{p} (filled dots) and X_{s} (empty dots) in each treatment and number of shared estimates τ.

Confirming previous findings [20, 21, 48], we find narrower distributions after social information sharing across all conditions. This narrowing amounts to second estimates X_{s} being, on average, closer to the truth than the X_{p}. The distributions of X_{p} (solid lines) are simulated by drawing the X_{p} from Laplace distributions, the center (median) and width (average absolute deviation from the median) of which are taken from the experimental distribution of estimates for each question. Former studies have shown that distributions of estimates are indeed well approximated by Laplace distributions [47, 21]. In Supplementary Figure S3, we show the distribution of X_{p} when all conditions are combined. The good agreement between the data and the simulation further supports the Laplace distributions assumption. The distributions of X_{s} (dashed lines) are the predictions of our model presented below. One

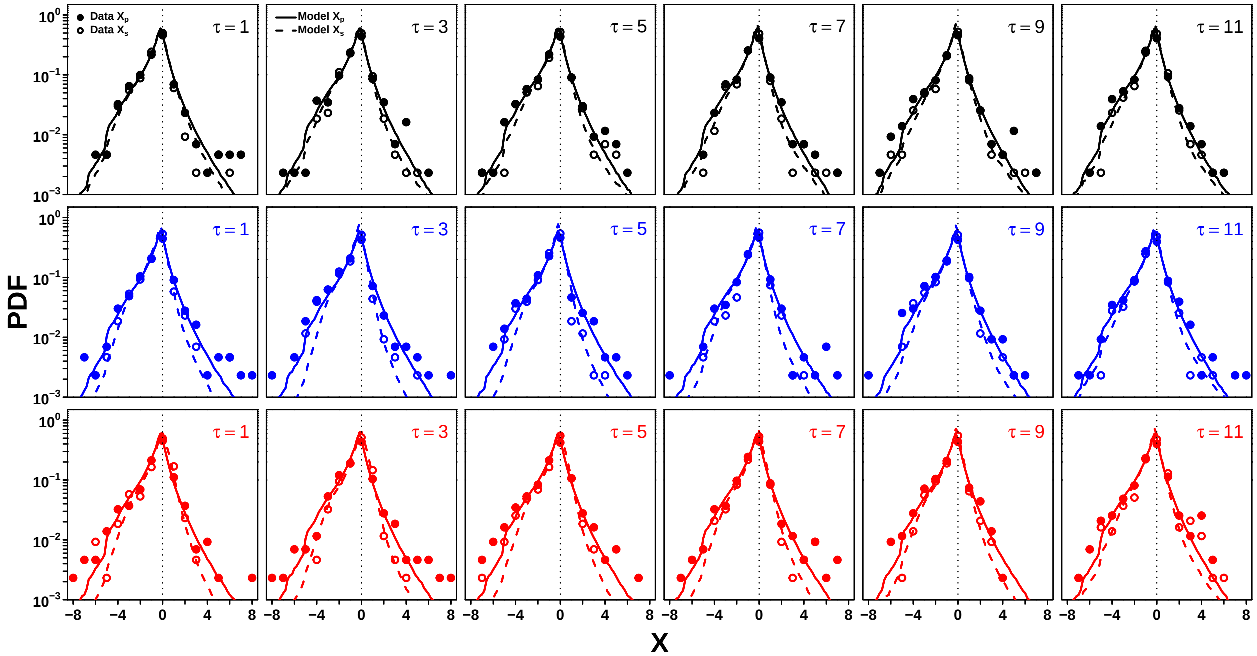


Figure 2: Probability density function (PDF) of personal estimates X_{p} (filled dots and solid lines) and second estimates X_{s} (empty dots and dashed lines) in the Random (black), Median (blue), and Shifted-Median (red) treatments, for each value of τ. Dots are the data and lines correspond to model simulations.

additional constraint was added in our simulations of both personal and second estimates: since in our experiment, actual estimates $E_{\mathrm{p,s}}$ are always greater than 1, we imposed that $X_{\mathrm{p,s}} > -\log(T)$, leading to a faster decay of the distribution for large negative log estimates.

Distribution of sensitivities to social influence S

Consistent with heuristic strategies under time and cognitive constraints [49, 50, 51], we assume that subjects, in evaluating a series of estimates, focus on the *central tendency* and *dispersion* of the estimates they receive as social information. These assumptions are also supported by other studies on estimation tasks [39, 52, 53]. Consistent with the logarithmic representation and Laplace distribution assumptions, we quantify the perceived central tendency and dispersion by the mean and average absolute deviation from the mean of the logarithms of the pieces of social information received, respectively.

We consider a subject's second estimate X_{s} as the weighted arithmetic mean[3] of their personal estimate X_{p} and the mean $M = \log(G)$ of the estimates received (G is the geometric mean of the actual estimates received): $X_{\mathrm{s}} = (1 - S)\, X_{\mathrm{p}} + S\, M$, where S is defined as the weight subjects assign to M, that we call the *sensitivity to social influence*. $S = 0$ thus implies that a subject keeps their personal estimate, and $S = 1$ that their second estimate equals the geometric mean of the estimates received. As we will show below, S depends on the number of estimates received and their dispersion.

In the following analysis of S, we will restrict S to the interval [-1, 2][4], thereby removing large values of S that may disproportionately affect measures based on S, in particular its average. Such large values of S are indeed meaningless as they are contingent on the way S is defined, and do not reflect a massive adjustment from X_{p} to X_{s}. Consider, for example, the case where $X_{\mathrm{p}} = 5$ and $M = 5.001$. Then, $X_{\mathrm{s}} = 5.1$ gives $S = 100$, while X_{s} is not very different from X_{p}. Such a restriction amounts to removing about 5.3% of the data.

Figure 3 shows that the distribution of S, in all treatments and values of τ, consists of a peak at $S = 0$ and a part that resembles a Gaussian distribution.

We thus assume that with probability P_0, subjects keep their initial estimate ($S = 0$), and with probability P_{g}, they draw an S in a Gaussian distribution of mean m_{g} and standard deviation σ_{g}. This assumption imposes the following relation:

$$\langle S \rangle = P_{\mathrm{g}}\, m_{\mathrm{g}}, \quad \text{i.e., } P_{\mathrm{g}} = \langle S \rangle / m_{\mathrm{g}}. \tag{1}$$

To determine the values of P_{g}, m_{g} and σ_{g} per condition (i.e., treatment and value of τ), we fit the distributions of S with the following distribution (using the "nls" function in R):

[3] Note that the arithmetic mean of the logs is equivalent to the log of the geometric mean.

[4] For plotting reasons, we actually restrict S to the interval [-1.05, 2.05].

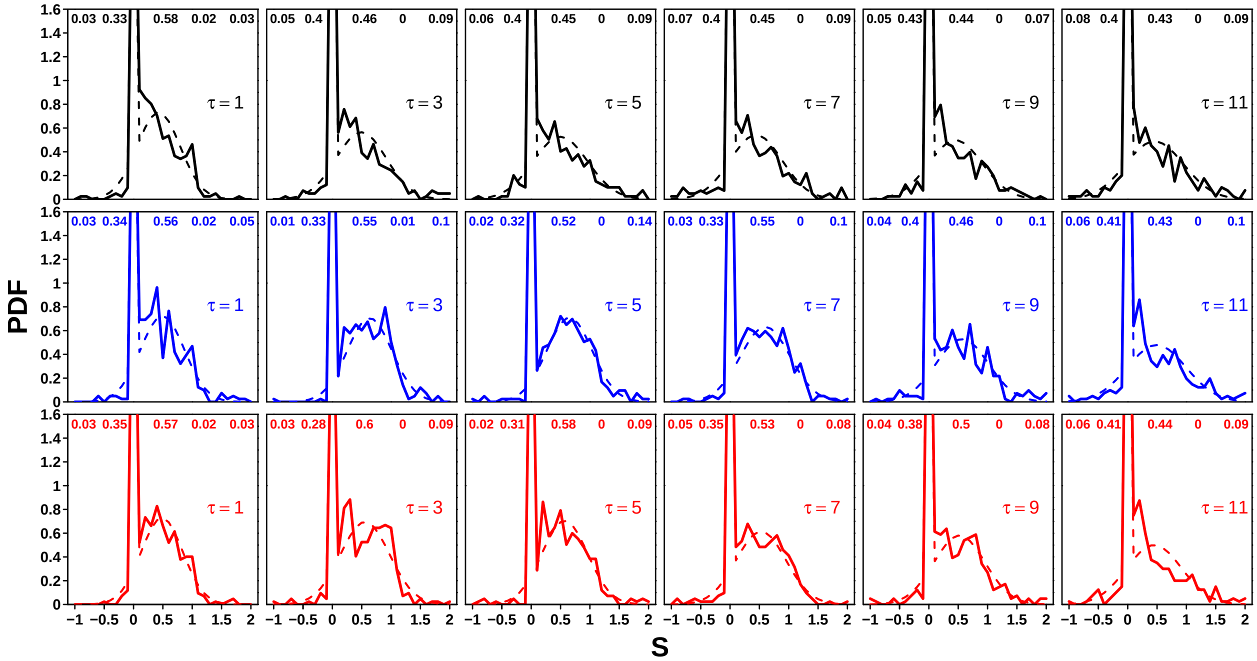


Figure 3: Probability density function (PDF) of sensitivities to social influence S in the Random (black), Median (blue), and Shifted-Median (red) treatments, for each value of τ. Solid lines are experimental data, and dashed lines fits using eq. 2. The experimental probabilities to contradict the social information ($S < 0$), to reject it ($S = 0$), to compromise with it ($0 < S < 1$), to adopt it ($S = 1$), and to overreact to it ($S > 1$) are shown on top of each graph.

$$f(S) = (1 - P_{\rm g})\ \delta(S) + P_{\rm g}\ \Gamma(S, m_{\rm g}, \sigma_{\rm g}), \text{with} \tag{2}$$

$$\Gamma(S, m_{\rm g}, \sigma_{\rm g}) = \frac{1}{\sqrt{2\pi}\,\sigma_{\rm g}}\ \exp\left[-\frac{(S - m_{\rm g})^2}{2\,\sigma_{\rm g}^2}\right], \tag{3}$$

where $P_{\rm g}$ is fixed by eq. 1, $\delta(S)$ is the Dirac distribution centered on 0, and $\Gamma(S, m_{\rm g}, \sigma_{\rm g})$ is the Gaussian distribution of mean $m_{\rm g}$ and standard deviation $\sigma_{\rm g}$.

Note that in previous studies, another peak was measured at $S = 1$, amounting to about 4% of answers [20, 21]. However, in our experiments, this peak was absent in almost all conditions, because when more than one estimate is shared, the second estimate is very unlikely to land exactly on the geometric mean of the social information. We, therefore, did not include it in the fit.

Dependence of P_g, m_g and σ_g on τ

Figure 4 shows P_g, m_g and σ_g against τ in each treatment. At $\tau = 1$, P_g and σ_g are comparable in all treatments, whereas m_g is higher in the Median and Shifted-Median treatments than in the Random treatment, indicating a higher tendency to follow the social information.

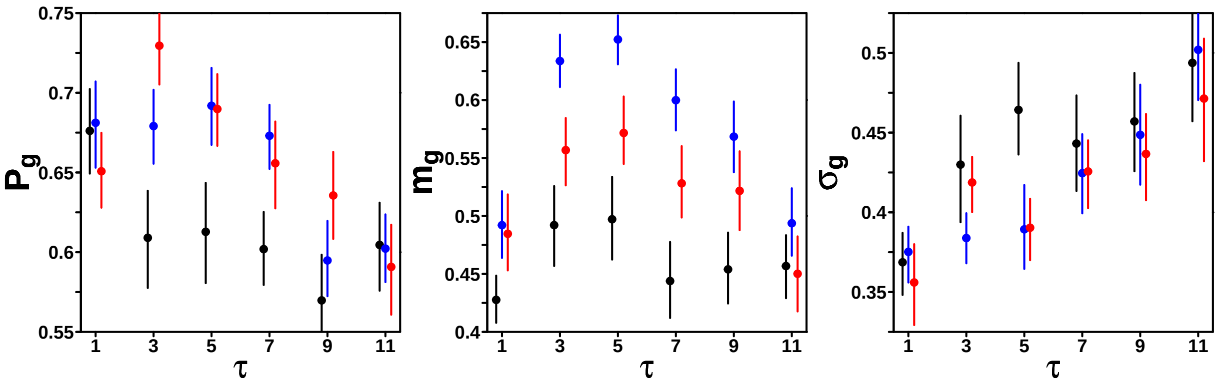


Figure 4: P_g, m_g and σ_g against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Error bars are computed using a bootstrap procedure described in the Materials and Methods, and roughly represent one standard error.

A similar pattern of social influence strength is observed at intermediate values of τ ($\tau = 3$, 5, 7, or 9), where P_g and m_g are substantially higher in the Median and Shifted-Median treatments than in the Random treatment. For σ_g, we observe a higher value in the Random treatment than both other treatments at $\tau = 3$ and 5, but not at higher levels of τ. Finally, at $\tau = 11$ the three measures are similar across treatments. This was expected since all three treatments are equivalent in this case (i.e., subjects receive all pieces of social information). It is worth noting that a previous study conducting a similar Random treatment [48], found very similar results.

Dependence of the dispersion σ on τ

One major difference between treatments that could help explain the above results lies in the dispersion $\sigma = \langle |X_\mathrm{SI} - M| \rangle$ of the estimates X_SI received as social information. Recall that the estimates received in the Median and Shifted-Median treatments were selected by proximity to a specific value (see Experimental Design), and are thus expected to be, on average, more similar to each other (i.e., to have a lower dispersion) than in the Random treatment. Figure 5 shows that, as expected, the average dispersion $\langle\sigma\rangle$ is substantially lower in the Median and Shifted-Median treatments than in the Random treatment.

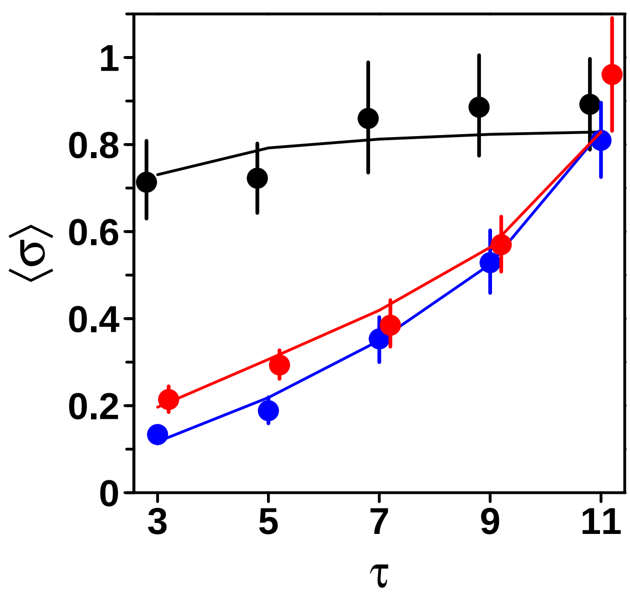


Figure 5: Average dispersion $\langle\sigma\rangle$ of the estimates received as social information against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments. $\langle\sigma\rangle$ is mostly independent of τ in the Random treatment, while it increases with τ in the Median and Shifted-Median treatments. Dots and error bars are the data and solid lines model simulations.

Moreover, $\langle\sigma\rangle$ increases with τ in these treatments, while it remains close to constant in the Random treatment. Expectedly, $\langle\sigma\rangle$ reaches a similar value in all treatments at $\tau = 11$. We thus expect the dependence of P_g, m_g and σ_g on τ observed in Figure 4 to be mediated by a dependence of these measures on σ.

Dependence of P_g, m_g and σ_g on the dispersion σ

Figure 6 shows P_g, m_g and σ_g as functions of the average dispersion of estimates received as social information $\langle\sigma\rangle$, for each combination of treatment and value of τ.

We find that P_g and m_g decrease linearly with $\langle\sigma\rangle$, reflecting a decreasing tendency to compromise with the social information as the dispersion of estimates received increases. On

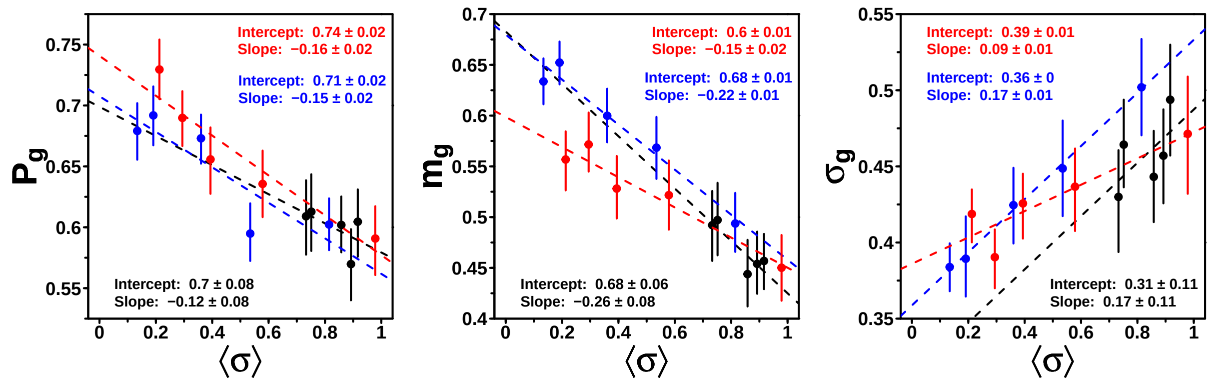


Figure 6: P_g, m_g and σ_g against the average dispersion of estimates received as social information $\langle\sigma\rangle$, in the Random (black), Median (blue) and Shifted-Median (red) treatments. Each dot corresponds to a specific value of τ. Values at $\tau = 1$ were excluded since there is no dispersion at $\tau = 1$. Dashed lines show linear fits per treatment.

the contrary, σ_g increases linearly with $\langle\sigma\rangle$, suggesting that the diversity of subjects' response to social influence increases with the diversity of pieces of social information received.

Dependence of S on the dispersion σ: similarity effect

As described above, P_g and m_g combined determine the average sensitivity to social influence. Figure 7 shows how $\langle S\rangle = P_\mathrm{g}\, m_\mathrm{g}$ – where the values of P_g and m_g are taken from Figure 6a and b – varies with the average dispersion of estimates received $\langle\sigma\rangle$.

Consistently with Figure 6a–b, $\langle S\rangle$ decreases linearly with $\langle\sigma\rangle$ in all treatments. We call this the *similarity effect*. Moreover, this linear dependence of $\langle S\rangle$ on σ appears to be treatment-independent, as a linear regression over all points fits the data very well.

Note that since we found a linear dependence of P_g ($P_\mathrm{g} = a + b\,\langle\sigma\rangle$) and m_g ($m_\mathrm{g} = a' + b'\,\langle\sigma\rangle$) on $\langle\sigma\rangle$, the dependence of $\langle S\rangle = P_\mathrm{g}\, m_\mathrm{g}$ on $\langle\sigma\rangle$ could have been quadratic. Yet, the quadratic term $b\,b'\,\langle\sigma\rangle^2$ is of the order $0.2 \times 0.2 \times 0.5^2 = 0.01$, and thus negligible.

Dependence of S on $D = M - X_\mathrm{p}$: distance and asymmetry effect

In previous studies where subjects received as social information the average estimate of other group members, S depended linearly on the distance $D = M - X_\mathrm{p}$ between the personal

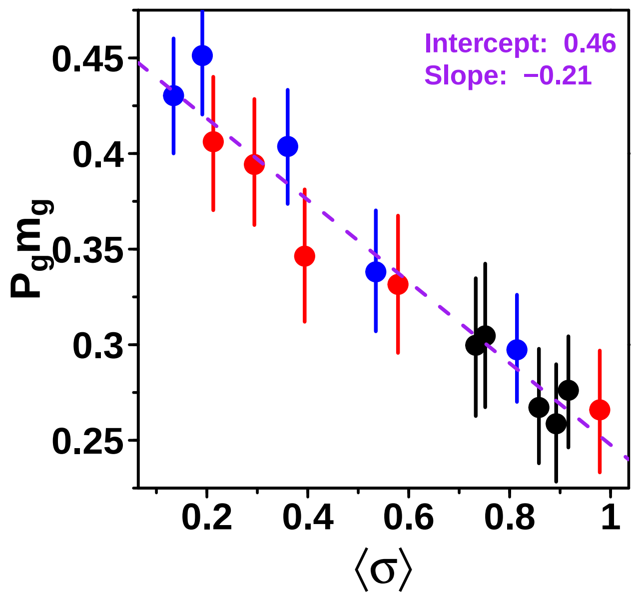


Figure 7: $P_{\mathrm{g}}\, m_{\mathrm{g}}$ against the average dispersion of estimates received as social information $\langle\sigma\rangle$ in the Random (black), Median (blue), and Shifted-Median (red) treatments. $P_{\mathrm{g}}\, m_{\mathrm{g}}$ decreases linearly with $\langle\sigma\rangle$ in all treatments. Each dot corresponds to a specific value of τ. The purple dashed line shows a linear regression over all points.

estimate X_{p} and the average social information M [20, 21]. This effect is known as the *distance effect*:

$$\langle S\rangle(D) = \alpha + \beta\,|D|. \tag{4}$$

Figure 8 shows the distance effect per condition, showing that the further the social information is away from the personal estimate, the stronger it is taken into account.

For each condition (and in agreement with a recent study [48]), we find that the center of the cusp relationship is located at $D = D_0 < 0$, rather than at $D = 0$. Moreover, the left and right slopes (coined β_- and β_+ respectively) are not always similar, requiring us to fit the slopes separately. These effects combined result in an asymmetric use of social information whereby social information that is higher than the personal estimate is weighted more than social information that is lower than the personal estimate. This effect is known as the *asymmetry effect* and we will discussed it in more detail below.

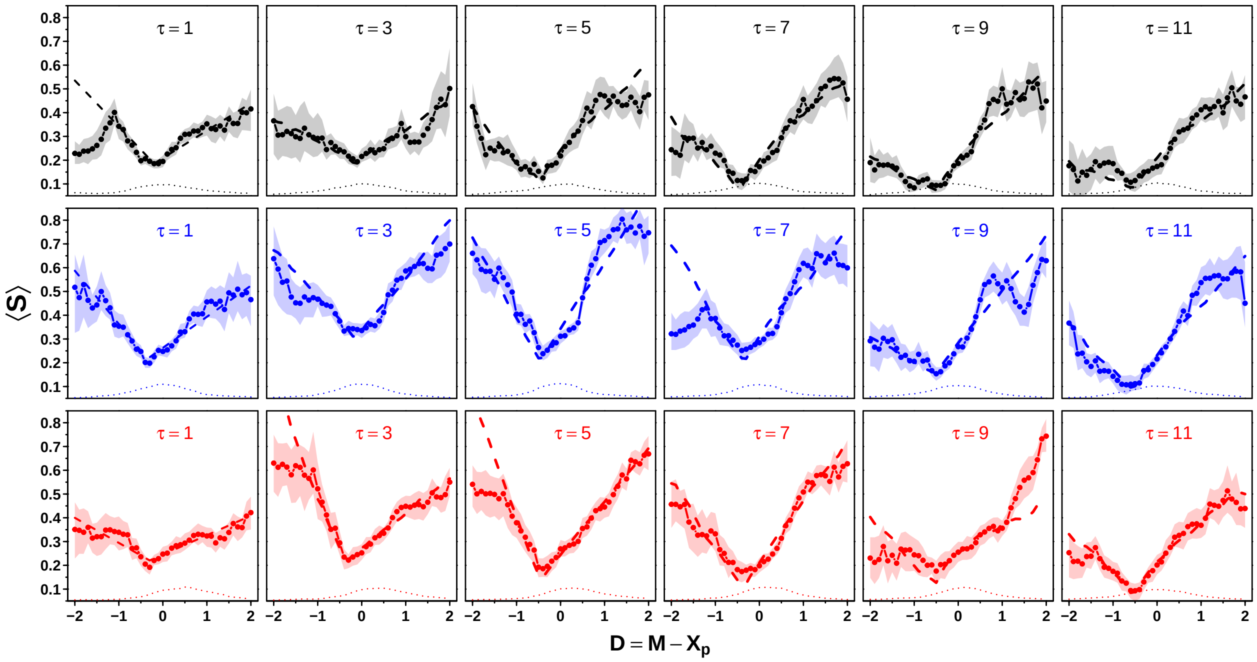


Figure 8: Average sensitivity to social influence $\langle S \rangle$ against the distance $D = M - X_{\mathrm{p}}$ between the personal estimate X_{p} and average social information M, in the Random (black), Median (blue), and Shifted-Median (red) treatments for all values of τ. Dots are the data, and shaded areas represent the error (computed using a bootstrap procedure described in the Materials and Methods) around the data. Dashed lines are fits using eq. 5, and dotted lines at the bottom of each panel show the density distribution of the data (in arbitrary units).

Finally, Figure 7 shows that we need to include a dependence of $\langle S \rangle$ on σ. Following Figure 7, we assume this dependence to be linear (with slope β'). Taking these results together, we thus arrive at the following fitting function:

$$\langle S \rangle(D, \sigma, \tau) = \alpha(\tau) + \beta_{\pm}(\tau)\,|D - D_0(\tau)| + \beta'(\tau)\,\sigma, \qquad (5)$$

where α, $\beta_{\pm}$, β' and D_0 can *a priori* depend on τ. Visual inspection was used to fix D_0, with a precision of less than 0.1, while all other parameters were fitted by minimizing least squares analytically (see Materials and Methods for details of the fitting procedure). At $\tau = 1$, $\sigma = 0$, therefore, β' was excluded from the parameter fitting for this case.

Figure 9 shows the fitted values against τ for each treatment, and suggests that these parameters do not systematically vary with τ. We next introduce a model of social information integration, in which we will, therefore, assume that these parameters are independent of τ, and equal to their average (when $\tau > 1$, see below).

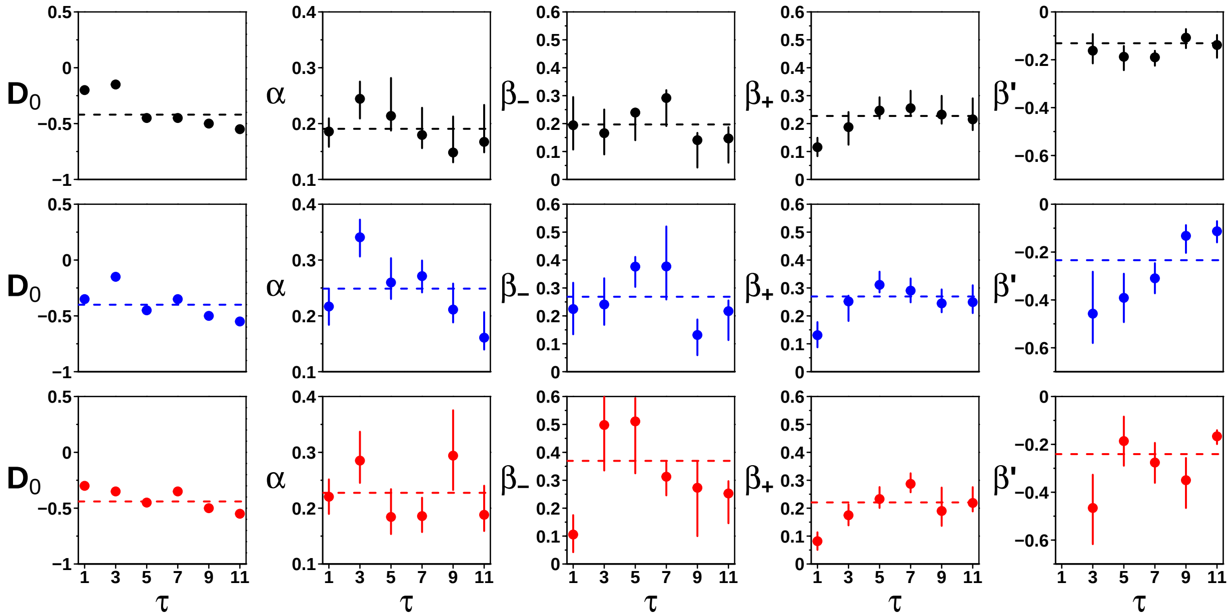


Figure 9: Fitted parameter values of D_0, α, β_-, β_+, and β' against τ in the Random (black), Median (blue), and Shifted-Median (red) treatments. Parameters do not show any clear dependence on τ in any treatments. Dashed lines are averages over all values of $\tau > 1$.

Model of social information integration

The model is based on eq. 5 and is an extension of a model developed in [48] (which itself builds on [20, 21]). The key effect we add is the dependence of subjects' sensitivity to social influence on the dispersion of estimates received as social information. This is done because the Median and Shifted-Median treatments select relatively similar pieces of social information to share, which heavily impacts social influence (Figures 6 and 7).

The model uses log-transformed estimates X as its basic variable, and each run of the model closely mimics our experimental design. For a given quantity to estimate in a given condition (i.e., treatment and number of shared estimates), $N = 12$ agents first provide their personal estimate X_{p}. Following Figure 2, these personal estimates are drawn from Laplace distributions, the center and width of which are respectively the median m_{p} and dispersion $\sigma_{\mathrm{p}} = \langle |X_{\mathrm{p}} - m_{\mathrm{p}}| \rangle$ of the experimental personal estimates of the quantity.

Next, agents receive as social information τ personal estimates from other agents in the group, selected according to the selection procedure of the respective treatment (see Experimental Design). Following Figure 3, agents either keep their personal estimate ($S = 0$) with probability P_0, or draw an S in a Gaussian distribution of mean m_g and standard deviation σ_g with probability P_g. According to eq. 1, $P_\mathrm{g} = \langle S \rangle / m_\mathrm{g}$, and $P_0 = 1 - P_\mathrm{g}$. The calculation of $\langle S \rangle$ is based on the mean M and dispersion σ of these estimates received, and follows eq. 5:

$$P_\mathrm{g}(D, \sigma, \tau) = \langle S \rangle (D, \sigma, \tau) / m_\mathrm{g}(\sigma) = \big(\alpha(\tau) + \beta_\pm(\tau)\ |D - D_0(\tau)| + \beta'(\tau)\ \sigma\big) / m_\mathrm{g}(\sigma), \qquad (6)$$

Finally, once an S is drawn for each agent, agents update their estimate according to:

$$X_\mathrm{s} = (1 - S)\, X_\mathrm{p} + S\, M. \qquad (7)$$

At $\tau = 1$, the values given to P_g, m_g and σ_g were taken from Figure 4. When sharing more than 1 estimate (i.e., $\tau > 1$), the linear dependences of these parameters on the dispersion of the social information $\langle \sigma \rangle$, shown in Figure 6, were used. Similarly, the values of D_0, α, β_- and β_+ at $\tau = 1$ were directly taken from Figure 9, while values of D_0, α, $\beta_\pm$ and β' at $\tau > 1$ were averaged over τ, and these averages were implemented in the model. This separation is done because the fitting was qualitatively different for $\tau > 1$ and $\tau = 1$, β' being absent in the latter (no dispersion at $\tau = 1$).

Next to this full model, we also evaluated two simpler models, leaving out either the similarity effect ($\beta' \sigma$ term) or the asymmetry effect ($D_0 < 0$ and $\beta_- \neq \beta_+$), to evaluate the importance of both effects in explaining the empirical patterns. Supplementary Figures S4 to S8 show the predictions when excluding the similarity effect, and Supplementary Figures S9 to S13 when excluding the asymmetry effect.

All model simulations results shown in the figures are averages over 10,000 runs. The full model reproduces well the distributions of estimates (Figure 2), and the dependence of $\langle\sigma\rangle$ on τ (Figure 5). We now use the model to analyze the impact of τ on sensitivity to social influence and estimation accuracy in each treatment.

Impact of τ on sensitivity to social influence S

Figure 10a shows how $\langle S\rangle$ varies with τ in all treatments. We find that in the Median and Shifted-Median treatments, $\langle S\rangle$ increases sharply between $\tau = 1$ and $\tau = 3$, before decreasing steadily, consistent with the patterns of P_{g} and m_{g} in Figure 4. In the Random treatment $\langle S\rangle$ is largely independent of τ. At $\tau = 11$, all conditions (again) converge.

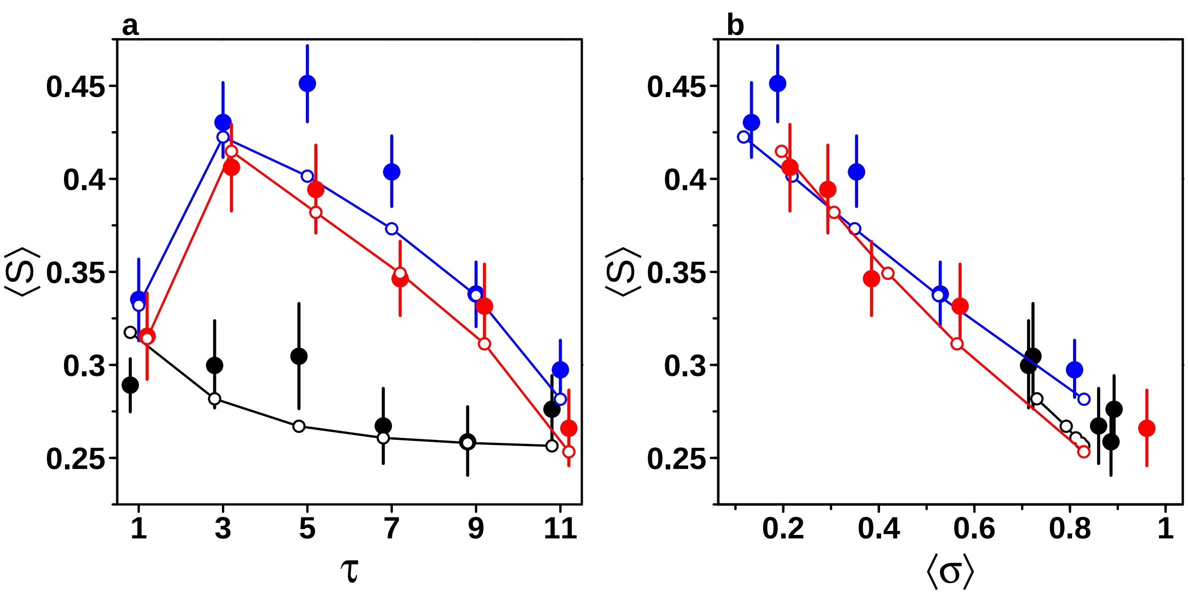


Figure 10: Average sensitivity to social influence $\langle S\rangle$ against (a) the number of shared estimates τ and (b) the average dispersion of estimates received $\langle\sigma\rangle$, in the Random (black), Median (blue), and Shifted-Median (red) treatments. (a) In the Random treatment there is only a minor dependence of $\langle S\rangle$ on τ. In the Median and Shifted-Median treatments, we find an inverse-U shape relationship with τ. This is due to the similarity effect as shown in (b): a linear decrease of $\langle S\rangle$ with $\langle\sigma\rangle$ when $\tau > 1$. Filled dots are the data, while empty dots and solid lines are model simulations.

These patterns result from the similarity effect shown in Figure 10b: $\langle S\rangle$ decreases as the dispersion of estimates received increases, when $\tau > 1$. Whereas in the Median and Shifted-Median treatments the different levels of τ correspond to different levels of dispersion (Figure 5), and thus different levels of $\langle S\rangle$, this effect is not present in the Random treatment.

Note that consistently with the relation $\langle S \rangle = P_{\mathrm{g}}\, m_{\mathrm{g}}$, the experimental values in Figure 10b are the same as those of Figure 6.

The full model reproduces the empirical results well. When removing the dependence on σ from the model (and re-fitting the parameters accordingly), the inverse-U shape in the Median and Shifted-Median is attenuated, and the decrease of $\langle S \rangle$ with $\langle \sigma \rangle$ is underestimated (Figure S4). This demonstrates that the similarity effect is key to explaining the patterns of sensitivity to social influence.

Impact of τ on S when $D < 0$ and $D > 0$

A more intuitive way to understand the result that $D_0 < 0$ and $\beta_+ > \beta_-$, is that subjects' sensitivity to social influence is on average higher when $D > 0$ (i.e., when the average social information received by subjects is higher than their personal estimate) than when $D < 0$ (i.e., when the average social information received by subjects is lower than their personal estimate). Figure 11 shows this so-called *asymmetry effect*, which is reproduced well by the full model.

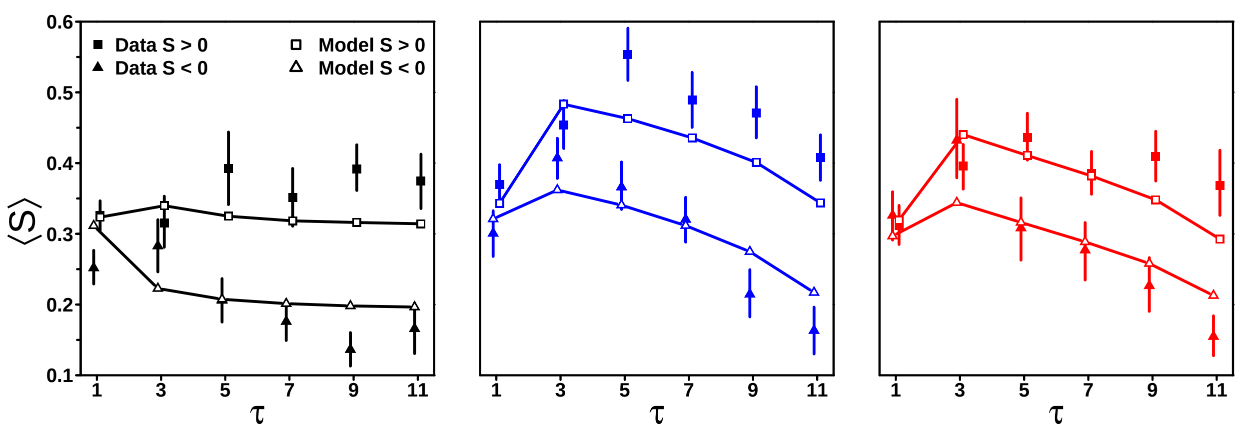


Figure 11: Average sensitivity to social influence $\langle S \rangle$ against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments, when the average social information M is higher than the personal estimate X_{p} ($D = M - X_{\mathrm{p}} > 0$; squares) and when it is lower ($D < 0$; triangles). Subjects follow the social information more on average when M is higher than X_{p}, than when it is lower. Filled symbols represent the data, while solid lines and empty symbols are model simulations. Supplementary Table S1 shows the percentage of cases when $D < 0$ and $D > 0$ in all conditions.

Below, we will show that this effect also drives improvements in estimation accuracy after social information sharing. Supplementary Figure S8 shows that the model without the asymmetry effect is unable to reproduce the higher sensitivity to social influence when $D > 0$ than when $D < 0$.

Improvements in estimation accuracy: herding effect

In line with previous works [20, 21, 48] we define, for a given group in a given condition, (i) the *collective accuracy* as the absolute value of the median of all individuals' estimates of all quantities in that group and condition: $\left|\text{Median}_{i,q}(X_{i,q})\right|$ (where i runs over individuals and q over quantities/questions), and (ii) the *individual accuracy* as the median of the absolute values of all individuals' estimates: $\text{Median}_{i,q}\big(|(X_{i,q})|\big)$. The closer to 0, the higher the accuracy. Collective accuracy represents the distance of the median estimate to the truth, and individual accuracy the median distance of individual estimates to the truth. Figure 12 shows how collective and individual accuracy depend on τ in each treatment.

Collective accuracy improves mildly – but not negligibly – in the Random and Median treatments. This improvement is due to the asymmetry effect (Figure 11), which partly counteracts the human tendency to underestimate quantities [20, 26, 27, 28]. Indeed, giving more weight to social information that is higher than one's personal estimate shifts second estimates toward higher values, thus improving collective accuracy. The model without the asymmetry effect is unable to predict this improvement in collective accuracy (Supplementary Figure S9).

In the Shifted-Median treatment the improvement in collective accuracy is substantially higher, especially at low values of τ. This is a consequence of the selection procedure of the pieces of social information. As shown in Figure 10, participants have a tendency to partially follow the social information ($0 < \langle S \rangle < 1$ in all conditions, a.k.a. *herding effect*). Although there are no substantial differences in $\langle S \rangle$ between the Median and Shifted-Median treatment, the estimated received as social information overestimate the group median in the

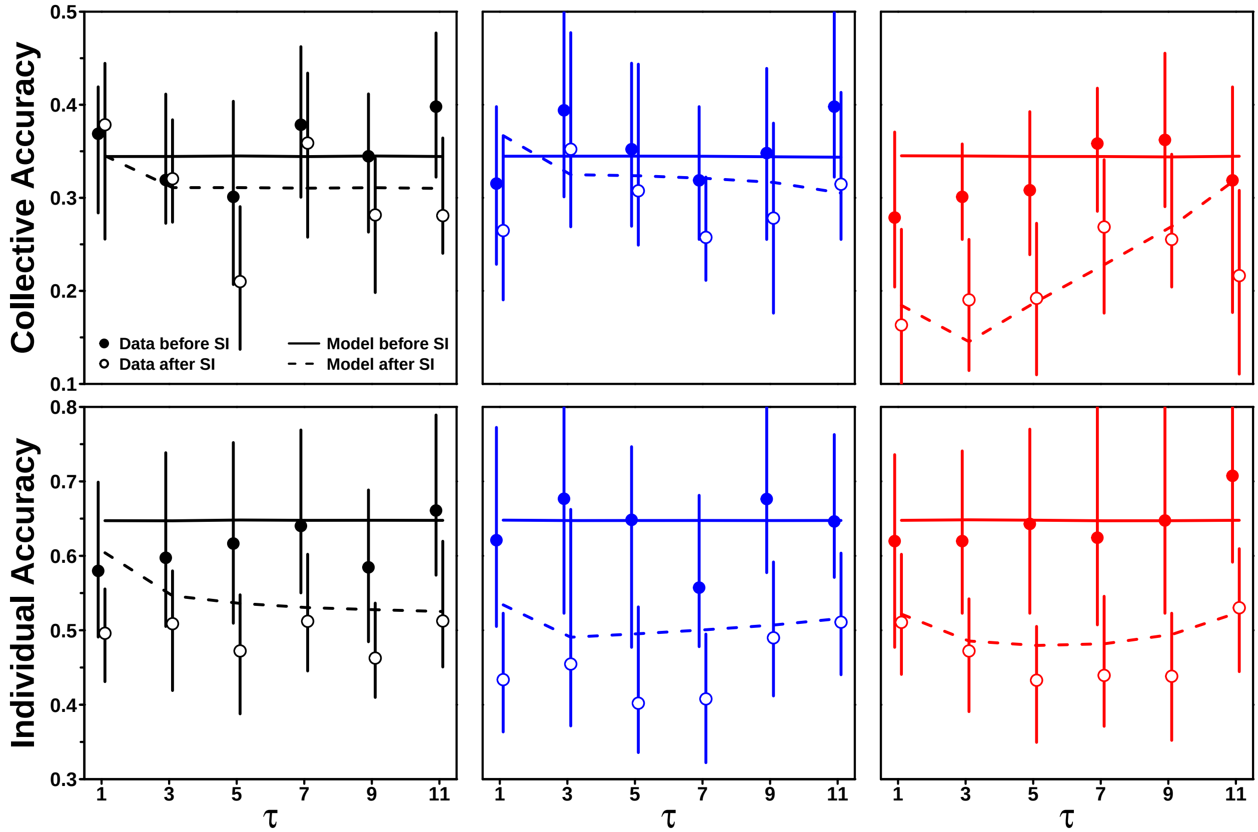


Figure 12: Collective and individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Values closer to 0 indicate higher accuracy. Solid and dashed lines are model simulations before and after social information sharing, respectively.

Shifted-Median treatment. A similar level of $\langle S \rangle$ thus shifts seconds estimates toward higher values (compared to the Median treatment), thereby partly countering the underestimation bias and boosting collective accuracy.

For individual accuracy we find substantial improvements in all conditions, with slightly higher improvements in the Median and Shifted-Median treatments than in the Random treatment, due to the similarity effect which boosts social information use in these treatments (Figure 10). This confirms previous studies showing that higher levels of social information use (when $0 < \langle S \rangle < 0.5$) increase the narrowing of the distribution of estimates (Figure 2), thereby increasing individual accuracy [19, 20].

Impact of D on estimation accuracy

Because subjects behave differently when receiving social information that is higher ($D > 0$) or lower ($D < 0$) than their personal estimate, we next study how these different scenarios impact accuracy. Figure 13 shows this for individual accuracy.

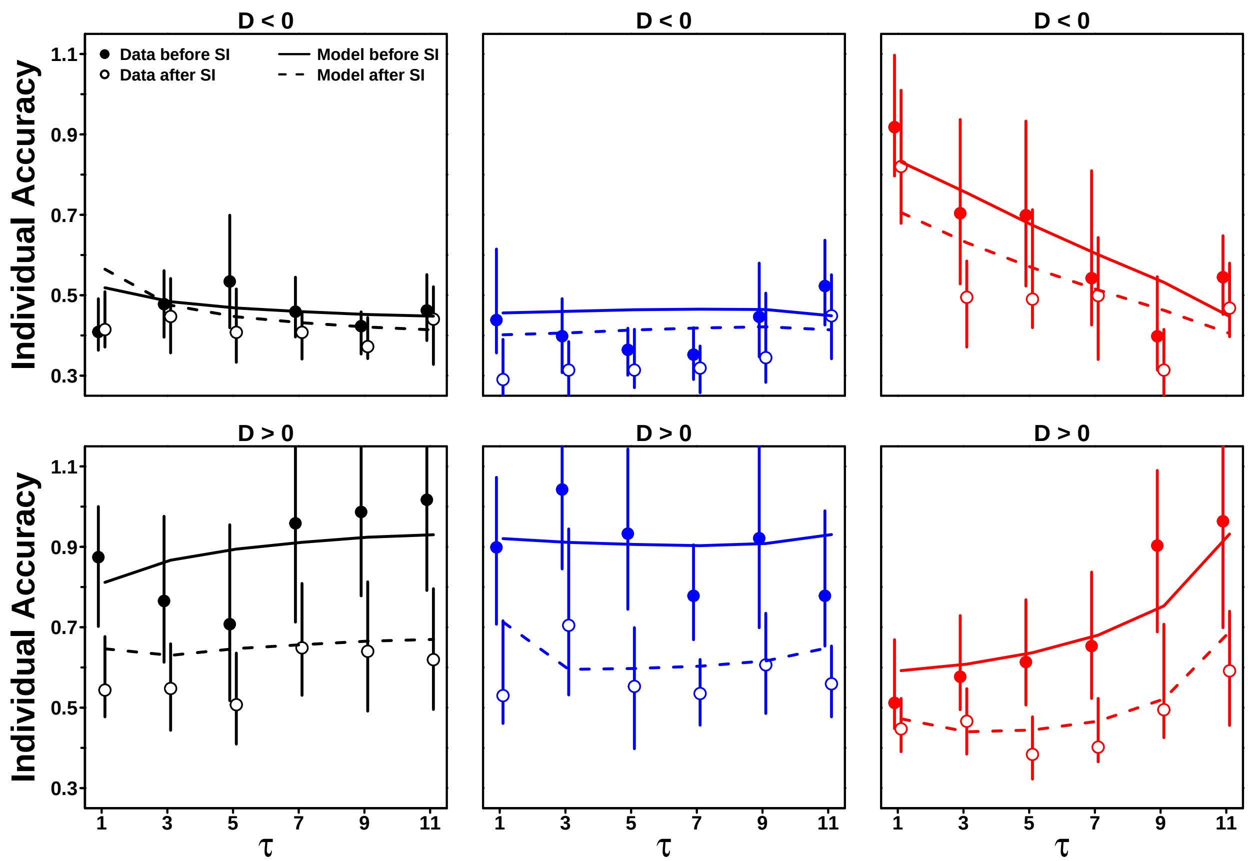


Figure 13: Individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. The population was separated into subjects' answers where the average social information received M was lower than their personal estimate X_p ($D = M - X_\mathrm{p} < 0$) and subjects' answers where the average social information received was higher than their personal estimate ($D > 0$). Solid and dashed lines are model simulations before and after social information sharing, respectively. Individual accuracy improves mildly for $D < 0$, but substantially for $D > 0$.

We find that, in the Random and Median treatments, subjects were significantly more accurate when $D < 0$ than when $D > 0$ before social information sharing. This is a consequence of the underestimation bias, as estimates in the former (latter) case are, on average, more likely to be above (below) the median estimate of the group – and therefore closer to (farther from) the truth. In the Shifted-Median treatment, however, we observe a more complex pattern: (i) at low values of τ, individual accuracy is worse before social information sharing in this treatment than in the Random and Median treatments when $D < 0$, while it is better when $D > 0$. This reversed pattern suggests that the shifted-median values tend, on average, to slightly overestimate the truth; (ii) individual accuracy improves with τ when $D < 0$, but declines with it when $D > 0$. As τ increases, the average social information indeed decreases until it is the same as in both other treatments at $\tau = 11$. In all conditions, individual accuracy improves mildly after social information sharing when $D < 0$, while it improves substantially when $D > 0$. The model is in good agreement with the data. Supplementary Figure S10 shows the equivalent figure for collective accuracy, showing qualitatively similar results.

Impact of S on estimation accuracy

Finally, we studied how subjects' sensitivity to social influence affects estimation accuracy, by separating subjects' answers into those for which S was either below or above the median value of S in that condition. Figure 14 shows individual accuracy for both categories.

Subjects in the below-median category provided more accurate personal estimates than those in the above-median category. It is well-known that more accurate individuals use less social information, and this insight has also been used to improve collective estimations [36]. This result is tied to the distance effect (Figure 8): subjects use social information the least when their initial estimate is close to the average social information, which is itself, on average, close to the truth.

Because subjects in the below-median category disregard, or barely use, social information, they do not improve in accuracy after social information sharing. On the contrary, subjects in the above-median category tend to compromise with the social information, thereby substantially improving in individual accuracy after social information sharing, and reaching similar levels of accuracy as the below-median category. The model accurately reproduces

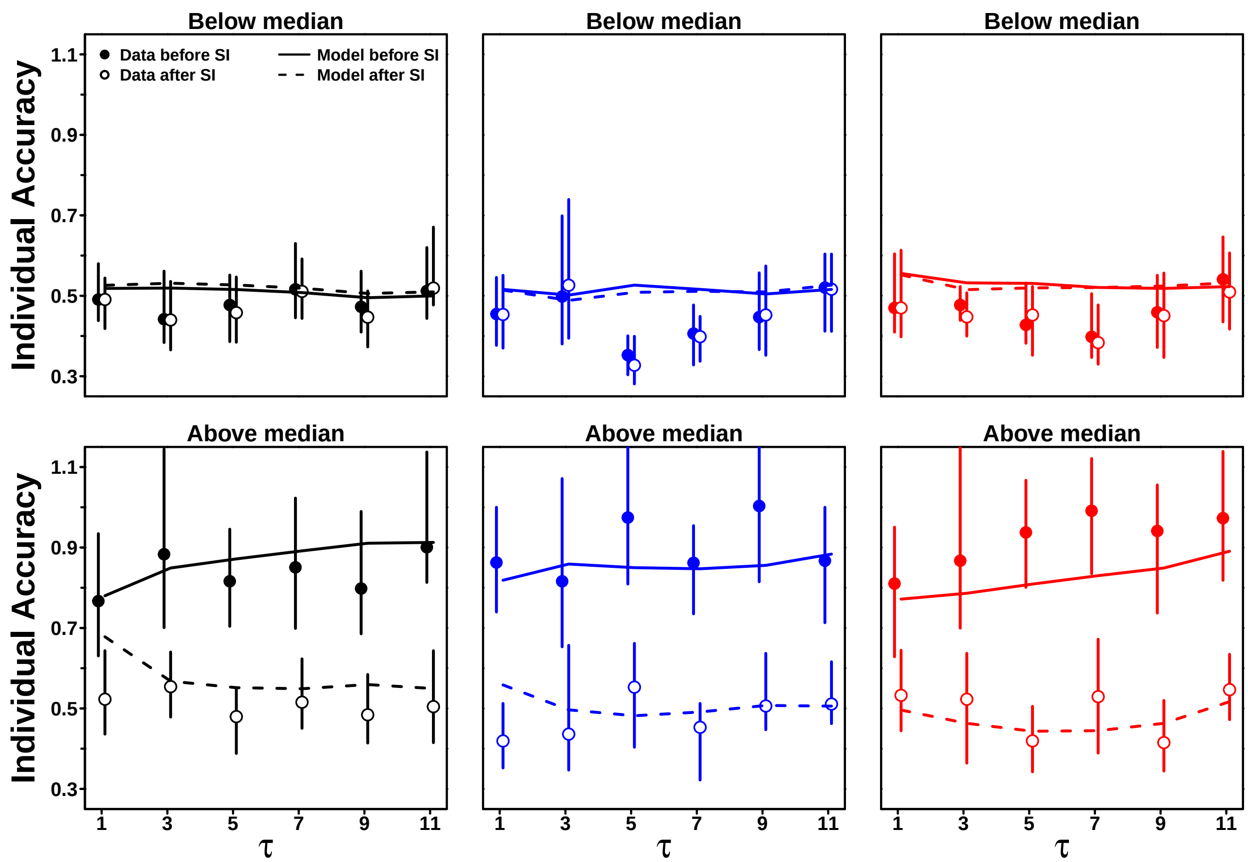


Figure 14: Individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. In each condition, the subjects' answers were separated according to their corresponding value of S with respect to the median of S. Solid and dashed lines are model simulations before and after social information sharing, respectively. When S is lower than the median, the subjects tend to keep their initial estimate, and individual accuracy therefore does not change. When S is higher than the median, the subjects tend to compromise with the social information, resulting in high improvements.

these results, which also are in agreement with former findings [20, 21, 48]. Supplementary Figure S11 shows the equivalent figure for collective accuracy, showing qualitatively similar patterns, albeit with substantially higher improvements in the Shifted-Median treatment for the above-median category.

Discussion

We studied the impact of the number of estimates presented to individuals in human groups, and the way they are selected, on collective and individual accuracy in estimating large quantities, and identified four key mechanisms underlying social information integration:

(i) subjects give more weight to the social information when the distance between the average social information and their own personal estimate increases (*distance effect*). This effect has been found in several previous studies [20, 21, 48];

(ii) subjects give more weight to the central tendency of multiple estimates when it is higher than their own personal estimate, than when it is lower. This asymmetry effect, also found in [26, 48], shifts second estimates toward higher values, thereby partly compensating the underestimation bias and improving collective accuracy. The asymmetry effect suggests that people are able to selectively use social information in order to counterbalance the underestimation bias, even without external intervention (Random treatment);

(iii) subjects follow social information more when the estimates are more similar to each other (*similarity effect*). Previous studies have shown that similarity in individuals' judgments correlates with judgment accuracy [54, 55], suggesting that following pieces of social information more when they are more similar is an adaptive strategy to increase the quality of one's judgments. Our selection method in the Median and Shifted-Median treatments capitalized on this effect as it selected relatively similar pieces of social information, thereby counteracting the human tendency to underuse social information [20, 56, 57], resulting in higher individual improvement in both treatments than in the Random treatment;

(iv) subjects tend to partially copy each other (*herding effect*), leading to a convergence of estimates after social information sharing, and therefore to an improvement in individual accuracy in all treatments. This effect is adaptive in most real-life contexts, as personal information is often limited and insufficient, such that relying on social information, at least partly, is an efficient strategy to make better judgments and decisions. Moreover, note that contrary to popular opinion, convergence of estimates need not yield negative outcomes (like impairing the Wisdom of Crowds [19, 31, 36]): even if the average opinion is biased, sharing opinions may temper extreme ones and improve the overall quality of judgments [58]. This tendency to follow the social information has another important consequence: it is possible to influence the outcome of collective estimation processes in a desired direction. In the Shifted-Median treatment, we showed that subjects' second estimates could be "pulled" towards the truth, thus improving collective accuracy. This is an example of *nudging*, also demonstrated in other contexts [59]. Previous studies have shown that the same tendency can also lead, under certain conditions, to dramatic situations in which everybody copies everybody else indiscriminately ("herd behavior") [60].

Next, we developed an agent-based model to study the importance of these effects in explaining the observed patterns. The model assumes that subjects have a fast and intuitive perception of the central tendency and dispersion of the estimates they receive, coherent with heuristic strategies under time and computational constraints [49, 50, 51], and consistent with previous findings [39, 52, 53]. By using simpler models excluding either the asymmetry effect or similarity effect, we demonstrated that these effects are key to explaining the empirical patterns of sensitivity to social influence and estimation accuracy. It is conceivable that the strategies used by people when integrating up to 11 pieces of social information in their decision-making process are very diverse and complex. Yet, despite its relative simplicity, our model is in good agreement with the data, underlining the core role of these effects in integrating several estimates of large quantities.

Our goal was to test a method to improve the quality of individual and collective judgments in social contexts. The method exploits available knowledge about cognitive biases in a given domain (here the underestimation of large quantities in estimation tasks) to select and provide individuals with relevant pieces of social information to reduce the negative effects of these biases. A previous study also manipulated the social information presented to the subjects in order to improve the accuracy of their second estimates [21]. However, at variance with our study, the correct answer to each question was known *a priori*, and exploited by "virtual influencers" providing (purposefully) incorrect social information to the subjects, specifically selected to counter the underestimation bias. Our method avoids such deception, and extends to situations in which the estimation context is known, but not the truth itself.

Another previous study exploited the underestimation bias by *recalibrating* personal estimates, thereby also successfully counteracting the underestimation bias [26]. Supplementary Figure S12 compares our Shifted-Median treatment to a direct recalibration of personal estimates, where all X_p are divided by $\gamma = 0.9$. Collective accuracy improves similarly under both methods. Individual accuracy, however, degrades with the recalibration method, while it strongly improves with the Shifted-Median method. Our method thus outperforms a mere recalibration of personal estimates. Moreover, note that recalibrating initial estimates may be useful from an external assessor's point of view, but does not provide participants with an opportunity to improve their accuracy, individually or collectively.

Our method may, in principle, be applied to different domains. Future work could, for instance, test this method in domains where overestimation dominates, by defining a shifted-median below the group median; or in domains where the quantities to estimate are negative (or at least not strictly positive) or lower than one (i.e., negative in log). Another interesting direction for future research would be to explore ways to refine our method. Supplementary Figure S13 and S14 show that collective and individual accuracy improved more for very large quantities than for moderately large ones, although the levels of underestimation are similar in both cases (Figure S2b). This suggests that the linear relationship between the median (log) estimates and the (log of the) true value may be insufficient to fully characterize this domain of estimation tasks. Considering other distributional properties, such as the dispersion, skewness and kurtosis of the estimates received, could help to fine tune the selection method to further boost accuracy.

Finally, let us point out that our population sample consisted of German undergraduate students. A previous cross-cultural study conducted in France and Japan, using a similar paradigm, found similar levels of underestimation in both countries, albeit slightly higher levels of social information use in Japan [20]. This suggests that our observed underestimation bias is widespread in this domain, though systematic comparison of the levels of bias and social information use in different (sub-)populations is still lacking. Filling this gap could represent a major step forward in research on social influenceability and cognitive biases.

To conclude, we believe that the mechanisms underlying social information use in estimation tasks share important commonalities with related fields (e.g., opinion dynamics [61]), and that our method has the potential to inspire research in such fields. For instance, one could imagine reducing the in-group bias by extending the amount of discrepant/opposite views presented to individuals in well-identified opinion groups. Implementing methods similar to ours in recommender systems and page-ranking algorithms may thus work against filter bubbles and echo chambers, and eventually reduce polarization of opinions [62]. Similarly, it is conceivable that the effects of well-known cognitive biases such as the confirmation [63] or overconfidence bias [64] could be dampened by strategically sharing social information.

Materials and Methods

Computation of the error bars

The error bars indicate the variability of our results depending on the $N_{\mathrm{Q}} = 36$ questions presented to the subjects. We call x_0 the actual measurement of a quantity appearing in the figures by considering all N_{Q} questions. We then generate the results of $N_{\mathrm{exp}} = 1,000$ new effective experiments. For each effective experiment indexed by $n = 1, ..., N_{\mathrm{exp}}$, we randomly draw $N'_{\mathrm{Q}} = N_{\mathrm{Q}}$ questions among the N_{Q} questions asked (so that some questions can appear several times, and others may not appear) and recompute the quantity of interest which now takes the value x_n. The upper error bar b_+ for x_0 is defined so that $C = 68.3\,\%$ (by analogy with the usual standard deviation for a normal distribution) of the x_n greater than x_0 are between x_0 and $x_0 + b_+$. Similarly, the lower error bar b_- is defined so that $C = 68.3\,\%$ of the x_n lower than x_0 are between $x_0 - b_-$ and x_0. The introduction of these upper and lower confidence intervals is adapted to the case when the distribution of the x_n is unknown and potentially not symmetric.

Fitting procedure used in Figure 8

Each combination of treatment and number of shared estimates contains 432 estimates. When binning data, one has to trade off the number of bins (thus displaying more detailed patterns) and the size of the bins (thus avoiding too much noise). In Figure 8, the noise within each condition was relatively high when using a bin size below 1. However, bins of size 1 were hiding the details of the relationship between $\langle S \rangle$ and D, especially the location of the bottom of the cusp.

To circumvent this problem, we use a procedure that is well adapted to such situations. First, remark that a specific binning leaves one free to choose on which values the bins are centered. For instance, a set of 5 bins centered on -2, -1, 0, 1 and 2 is as valid as a set of 5 bins centered on -2.5, -1.5, -0.5, 0.5, and 1.5, as the *same* data are used in both cases. Both sets of points produced are replicates of the same data, but we now have 10 points instead of 5.

In each panel of Figure 8, we used such a moving center starting the first bin at -2, and the last one at +2, producing histograms (of bin size 1) in steps of 0.1 for the bin center. This replicated the data 9 times, thus having overall 10 replicates and 50 points, instead of 5. We then removed the values beyond $D = 2$, thus keeping 41 points ($D = -2$ to $D = 2$).

Next, we fitted these points using the following function at $\tau = 1$:

$$S_{\text{fit}} = \alpha + \beta_{\pm} \, |D - D_0|,$$

where α, β_- and β_+ are the fitting parameters, while D_0 was fixed using visual inspection.

At $\tau > 1$, we used the following function, including the dispersion σ:

$$S_{\text{fit}} = \alpha + \beta_{\pm}\,|D - D_0| + \beta'\,\sigma,$$

where α, β_-, β_+ and β' are the fitting parameters, while D_0 was fixed using visual inspection.

For the fitting, we used the entire interval shown in Figure 8, namely [-2.5, 2.5] (bins are of size one, so the dot at $D = 2$, for instance, shows the average of S between 1.5 and 2.5). In a few cases only did we slightly restrict the fitting interval in order to obtain better results:

- Random treatment, $\tau = 1$: [-1.65, 2.5]
- Median treatment, $\tau = 7$ and Shifted-Median treatment, $\tau = 3$ and 5: [-1.5, 2.5]
- Shifted-Median treatment, $\tau = 9$: [-1.2, 1]

For the fitting, we wrote a program to perform the minimization of least squares. Let $Q = \sum_i \left(S_i - S_{i\text{fit}}\right)^2 = \sum_i \left(S_i - \alpha - \beta_{\pm}\,|D_i - D_0| - \beta'\,\sigma_i\right)^2$ be the sum, over all the data in the chosen interval (indexed by i), of squared distances between S and S_{fit}. We then equated to 0 the partial derivatives of Q with respect to α, β_-, β_+ and β' (when $\tau > 1$) to obtain the values of these parameters.

Acknowledgments: We are grateful to Felix Lappe for programming the experiment, and thank Alan Tump, Lucienne Eweleit, Klaus Reinhold and Oliver Krüger for their support in the organization of our study. We are grateful to the ARC research group for their constructive feedback. This work was partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2002/1 "Science of Intelligence" – project number 390523135.

Author contributions: B. J. and R. K. designed research; B. J. and R. K. performed research; B. J. and C. S. analyzed the data and designed the model; the three authors wrote the article.

Corresponding author: Correspondence to Bertrand Jayles – jayles@mpib-berlin.mpg.de

Competing interests: The authors declare no competing interests.

Data accessibility: The data supporting the findings of this study are available at figshare: https://doi.org/10.6084/m9.figshare.12472034.v1

References

[1] Ehrlinger J, Readinger WO, Kim B (2016) Decision-making and cognitive biases. Reference Module in Neuroscience and Biobehavioral Psychology, Encyclopedia of Mental Health (Second Edition), pp. 5–12

[2] Mahmoodi A, et al. (2015) Equality bias impairs collective decision-making across cultures. *Proceedings of the National Academy of Science of the USA* 112(12):3835–3840.

[3] Cha M, et al. (2010) Measuring user influence in twitter: The million follower fallacy. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* pp. 10–17.

[4] Jansen BJ, et al. (2009) Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology* 60(11):2169–2188.

[5] Gonçalves B, Perra N (2015) *Social phenomena: From data analysis to models.* (Heidelberg, New-York: Springer International Publishing AG.).

[6] Cheng M, Jin X (2019) What do airbnb users care about? an analysis of online review comments. *International Journal of Hospitality Management* 76(A):58–70.

[7] Schafer JB, Konstan JA, Riedl J (2001) E-commerce recommendation applications. *Data Mining and Knowledge Discovery* 5(1–2):115–153.

[8] O'Connor P (2008) User-generated content and travel: A case study on tripadvisor.com. *Information and Communication Technologies in Tourism 2008* pp. 47–58.

[9] Fowler JH, Christakis NA (2010) Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences of the United States of America* 107(12):5334–5338.

[10] Salminen J (2012) Collective intelligence in humans: A literature review. *arXiv:1204.3401*.

[11] Bonabeau E (2009) Decisions 2.0: the power of collective intelligence. *MIT Sloan Management Review, Cambridge* 50(2):45–52.

[12] Woolley AW, Aggarwal I, Malone TW (2015) Collective intelligence and group performance. *Current Directions in Psychological Science* 24(6):420–424.

[13] Kurvers RH, et al. (2016) Boosting medical diagnostics by pooling independent judgments. *Proceedings of The National Academy of Sciences of the United States of America* 113(31):8777–8782.

[14] Brewer MB (1979) In-group bias in the minimal intergroup situation: a cognitive-motivational analysis *Psychological Bulletin* 86(2):307–324.

[15] Garrett RK (2009) Echo chambers online: Politically motivated selective exposure among internet news users. *Journal of Computer-Mediated Communication* 14(2):265–285.

[16] Flaxman S, Goel S, Rao JM (2016) Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly* 80(Special issue):298–320.

[17] Yaniv I (2004) Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes* 93(1):1–13.

[18] Soll JB, Larrick RP (2009) Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(3):780–805.

[19] Lorenz J, 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.

[20] Jayles B, et al. (2017) How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences of the United States of America* 114(47):12620–12625.

[21] Jayles B, et al. (2020) The impact of incorrect social information on collective wisdom in human groups. *Journal of the Royal Society Interface* 17(170):20200496.

[22] Indow T, Ida M (1977) Scaling of dot numerosity. *Perception & Psychophysics* 22(3):265–276.

[23] Krueger, L. E. Single judgements of numerosity. *Perception & Psychophysics* **31(2),** 175–182 (1982).

[24] Izard, V. & Dehaene, S. Calibrating the mental number line. *Cognition* **106(3),** 1221–1247 (2008).

[25] Crollen V, Castronovo J, Seron X (2011) Under- and Over-Estimation: A Bi-Directional Mapping Process Between Symbolic and Non-Symbolic Representations of Number? *Experimental Psychology* 58(1):39–49.

[26] Kao AB, et al. (2018) Counteracting estimation bias and social influence to improve the wisdom of crowds. *Journal of the Royal Society Interface* 15(141).

[27] Lichtenstein S, et al. (1978) Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory* 4(6):551–578.

[28] Hertwig R, Pachur T, Kurzenhäuser S (2005) Judgments of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(4):621–642.

[29] Scheibehenne B (2018) The psychophysics of number integration: Evidence from the lab and from the field. *Decision. Advance online publication.*

[30] Mavrodiev P, Tessone CJ, Schweitzer F (2013) Quantifying the effects of social influence. *Scientific Reports* 3:1360.

[31] Kerckhove CV, et al. (2016) Modelling influence and opinion evolution in online collective behaviour. *PLoS ONE* 11(6):e0157685.

[32] Becker J, Brackbill D, Centola D (2016) Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences of the United States of America* 114(26), E5070–E5076.

[33] Luo Y, Iyengar G, Venkatasubramanian V (2018) Social influence makes self-interested crowds smarter: an optimal control perspective. *IEEE Transactions on Computational Social Systems* 5(1):200–209.

[34] Faria JJ, Dyer JR, Tosh CR, Krause J (2010) Leadership and social information use in human crowds. *Animal Behaviour* 79(4).

[35] King AJ, et al. (2012) Is the true 'wisdom of the crowd' to copy successful individuals? *Biology Letters* 8(2):197–200.

[36] Madirolas G, de Polavieja GG (2015) Improving collective estimations using resistance to social influence. *PLOS Computational Biology* 11(11):e1004594.

[37] Moussaïd M, et al. (2013) Social influence and the collective dynamics of opinion formation. *PLoS ONE* 8(11):e78433.

[38] Chacoma A, Zanette DH (2015) Opinion formation by social influence: From experiments to modeling. *PLoS One* 10(10):e0140406.

[39] Yaniv I, Milyavsky M (2007) Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision Processes* 103:104–120.

[40] Rand DG, Arbesman S, Christakis NA (2011) Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Science of the United States of America* 108(48):19193–19198.

[41] Analytis PP, Barkoczi D, Herzog SM (2018) Social learning strategies for matters of taste. *Nature Human Behavior* 2:415–424.

[42] Galton F (1907) Vox populi. *Nature* 75:450–451.

[43] Surowiecki J (2005) *The wisdom of crowds.* (Anchor Books, New York, NY).

[44] Herzog SM, Litvinova A, Yahosseini KS, Tump AN, Kurvers RHJM (2019) *The ecological rationality of the wisdom of crowds.* (In R. Hertwig, T. J. Pleskac, T. Pachur, & The Center for Adaptive Rationality, Taming uncertainty (pp. 245–262). Cambridge, MA: MIT Press.).

[45] Dehaene S, et al. (2008) Log or linear? distinct intuitions of the number scale in western and amazonian indigene cultures. *Science* 320(5880):1217–1220.

[46] Ioannou CC, Madirolas G, Brammer FS, Rapley HA, de Polavieja GG (2018) Adolescents show collective intelligence which can be driven by a geometric mean rule of thumb. *PLoS ONE* 13:e0204462.

[47] Lobo MS, Yao D (2010) Human judgment is heavy tailed: Empirical evidence and implications for the aggregation of estimates and forecasts. *INSEAD working paper series.*

[48] Jayles B, Sire C, Kurvers HJMR (2020) Impact of sharing full versus averaged social information on social influence and estimation accuracy. https://doi.org/10.31234/osf.io/4n8bh.

[49] Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185(4157):1124–1131.

[50] Simon HA (1982) *Models of Bounded Rationality.* (MIT Press).

[51] Gigerenzer G, Gaissmaier W (2011) Heuristic decision making. *Annual Review Psychology* 62:451–482.

[52] Harries C, Yaniv I, Harvey N (2004) Combining advice: The weight of a dissenting opinion in the consensus. *Journal of Behavioral Decision Making* 17:333–348.

[53] Molleman L, Tump AN, et al. (2020) Strategies for integrating disparate social information. https://doi.org/10.31234/osf.io/wgzna.

[54] Kurvers RHJM, et al. (2019) How to detect high-performing individuals and groups: Decision similarity predicts accuracy. *Science Advances* 5(11):eaaw9011

[55] Lim J, Lee S-H (2020) Utility and use of accuracy cues in social learning of crowd preferences. *https://doi.org/10.31234/osf.io/jrynu*

[56] Yaniv I, Kleinberger E (2000) Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes* 83(2):260–281.

[57] Tump AN, Wolf M, Krause J, Kurvers RHJM (2018) Individuals fail to reap the collective benefits of diversity because of over-reliance on personal information. *Journal of the Royal Society Interface* 15:20180155.

[58] Davis-Strober CP, Budescu DV, Dana J, Broomell SB (2014) When Is a Crowd Wise? *Decision* 1(2):79–101.

[59] Thaler R, Sunstein C (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness.*

[60] Banerjee AV (1992) A simple model of herd behavior. *The Quarterly Journal of Economics, Oxford University Press* 107(3):797-817.

[61] Lorenz J (2007) Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C* 18(12):1819–1838.

[62] Sunstein, CR (2018) *#Republic: Divided democracy in the age of social media.* (Princeton, NJ: Princeton University Press).

[63] Nickerson RS (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2(2):175–220.

[64] Dunning D (2012) *Self-Insight: Roadblocks and Detours on the Path to Knowing Thyself.*

Supporting information captions

Table S1: Distribution of cases when the social information provided to an individual was higher ($D > 0$) or lower ($D < 0$) than their personal estimate, for each combination of treatment and number of estimates received τ. As expected, the proportions are roughly equal in the Random treatment, while the social information is more often lower than the personal estimate in the Median treatment, and more often higher in the Shifted-Median treatment.

Figure S1: Experimental procedure for an example question. The left panel shows the first screen in which subjects had to provide their personal estimate. The question was asked on the first line, and the answer could be typed on the second line, using a keyboard that appeared when clicking on "Ihre Antwort" ("Your answer" in German). Subjects submitted their estimates by pushing the "OK" button. A timer was displayed in the top right corner of the screen to remind subjects to answer within 30 seconds. The right panel shows the second screen in which subjects could revise their estimate after observing answers from other group members (in this example 5 answers). As a reminder, the original question, as well as the subject's personal estimate were shown. Subjects provided their second estimate in the same way as the first one and the countdown timer was again set on 30 seconds.

Figure S2: Median of the logarithm of estimates against the logarithm of the correct answer for the 36 questions asked in our experiment (one dot per question). (a) Green colors represent general knowledge questions, and orange numerosity questions, i.e., estimating the number of objects in an image. The slopes of the linear regression lines are 0.9 and 0.93 respectively, suggesting a similar relationship for both classes; (b) Green colors represent the 18 questions with the largest true values, and orange the 18 questions with the smallest true values. The slopes of the linear regression lines are 0.91 and 0.86 respectively, suggesting that the degree of underestimation is robust across different magnitudes.

Figure S3: Probability density function (PDF) of personal estimates X_{p} for all conditions combined. Dots are the data and the line model simulations.

Figure S4: Average sensitivity to social influence $\langle S \rangle$ against (a) the number of shared estimates τ and (b) the average dispersion $\langle \sigma \rangle$ of the social estimates, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Filled dots are the data, while empty dots and solid lines are simulations of the model without the similarity effect. This model underestimates the inverse-U shape in panel a and the decrease of $\langle S \rangle$ with $\langle \sigma \rangle$ in panel b.

Figure S5: Average sensitivity to social influence $\langle S \rangle$ against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments, when the average social information M is higher than the personal estimate X_{p} ($D = M - X_{\mathrm{p}} > 0$; squares) and when it is lower ($D < 0$; triangles). Filled symbols represent the data, while solid lines and empty symbols are simulations of the model without the similarity effect. This model is unable to reproduce the empirical results and predicts flatter curves instead.

Figure S6: Collective and individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Values closer to 0 indicate higher accuracy. Solid and dashed lines are simulations of the model without the similarity effect, before and after social information sharing, respectively.

Figure S7: Average sensitivity to social influence $\langle S \rangle$ against (a) the number of shared estimates τ and (b) the average dispersion $\langle \sigma \rangle$ of the social estimates, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Filled dots are the data, while empty dots and solid lines are simulations of the model without the asymmetry effect.

Figure S8: Average sensitivity to social influence $\langle S \rangle$ against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments, when the average social information M is higher than the personal estimate X_{p} ($D = M - X_{\mathrm{p}} > 0$; squares) and when it is lower ($D < 0$; triangles). Filled symbols represent the data, while solid lines and empty symbols are simulations of the model without the asymmetry effect. This model is unable to reproduce the empirical discrepancy between $\langle S \rangle$ when $D < 0$ and when $D > 0$.

Figure S9: Collective and individual accuracy, against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Values closer to 0 indicate higher accuracy. Solid and dashed lines are simulations of the model without the asymmetry effect, before and after social information sharing, respectively. This model is unable to reproduce the improvement in collective accuracy in the Random and Median treatments.

Figure S10: Collective accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. The population was separated into subjects' answers where the average social information received M was lower than their personal estimate X_{p} ($D = M - X_{\mathrm{p}} < 0$) and subjects' answers where the average social information received was higher than their personal estimate ($D > 0$). Solid and dashed lines are model simulations before and after social information sharing, respectively.

Figure S11: Collective accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. The population was separated, in each condition, into subjects whose sensitivity to social influence S was lower than the median value of S in that condition, and subjects whose sensitivity to social influence S was higher than the median value of S in that condition. Solid and dashed lines are model simulations before and after social information sharing, respectively.

Figure S12: Collective and individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Shifted-Median treatment. Squares denote the results of the recalibration of personal estimates (see Discussion for details). Collective accuracy improves similarly with this recalibration method as in the Shifted-Median treatment. However, individual accuracy decays with the recalibration method, while it improves substantially in the Shifted-Median treatment.

Figure S13: Collective accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles or squares) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. Top/bottom panels indicate the results of the half of our questions with lowest/highest true values. Before social information sharing, collective accuracy is higher (i.e., closer to 0) for moderately large values than for very large values, but improves more in the latter than in the former.

Figure S14: Individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles or squares) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. Top/bottom panels indicate the results of the half of our questions with lowest/highest true values. Before social information sharing, individual accuracy is higher (i.e., closer to 0) for moderately large values than for very large values, but improves more in the latter than in the former.

SI Appendix. Belonging to: Debiasing the crowd: how to select social information to improve judgment accuracy?

Bertrand Jayles[1,*], Clément Sire[2], Ralf H.J.M Kurvers[1]

[1]Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

[2]Laboratoire de Physique Théorique, Centre National de la Recherche Scientifique (CNRS), Université de Toulouse – Paul Sabatier (UPS), Toulouse, France

1 Experimental design

Participants were 216 students, distributed over 18 groups of 12 individuals, from the Bielefeld University, taking an Introductory Biology course (16-18 April 2018). Prior to participation, all participants signed an informed consent form and the experiment was approved by the Institutional Review Board of the Max Planck Institute for Human Development (A 2018/11).

Each of the 12 subjects—in each of the 18 groups—was confronted with 36 estimation questions (see the list in section 2) on a tactile tablet (Lenovo TAB 2 A10-30). Each question was asked twice: first, subjects provided their personal estimate $E_{\rm p}$. Next, they received as social information the estimate(s) of one or more group members (*i.e.* other subjects in the same room at the same time), and were asked to provide a second estimate $E_{\rm s}$. As a

*Corresponding author – jayles@mpib-berlin.mpg.de

reminder, their personal estimate was also shown during the second answering of a question. Supplementary Fig. S1 illustrates how social information was displayed on the tablets: on the right side of the screen was a blue panel showing all pieces of social information, sorted in increasing order. All tablets were controlled by a central server, and participants could only proceed to the next question once all individuals provided their estimate. A 30 seconds count down timer was shown on the screen to motivate subjects to answer within this time window, although they were allowed to take more time.

When providing social information, we varied (i) the number of estimates selected (1, 3, 5, 7, 9, or 11), and (ii) the selection procedure (Random, Median, and Shifted-Median). In the Random treatment, subjects received random estimates from their 11 group members. In the Median treatment, we presented the estimates of which logarithm[1] was closest to the median of the logarithms of the 12 personal estimates. In the Shifted-Median treatment, subjects were provided the estimates of which logarithm was closest to a shifted (overestimated) value of the median of the logarithms of the 12 personal estimates (see Main Text, Material and Methods). The participants were not aware of these different treatments.

Importantly, in all treatments, subject did not receive their own estimate as social information. In total, there were 6 different numbers of estimates selected $\times$ 3 treatments = 18 unique conditions. In every session, the 36 questions were randomly assigned to six blocks of six questions. Across groups, the order of the blocks, and the questions within a block, were randomized. A block always contained each number of estimates to be shown (1, 3, 5, 7, 9 and 11) once and was assigned one of the three treatments (Random, Median or Shifted-Median). Each group experienced two blocks of each treatment, and thus each of the 18 unique conditions twice. The randomization was constrained in such a way that at the end of the whole experiment, *all* 36 questions were asked once in *all* 18 different conditions, resulting in 36 estimates (1 per question) $\times$ 12 subjects = 432 estimates ($\times$2: before and after receiving social information) per condition.

[1] The logarithmic scale is consistent with the logarithmic perception of numbers [43].

Students received course credits for participation. Additionally, we incentivized them based on their performance P, defined as:

$$P_i = \frac{1}{2}\left(\text{Median}_q \left|\log\left(\frac{E_{\mathrm{p}_{i,q}}}{T_q}\right)\right| + \text{Median}_q \left|\log\left(\frac{E_{\mathrm{s}_{i,q}}}{T_q}\right)\right|\right),$$

where i and q index individuals and questions, E_p and E_s are estimates before ("personal") and after ("second") receiving social information, and T is the correct answer to the question at hand. This performance criterion measures, for each individual, the median distance (in terms of orders of magnitude) of their estimates to the corresponding correct answers to all questions, averaged over the two estimates (before and after receiving social information). The payments were defined according to the distribution of performances measured in [20]:

- $P_i < 0.4$: 5€ ($\sim$ 20% of subjects)
- $0.4 \leq P_i < 0.5$: 4€ ($\sim$ 30% of subjects)
- $P_i \geq 0.5$: 3€ ($\sim$ 50% of subjects)

2 List of questions

Below is the list of questions used in the experiment and the corresponding true values T. In the original experiment, the questions were asked in German. Questions were a mix of general knowledge and numerosity, i.e., estimating the number of objects (e.g. marbles, matches, animals) in an image. Images were shown for 6 seconds. 18 questions were taken from a previous study [20], and 18 were new (shown in italic). Questions 21 and 32 were the same in [20], but were asked in different units, such that the true answer and corresponding estimates were substantially different. Therefore, we considered these as new.

1. What is the population of Tokyo and its agglomeration? $T = 38,000,000$
2. What is the population of Shanghai and its agglomeration? $T = 25,000,000$

3. What is the population of Seoul and its agglomeration? $T = 26,000,000$

4. What is the population of New-York City and its agglomeration? $T = 21,000,000$

5. What is the population of Madrid and its agglomeration? $T = 6,500,000$

6. What is the population of Melbourne and its agglomeration? $T = 4,500,000$

7. *How many ebooks were sold in Germany in 2016?* $T = 28,100,000$

8. How many books does the American library of Congress hold? $T = 16,000,000$

9. How many people died from cancer in the world in 2015? $T = 8,800,000$

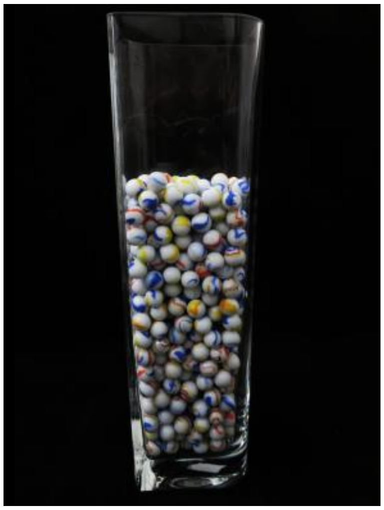
10. *How many smartphones were sold in Germany in 2017?* $T = 24,100,000$

11. *What was the total distance of the 2016 Tour de France (in kilometers)?* $T = 3,529$

12. *How many insured cars were stolen in Germany in 2016?* $T = 18,227$

13. Marbles 1: How many marbles do you think are in the jar in the following image? $T = 100$

14. Marbles 2: How many marbles do you think are in the jar in the following image? $T = 450$



15. Matches 1: How many matches do you think are present in the following image? $T = 240$

16. Matches 2: How many matches do you think are present in the following image? $T = 480$

17. *How many people identify as indigenous in Mexico?* $T = 6,000,000$

18. *How many cars were registered in Germany in 2016?* $T = 45,071,000$

19. *What is the diameter of the Sun (in kilometers)?* $T = 1,391,400$

20. What is the distance between Earth and the Moon (in kilometers)? $T = 384,400$

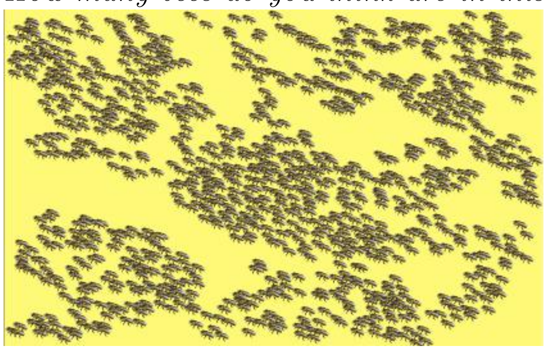
21. How many stars does the Milky way hold? $T = 235,000,000,000$

22. How many kilometers is one light-year (in billion kilometers)? $T = 9,460$

23. How much is the per-day income of Mark Zuckerberg (in dollars)? $T = 4,400,000$

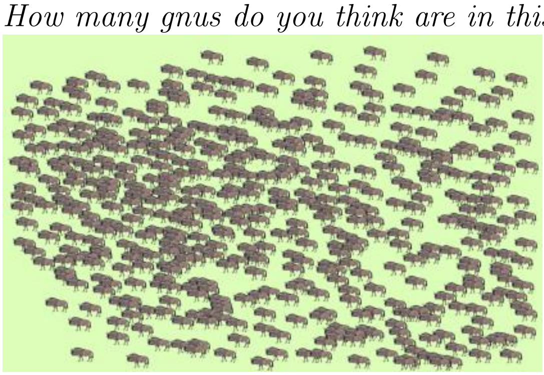
24. How many cells are there in the human body (in billion cells)? $T = 100,000$

25. How many bees do you think are in this picture? $T = 976$



26. What is the average annual salary of a player in the Bundesliga (in euros)? $T = 1,456,565$

27. How many gnus do you think are in this picture? $T = 483$



28. How many bikes do you think there are in Germany? $T = 62,000,000$

29. What is the distance from planet Mercury to the Sun (in kilometers)? $T = 58,000,000$

30. What is the total length of the metal threads used in the braided cables of the Golden Gate Bridge (in kilometers)? $T = 129,000$

31. What is the mass of the pyramid of Cheops (in tons)? $T = 5,000,000$

32. How much did the building of the Burj Khalifa tower in Dubai cost (in dollars)? $T = 1,500,000,000$

33. What is the average salary for players at Bayern Munich (in euros)? $T = 5,460,000$

34. What is the distance from Berlin to New-York (in kilometers)? $T = 6,188$

35. How many tourists were recorded in France in 2016? $T = 82,600,000$

36. How many UFO sightings have been reported to the National UFO Reporting Center in its history? $T = 90,000$

3 Supplementary table

Treatment	τ	$D > 0$ (%)	$D < 0$ (%)
Random	1	50	50
Random	3	52	48
Random	5	52	48
Random	7	49	51
Random	9	56	44
Random	11	53	47
Median	1	48	52
Median	3	44	56
Median	5	37	63
Median	7	32	68
Median	9	41	59
Median	11	52	48
Shifted-Median	1	68	32
Shifted-Median	3	65	35
Shifted-Median	5	56	44
Shifted-Median	7	45	55
Shifted-Median	9	42	58
Shifted-Median	11	49	51

Table S1: Distribution of cases when the social information provided to an individual was higher ($D > 0$) or lower ($D < 0$) than their personal estimate, for each combination of treatment and number of estimates received τ. As expected, the proportions are roughly equal in the Random treatment, while the social information is more often lower than the personal estimate in the Median treatment, and more often higher in the Shifted-Median treatment.

4 Supplementary figures

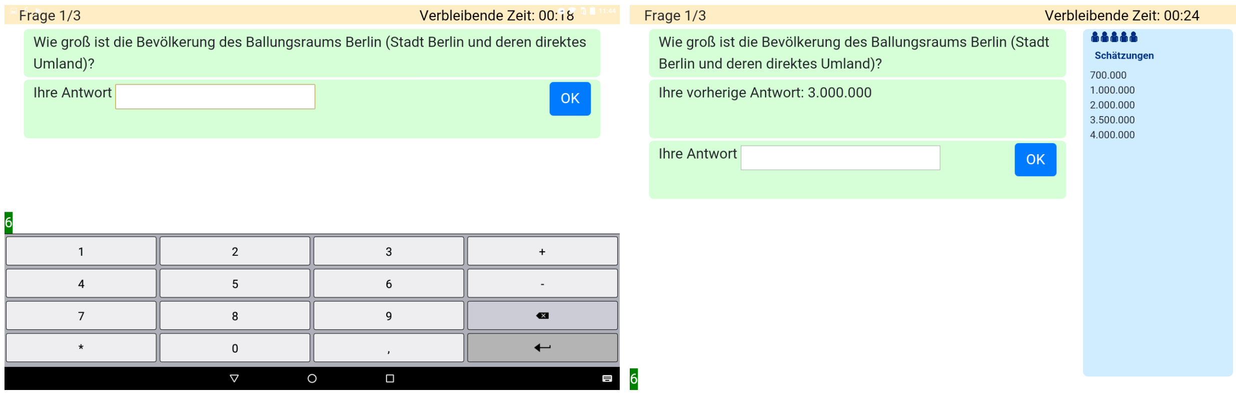


Fig. S1: Experimental procedure for an example question. The left panel shows the first screen in which subjects had to provide their personal estimate. The question was asked on the first line, and the answer could be typed on the second line, using a keyboard that appeared when clicking on “Ihre Antwort” (“Your answer” in German). Subjects submitted their estimates by pushing the “OK” button. A timer was displayed in the top right corner of the screen to remind subjects to answer within 30 seconds. The right panel shows the second screen in which subjects could revise their estimate after observing answers from other group members (in this example 5 answers). As a reminder, the original question, as well as the subject’s personal estimate were shown. Subjects provided their second estimate in the same way as the first one and the countdown timer was again set on 30 seconds.

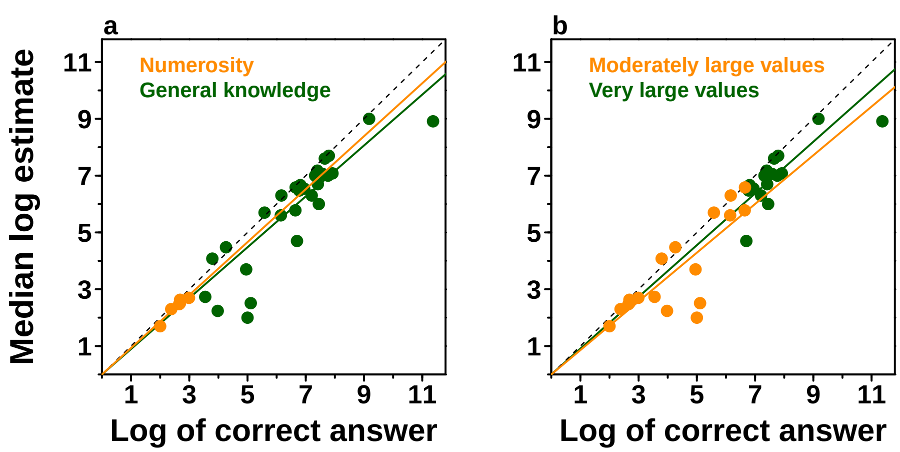


Fig. S2: Median of the logarithm of estimates against the logarithm of the correct answer for the 36 questions asked in our experiment (one dot per question). (a) Green colors represent general knowledge questions, and orange numerosity questions, i.e., estimating the number of objects in an image. The slopes of the linear regression lines are 0.9 and 0.93 respectively, suggesting a similar relationship for both classes; (b) Green colors represent the 18 questions with the largest true values, and orange the 18 questions with the smallest true values. The slopes of the linear regression lines are 0.91 and 0.86 respectively, suggesting that the degree of underestimation is robust across different magnitudes.

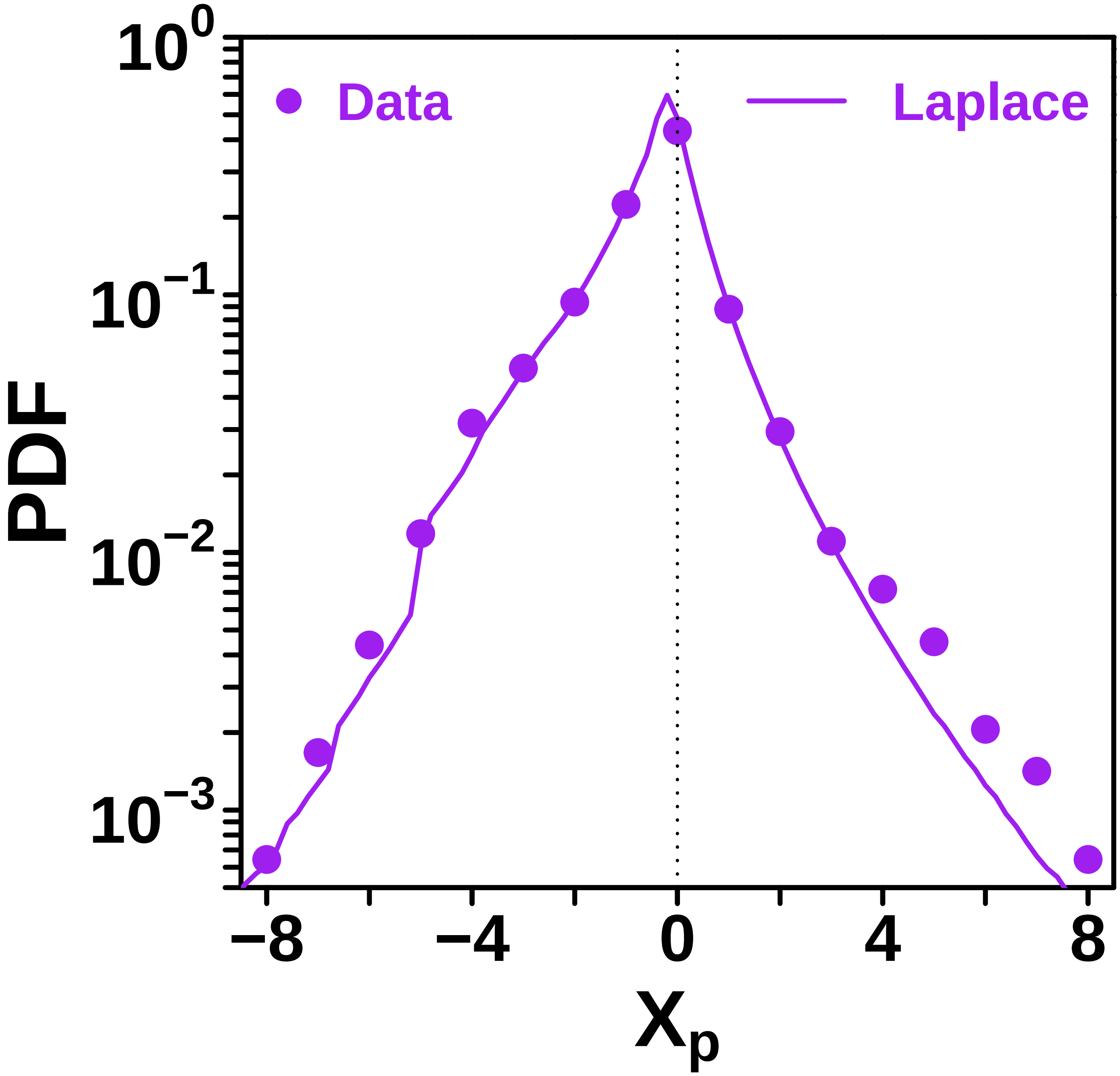


Fig. S3: Probability density function (PDF) of personal estimates X_{p} for all conditions combined. Dots are the data and the line model simulations.

Model without similarity effect: Figures S4 to S6

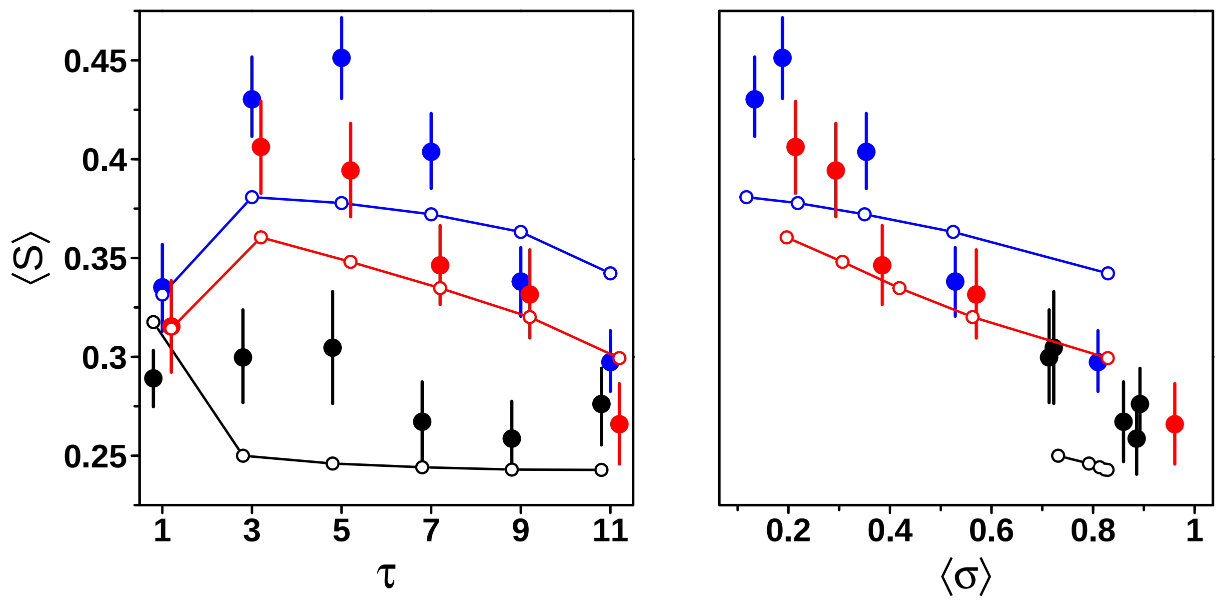


Fig. S4: Average sensitivity to social influence $\langle S \rangle$ against (a) the number of shared estimates τ and (b) the average dispersion $\langle \sigma \rangle$ of the social estimates, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Filled dots are the data, while empty dots and solid lines are simulations of the model without the similarity effect. This model underestimates the inverse-U shape in panel a and the decrease of $\langle S \rangle$ with $\langle \sigma \rangle$ in panel b.

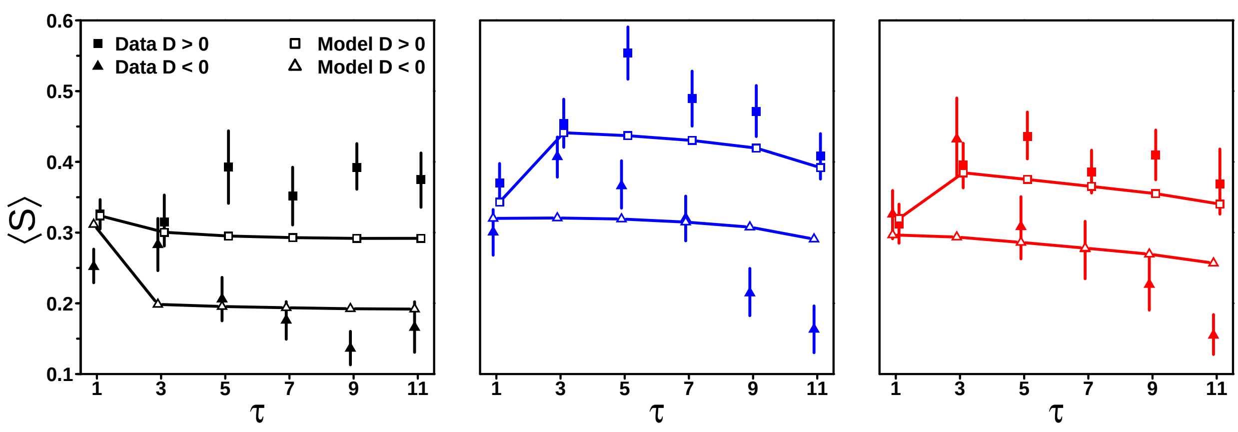


Fig. S5: Average sensitivity to social influence $\langle S \rangle$ against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments, when the average social information M is higher than the personal estimate X_p ($D = M - X_p > 0$; squares) and when it is lower ($D < 0$; triangles). Filled symbols represent the data, while solid lines and empty symbols are simulations of the model without the similarity effect. This model is unable to reproduce the empirical results and predicts flatter curves instead.

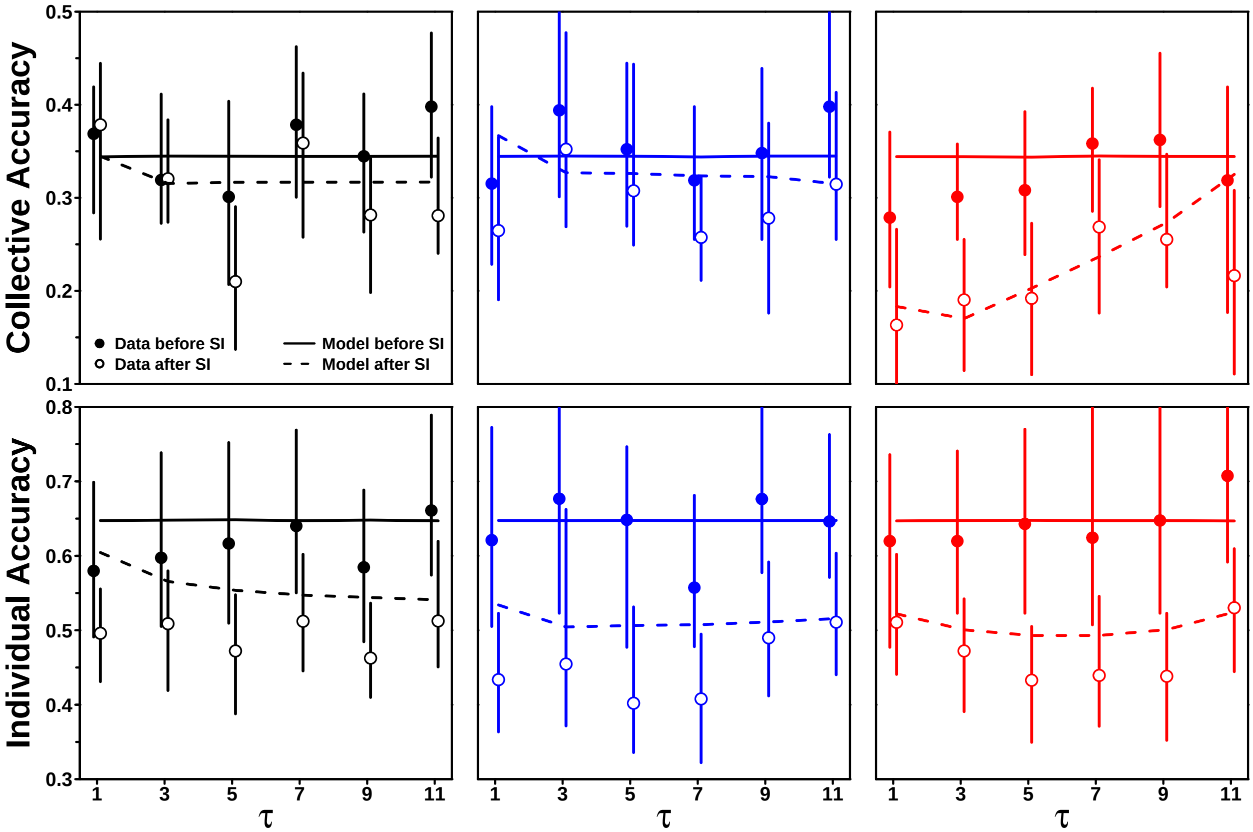


Fig. S6: Collective and individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Values closer to 0 indicate higher accuracy. Solid and dashed lines are simulations of the model without the similarity effect, before and after social information sharing, respectively.

Model without asymmetry effect: Figures S7 to S9

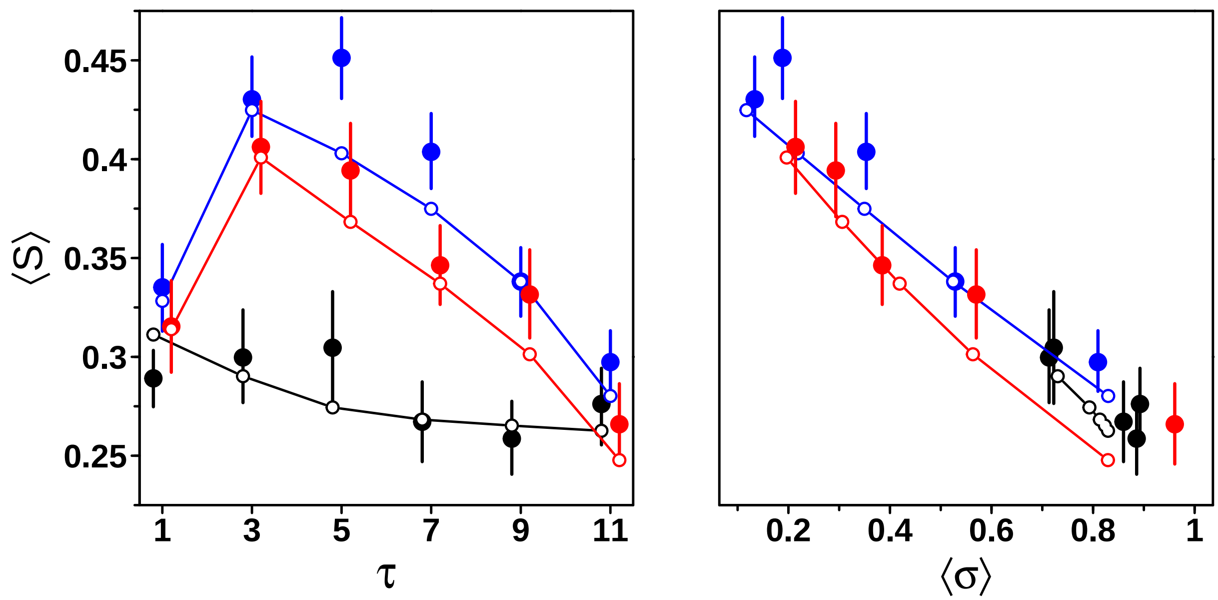


Fig. S7: Average sensitivity to social influence $\langle S \rangle$ against (a) the number of shared estimates τ and (b) the average dispersion $\langle \sigma \rangle$ of the social estimates, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Filled dots are the data, while empty dots and solid lines are simulations of the model without the asymmetry effect.

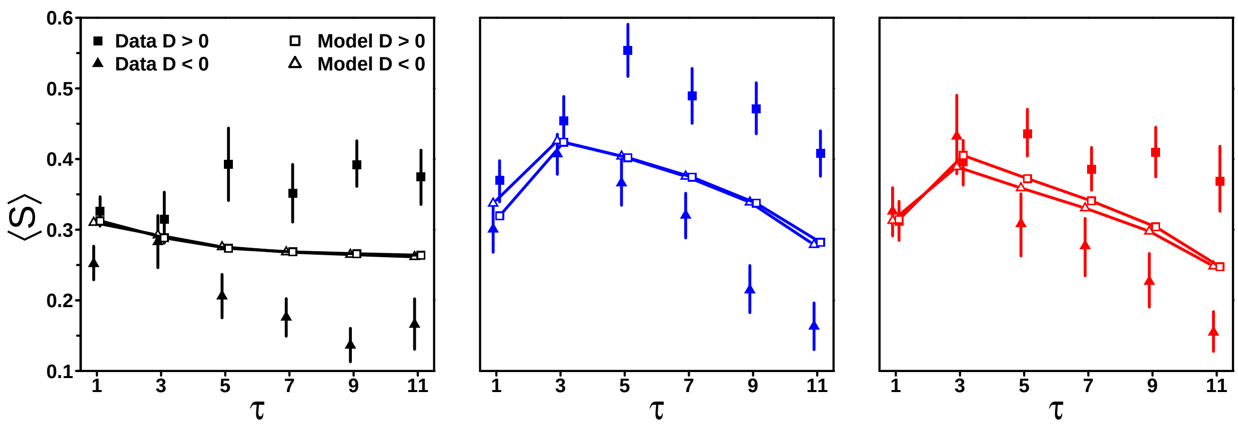


Fig. S8: Average sensitivity to social influence $\langle S \rangle$ against the number of shared estimates τ, in the Random (black), Median (blue), and Shifted-Median (red) treatments, when the average social information M is higher than the personal estimate X_{p} ($D = M - X_{\mathrm{p}} > 0$; squares) and when it is lower ($D < 0$; triangles). Filled symbols represent the data, while solid lines and empty symbols are simulations of the model without the asymmetry effect. This model is unable to reproduce the empirical discrepancy between $\langle S \rangle$ when $D < 0$ and when $D > 0$.

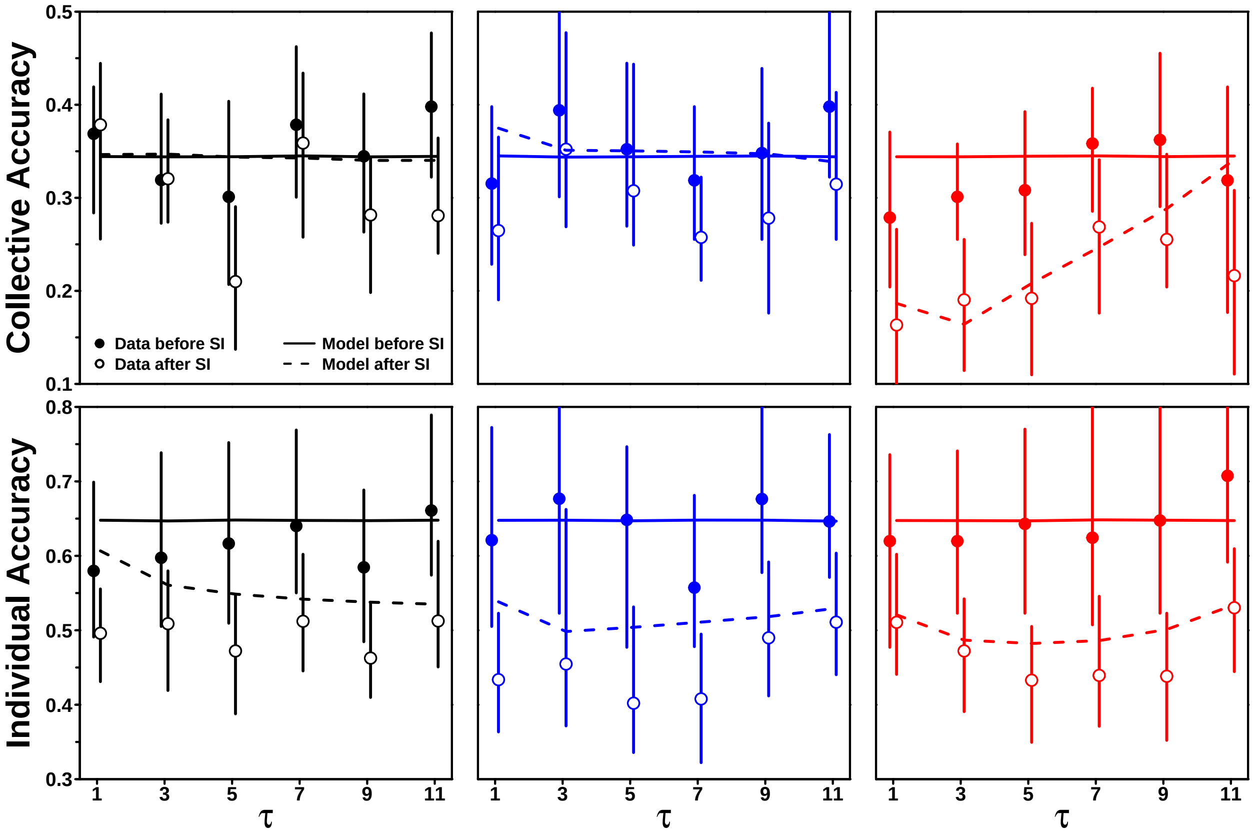


Fig. S9: Collective and individual accuracy, against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue), and Shifted-Median (red) treatments. Values closer to 0 indicate higher accuracy. Solid and dashed lines are simulations of the model without the asymmetry effect, before and after social information sharing, respectively. This model is unable to reproduce the improvement in collective accuracy in the Random and Median treatments.

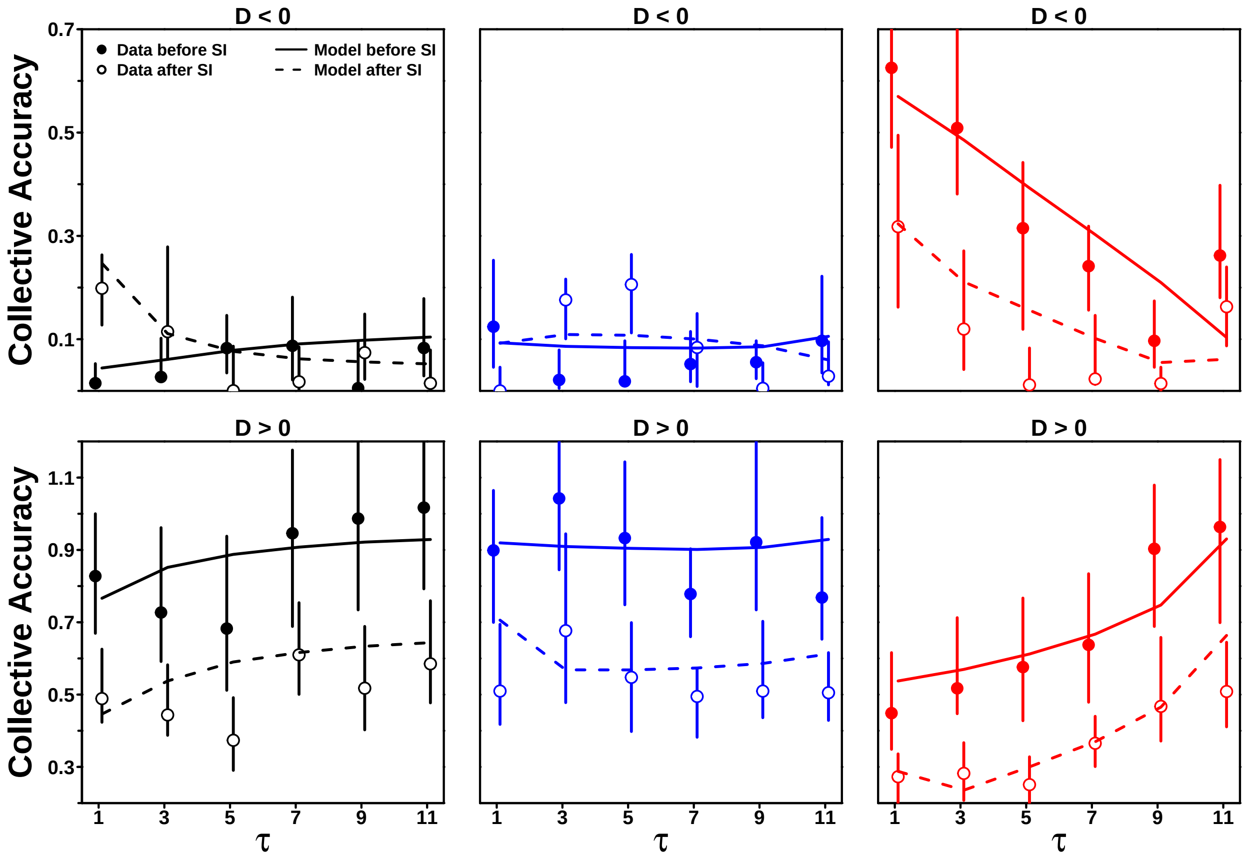


Fig. S10: Collective accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. The population was separated into subjects' answers where the average social information received M was lower than their personal estimate X_{p} ($D = M - X_{\mathrm{p}} < 0$) and subjects' answers where the average social information received was higher than their personal estimate ($D > 0$). Solid and dashed lines are model simulations before and after social information sharing, respectively.

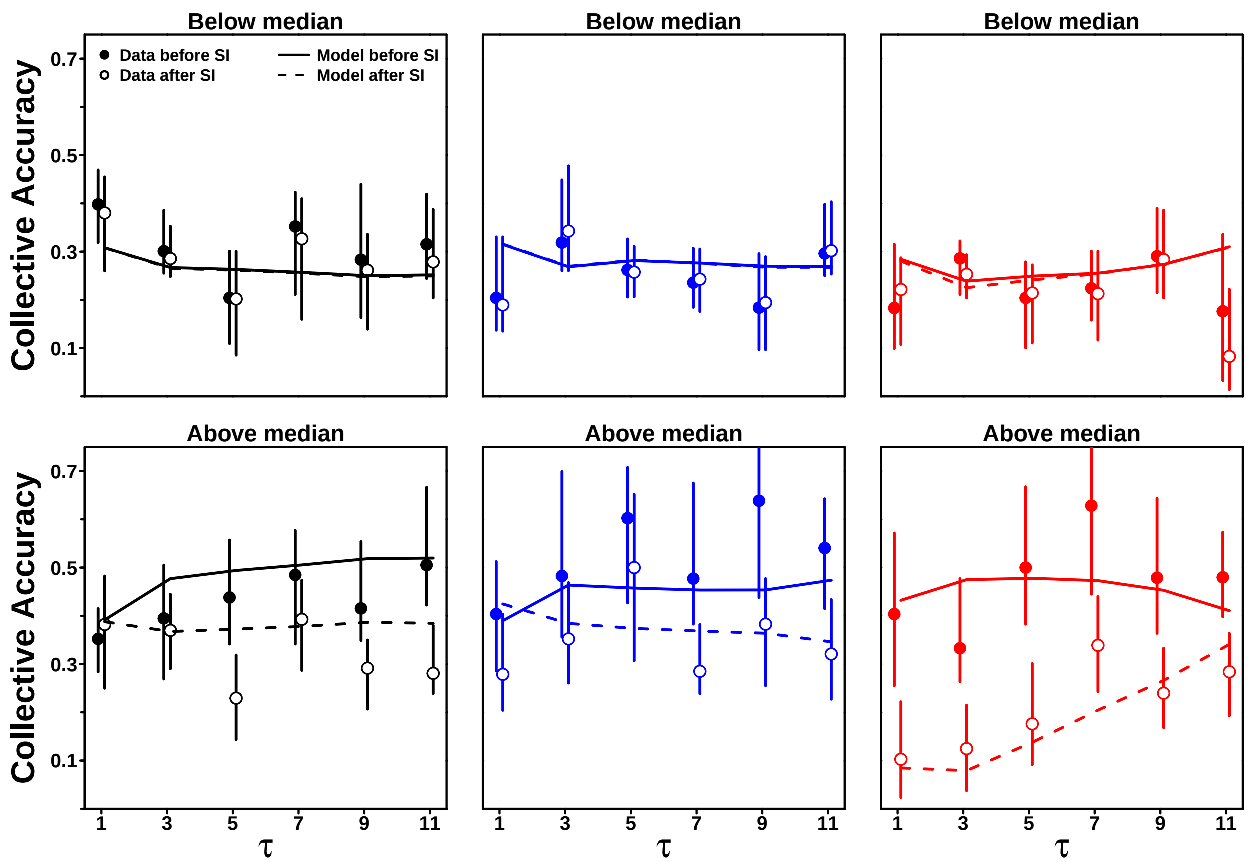


Fig. S11: Collective accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. The population was separated, in each condition, into subjects whose sensitivity to social influence S was lower than the median value of S in that condition, and subjects whose sensitivity to social influence S was higher than the median value of S in that condition. Solid and dashed lines are model simulations before and after social information sharing, respectively.

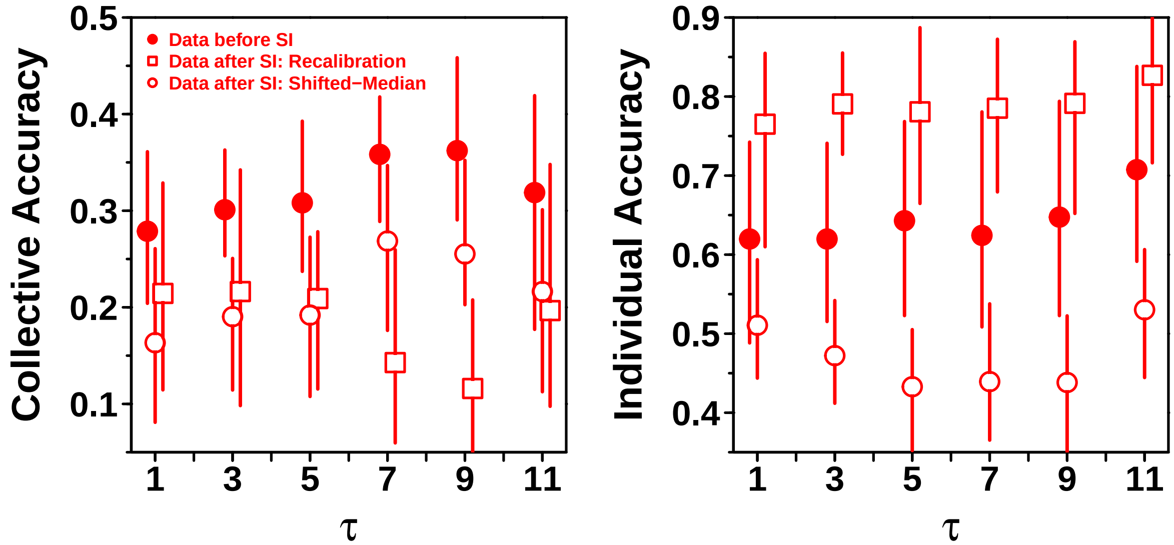


Fig. S12: Collective and individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles) social information sharing, in the Shifted-Median treatment. Squares denote the results of the recalibration of personal estimates (see Discussion for details). Collective accuracy improves similarly with this recalibration method as in the Shifted-Median treatment. However, individual accuracy decays with the recalibration method, while it improves substantially in the Shifted-Median treatment.

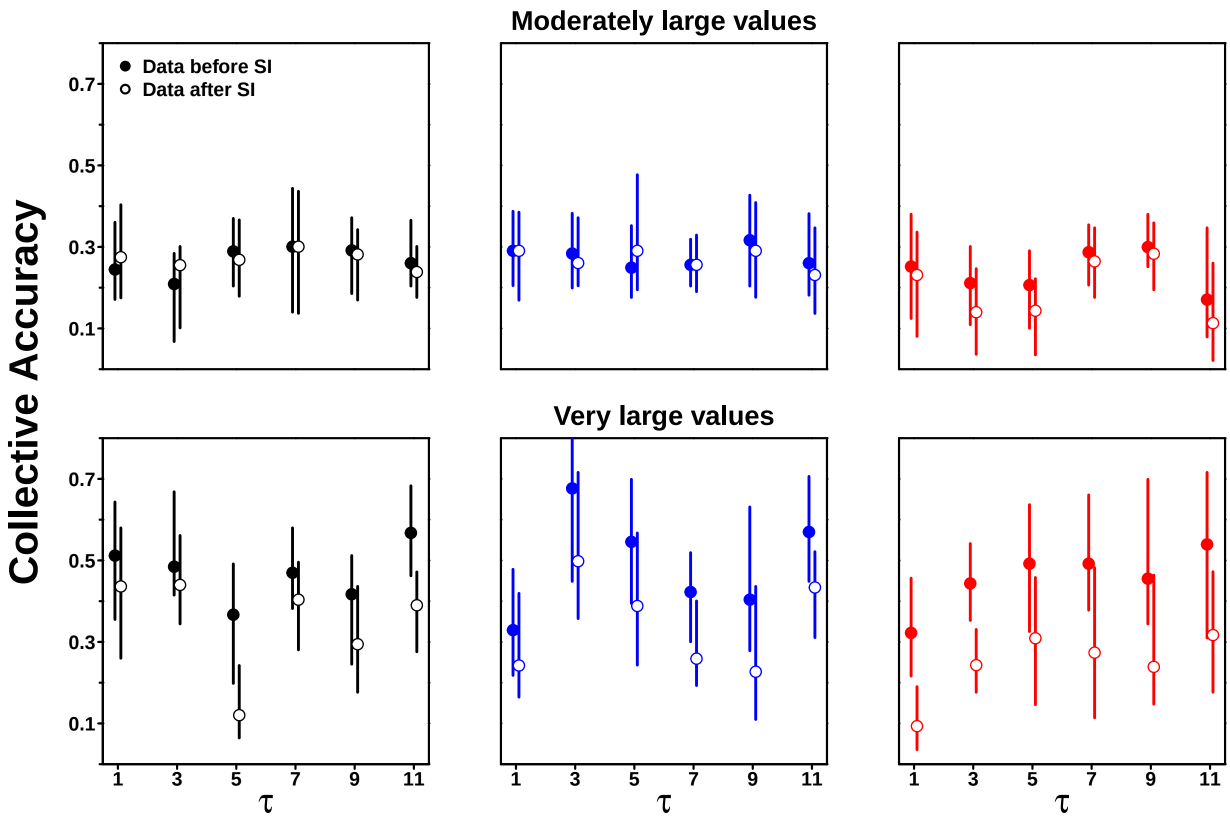


Fig. S13: Collective accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles or squares) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. Top/bottom panels indicate the results of the half of our questions with lowest/highest true values. Before social information sharing, collective accuracy is higher (i.e., closer to 0) for moderately large values than for very large values, but improves more in the latter than in the former.

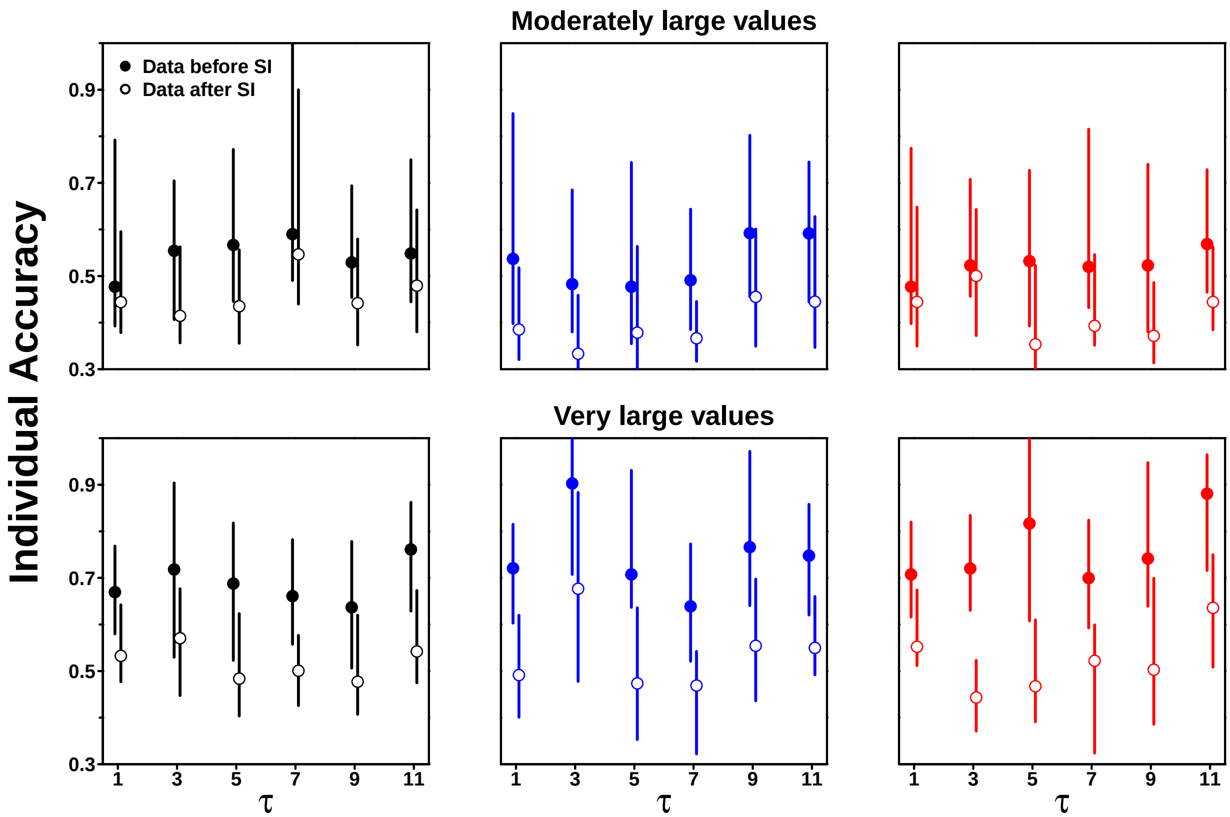


Fig. S14: Individual accuracy against the number of shared estimates τ, before (filled dots) and after (empty circles or squares) social information sharing, in the Random (black), Median (blue) and Shifted-Median (red) treatments. Top/bottom panels indicate the results of the half of our questions with lowest/highest true values. Before social information sharing, individual accuracy is higher (i.e., closer to 0) for moderately large values than for very large values, but improves more in the latter than in the former.