Supporting Information Appendix

Bertrand Jayles[a,b], **Hye-rin Kim**[c], **Ramón Escobedo**[b], **Stéphane Cezera**[d], **Adrien Blanchet**[e,f], **Tatsuya Kameda**[g], **Clément Sire**[a], **and Guy Theraulaz**[b,e]

[a]Laboratoire de Physique Théorique, Centre National de la Recherche Scientifique (CNRS), Université de Toulouse (Paul Sabatier), 31062 Toulouse, France; [b]Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), CNRS, Université de Toulouse (UPS), 31062 Toulouse, France; [c]Department of Behavioral Science, Hokkaido University, Kita 8-jo Nishi 11-chome Kita-ku, Sapporo, Hokkaido, Japan; [d]Toulouse School of Economics, INRA, Université de Toulouse (Capitole), 31000 Toulouse, France; [e]Institute for Advanced Study in Toulouse, 31015 Toulouse, France; [f]Toulouse School of Economics, Université Toulouse Capitole, 31000 Toulouse, France; [g]Department of Social Psychology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan 113-0033

Ethics statement

The aims and procedures of the experiments conformed to the Toulouse School of Economics Ethical Rules. All subjects provided written consent for their participation.

Informed consent was obtained from each participant in Japan using a consent form approved by the Institutional Review Board of the Center for Experimental Research in Social Sciences at Hokkaido University.

Materials and Methods

We conducted two series of experiments: the first one in Japan, at Hokkaido University, the second one in France, at the Toulouse School of Economics. The experimental protocol was similar in both cases and the results obtained in Japan (along with a simple analytical model presented in this appendix) were used to fix the parameters for which we conducted more in-depth investigation. An additional control (but partial) experiment was conducted in Japan with exactly the same questions as in France.

Experimental procedure in Japan. We recruited 180 participants, all of them were students from Hokkaido University. They were divided in 5 groups of 36 individuals, and one group per day performed the experiment. Everyday, each group was divided into 6 subgroups of 6 individuals, who participated to the experiment at the same time, what we call a session. Subjects could not participate to more than one session.

After entering the experimental room, the subjects were placed in cubicles that prevented interactions between them (see Figure S1B). Before starting the experiment, they were explained the rules, the payment conditions, the anonymity warranty, and were asked to shut down their mobile phones.

They were then asked to estimate 14 different quantities on a computer (see the list of questions below) and their answering time was limited (25 s). If they exceeded the time limit, a warning message would appear in red on the screen. The experimental program was written using the Z-Tree software[1].

Each question (quantity to estimate) involved two steps: first, subjects had to give their personal estimate E_p; then, after receiving the social information I, defined as the arithmetic mean of the τ previous participants' estimates (with $\tau = 1$, 3, or 5), they were asked to give a new estimate E. Subjects were not told the value of τ. They answered the questions sequentially (see Figure S1A), the sequence order being that in which they gave their personal estimate.

Virtual "experts" providing the true value T[2] for each question asked were inserted at random into the sequence of participants (see Figure S1A). For each sequence involving 36 participants, we controlled the number $n = 0$ or 18 and hence the percentage $\rho = \frac{n}{n+36} = 0\,\%$ or $33\,\%$ of virtual experts. The social information delivered to human participants, being the average of previous estimates, is hence strongly affected by these virtual experts. 18 experts were inserted in 6 questions out of 14 (see list of questions below), at locations chosen randomly prior to the experiment, different for all questions, but the same everyday (for all groups). 6 other questions had no experts, and the two last ones had a special treatment (see remarks below). The subjects knew nothing about these virtual experts.

When providing their estimates E_p and E, subjects had to report their confidence level in their answer, on a Likert scale ranging from 1 (very low) to 5 (very high).

At the end of each session, subjects received monetary payments in Japanese Yen (¥) according to their performance, defined as the average closeness to the true value over all questions. Three different intervals for monetary payments were used: ¥2000 for the best, ¥1500 for the two next ones, and ¥1000 for the three last ones.

Remarks:

- An initial condition was provided to the first subject of each group, as social information. The initial condition was purposely different from the true value, thus creating an artificial bias, to add some difficulty to the collective task. It was chosen from a Gaussian distribution centered around a certain value E_0, with standard deviation σ_0. We defined E_0 and σ_0 such that it would seem reasonable for the participants (see list of questions for the actual values used in the experiments). The second subject in the sequence was provided the average of the initial condition and the first estimate as social information. The third one was provided with the average of the initial condition and the two first answers, and so on until the τ^{th} participant, who was given the average of the τ previous estimates, including the initial condition. After her, all subjects were given the average of the τ previous estimates.

- For the 2 "special" questions, we were interested only in the impact of the distance between personal and group opinion (social information), on subjects' sensitivity to social influence. Therefore, no initial condition nor couple (ρ, τ) were associated to these questions. Instead, subjects were given, as social information, a percentage of their

[1] Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10:171–178.

[2] In the experiment performed in Japan, the value provided by the experts was actually narrowly distributed around the true value (with standard deviation $\sigma_e = 5\% T$).

DRAFT

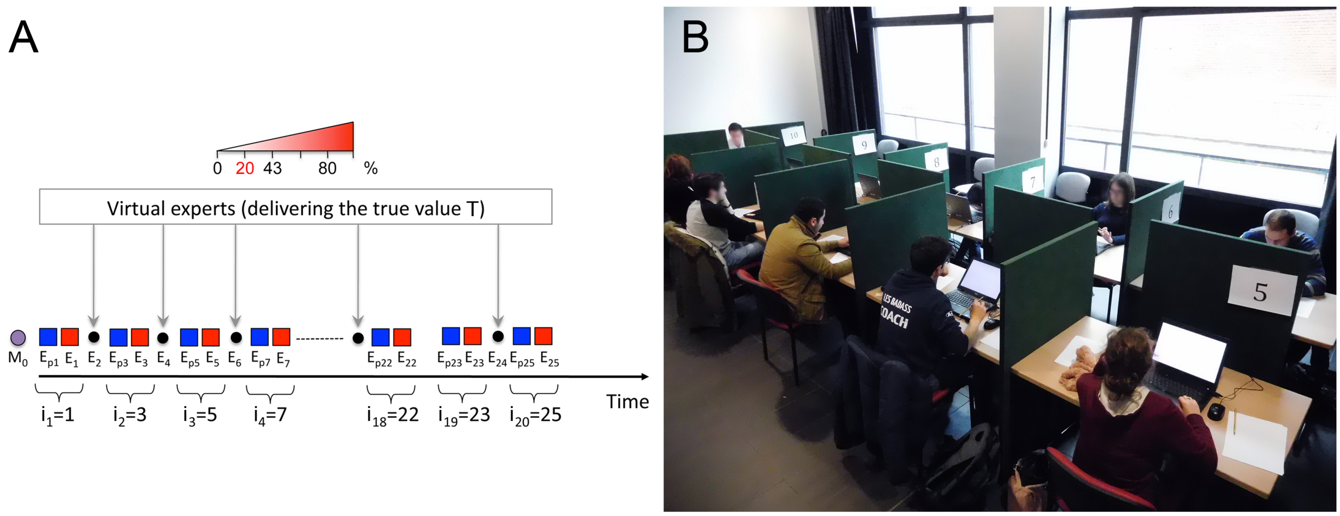

Fig. S1. The experimental protocol is summarized on the left panel A. For a given question, each individual i_k ($k = 1, \ldots, 20$) gives first her personal estimate E_p (in blue), then, after receiving social information (the arithmetic mean of the τ previous estimates in Japan, the geometric mean in France), she gives her new estimate (in red). Virtual experts (black dots) are added randomly into the sequence, with percentage ρ ($\rho = 20\%$ in this example, as highlighted in red), without subjects being aware of it. They impact the social information given to the subjects, and can thus be seen as an information input. On the right panel B is a picture of the experimental room.

own personal estimate (20% to 500%). For example, if a subject answered 1000 to a question, she was proposed as social information a number between 200 and 5000. In one session, the 6 subjects were associated 6 different percentages (among 12: 20%, 50%, 66%, 77%, 83%, 91%, 110%, 120%, 130%, 150%, 200%, 500%). Also, the subjects were given 2 different percentages for each special question: one below 100%, and one above. These "special questions" were purposely chosen such that the subjects were extremely unlikely to have any idea of the answers.

Experimental procedure in France. 186 subjects were recruited, the large majority of whom were students from the University of Toulouse. The experimental procedure was very similar to the one described above, but organized in a slightly different way. 20 sessions were organized over 5 days, in which 8 to 10 subjects had to answer 29 questions. The answering time was limited to 40 s per estimate, after which a blinking text urging the subject to speed up would appear on the screen. In each session, the 8 subjects were part of 8 different sequences (defined as in the experiment performed in Japan) associated with the 8 different couples (ρ, τ). The (1 or 2) possible "extra" subjects were given a different treatment (see remarks below).

As in the experiment performed in Japan, each question involved two steps: the subjects had to first give their personal estimate, and could then revise it after receiving social information. However, the social information in this experiment was defined as the *geometric mean* of the τ previous participants' estimates (with $\tau = 1$ or 3). This is because humans think in terms of orders of magnitude (see main text). The order of the questions was randomized up to the 24^{th} in all sessions (questions 25 to 29 were always in the same order).

Each question was sequentially answered by 20 human subjects. As in the experiment performed in Japan, virtual experts were inserted into the sequence (Figure S1A), with percentages $\rho = 0\,\%$; $20\,\%$, $43\,\%$ and $80\,\%$ (corresponding respectively to $n = 0$, 5, 15 and 80 experts). The 8 conditions (ρ, τ) were randomly associated to the 8 participants for each question. It is important to stress that in this experiment, and contrary to the one performed in Japan, all questions were asked in all conditions.

As in Japan, subjects had to report their confidence in their estimates, both before and after social influence, on the same Likert scale as previously described.

In addition, a third step was added for each question, for the experiment in France: after giving their second estimate, subjects were asked to evaluate the social information they received, on a 4-points scale of pertinence:

1. Absurd or not pertinent
2. Little pertinent
3. Quite pertinent
4. Very pertinent

They also had to choose, among 8 options, the one that best explained their choice to keep or change their opinion, after they received social information:

1. I had no idea of the answer
2. I had little idea of the answer
3. I was pretty sure to know the answer
4. I remembered something that eluded me the first time
5. It seemed doubtful to me that the other participants knew the answer better than I
6. I made a mistake in my first answer
7. I did not understand the question the first time
8. Other (here, they could briefly explain the reason)

DRAFT

According to the way they changed their opinion in the first 24 questions, the subjects were divided into 3 subgroups, in questions 25 to 29: the 2 subjects (out of 8 in a session) minimizing $\langle|1 - S_q|\rangle_q$ (i.e. the ones who were on average closest to $S = 1$), where q labels the questions, were defined as "followers"; the 2 ones minimizing $\langle|S_q|\rangle_q$ (i.e. the ones who were on average closest to $S = 0$) were defined as "confident"; the 4 other ones were defined as "average". We expected to find behavioral consistencies among these three subgroups (see Fig. S6): would the categorization built from the first 24 questions still be consistent in the following 5 questions?

At the end of every session, subjects were paid according to their performance (defined similarly to Japan): the first one was paid 20€, the two next ones 15€ and the four last ones 10€.

Remarks:

- The initial condition was proportional to the true value, with a factor $\kappa = 2$, 5 or 50, depending on the questions (see the list of questions). We alternated the factor between κ and $\frac{1}{\kappa}$, to control for a possible asymmetry between larger and smaller values.
- We added restrictions to the answers subjects could possibly give. These restrictions were defined up to a factor of $\lambda = 2$, 3, 4 or 7 orders of magnitude smaller or larger than the true value (see list of questions). These values of λ were chosen such that answers beyond these bounds could be safely considered as totally absurd (or a typing error).
- As a safeguard against subjects not showing up at a given session, we recruited 10 subjects per session, instead of 8. Overall, we needed 160 subjects, such that 26 were "extra" subjects (we had 186 in total). They were not part of any sequence, nor associated to a condition (ρ, τ), but were not aware to be given any difference in treatment. They were provided the value κT or $\frac{T}{\kappa}$ (with probability 50 %; T is the true answer) as social information. We recorded all their answers for the statistics. They were paid 10€ whatever their performance.

Second experiment in Japan. After the two main experiments, a partial experiment was performed in Japan with the same questions used in the experiment performed in France, but where subjects – 115 were recruited – only had to provide their personal estimate E_p. This experiment aimed at controlling the response of subjects from different cultures to the same questions. We indeed found that the same questions lead to the same distribution of personal estimates (see Figure S3B).

Questions used in the experiment.

Questions in Japan (asked in Japanese).

1. How old was Gandhi when he died? ($\tau = 3$ and $\rho = 0\%$)
   - $T = 78$
   - $E_0 = 1.2T$
   - $\sigma_0 = 10\%$
2. How much people live in Bayonne (town in the south west of France)? (special question)
3. How old was Yasujiro Ozu when he died? ($\tau = 3$ and $\rho = 1/3$)
   - $T = 60$
   - $E_0 = 1.2T$
   - $\sigma_0 = 15\%$
4. Jar 1: How many balls do you think are in this jar? ($\tau = 1$ and $\rho = 1/3$)
   - $T = 450$
   - $E_0 = 1.5T$
   - $\sigma_0 = 30\%$
5. How much people live in Tokyo (in terms of 10,000)? ($\tau = 5$ and $\rho = 1/3$)
   - $T = 1,346$
   - $E_0 = 1.25T$
   - $\sigma_0 = 30\%$
6. How many cell phones were sold in Japan in 2014 (in terms of 10,000)? ($\tau = 1$ and $\rho = 0\%$)
   - $T = 4,000$
   - $E_0 = 0.75T$
   - $\sigma_0 = 30\%$
7. What is the median income per month in Japan, in yen (in terms of 10,000 yens) ? ($\tau = 5$ and $\rho = 0\%$)
   - $T = 36$
   - $E_0 = 0.8T$
   - $\sigma_0 = 20\%$
8. Jar 2: How many balls do you think are in this jar? ($\tau = 1$ and $\rho = 1/3$)
   - $T = 100$
   - $E_0 = 1.5T$
   - $\sigma_0 = 30\%$
9. How many matches do you think there are? ($\tau = 1$ and $\rho = 0\%$)
   - $T = 400$
   - $E_0 = 0.75T$
   - $\sigma_0 = 30\%$
10. What is the average distance between the Earth and the Moon in km (in terms of 1,000 km)? ($\tau = 5$ and $\rho = 1/3$)
    - $T = 384$
    - $E_0 = 1.25T$
    - $\sigma_0 = 30\%$
11. How many books does Umberto Eco's library hold? (special question)
12. What is the average life expectancy for men in Ethiopia? ($\tau = 3$ and $\rho = 0\%$)

DRAFT

- $T = 63$
- $E_0 = 1.25T$
- $\sigma_0 = 10\%$

13. How many books does the American Congress Library hold (in terms of 10,000)? ($\tau = 5$ and $\rho = 0\%$)
    - $T = 2,389$
    - $E_0 = 0.8T$
    - $\sigma_0 = 20\%$
14. What is the distance between Tokyo and Pyongyang? ($\tau = 3$ and $\rho = 1/3$)
    - $T = 1,287$
    - $E_0 = 1.25T$
    - $\sigma_0 = 10\%$

Questions in France (asked in French) and in the second experiment in Japan (asked in Japanese).

1. What is the population of Tokyo and its agglomeration?
    - $T = 38,000,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^3$
2. What is the population of Shanghai and its agglomeration?
    - $T = 25,000,000$
    - $E_0 = 5T$
    - $\lambda = 10^3$
3. What is the population of Seoul and its agglomeration?
    - $T = 26,000,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^3$
4. What is the population of New-York City and its agglomeration?
    - $T = 21,000,000$
    - $E_0 = 5T$
    - $\lambda = 10^3$
5. What is the population of Madrid and its agglomeration?
    - $T = 6,500,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^3$
6. What is the population of Melbourne and its agglomeration?
    - $T = 4,500,000$
    - $E_0 = 5T$
    - $\lambda = 10^3$
7. How many ebooks were sold in France in 2014?
    - $T = 5,000,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^4$
8. How many books does the American Congress library hold?
    - $T = 23,000,000$
    - $E_0 = 5T$
    - $\lambda = 10^4$
9. How many people die from cancer in the world every year?
    - $T = 15,000,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^4$
10. How many cell phones are sold in France every year?
    - $T = 22,000,000$
    - $E_0 = 5T$
    - $\lambda = 10^4$
11. How many kilometers does a professional cyclist bikes a year?
    - $T = 40,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^2$
12. How many cars are stolen in France every year?
    - $T = 110,000$
    - $E_0 = 5T$
    - $\lambda = 10^3$
13. Marbles 1: How many marbles can you see in this jar?
    - $T = 100$
    - $E_0 = \frac{T}{2}$
    - $\lambda = 10^2$
14. Marbles 2: How many marbles can you see in this jar?
    - $T = 450$
    - $E_0 = 2T$
    - $\lambda = 10^2$
15. Matches 1: How many matches can you see?
    - $T = 240$
    - $E_0 = \frac{T}{2}$
    - $\lambda = 10^2$
16. Matches 2: How many matches can you see?
    - $T = 480$
    - $E_0 = 2T$
    - $\lambda = 10^2$

DRAFT

17. Rope 1: In your opinion, how long is this rope (in cm)?
    - $T = 200$
    - $E_0 = \frac{T}{2}$
    - $\lambda = 10^2$
18. Rope 2: In your opinion, how long is this rope (in cm)?
    - $T = 700$
    - $E_0 = 2T$
    - $\lambda = 10^2$
19. What is the radius of the Sun (in km)?
    - $T = 700,000$
    - $E_0 = \frac{T}{50}$
    - $\lambda = 10^7$
20. What is the distance between Earth and the Moon (in km)?
    - $T = 385,000$
    - $E_0 = 50T$
    - $\lambda = 10^7$
21. How many stars does the Milky way hold (in millions of stars)?
    - $T = 235,000$
    - $E_0 = \frac{T}{50}$
    - $\lambda = 10^7$
22. How many billions kilometers is worth a light-year?
    - $T = 9,000$
    - $E_0 = 50T$
    - $\lambda = 10^7$
23. How many galaxies does the visible universe hold (in millions of galaxies)?
    - $T = 100,000$
    - $E_0 = \frac{T}{50}$
    - $\lambda = 10^7$
24. How many cells are there in the human body (in billions of cells)?
    - $T = 100,000$
    - $E_0 = 50T$
    - $\lambda = 10^7$
25. What is the population of Amsterdam and its agglomeration?
    - $T = 1,600,000$
    - $E_0 = 5T$
    - $\lambda = 10^3$
26. What is the annual gross salary of a professional league 1 soccer player in France (in euros)?
    - $T = 540,000$
    - $E_0 = \frac{T}{5}$
    - $\lambda = 10^3$
27. Matches 3: How many matches can you see?
    - $T = 700$
    - $E_0 = 2T$
    - $\lambda = 10^2$
28. What is the total mass of oceans on Earth (in billions of tons)?
    - $T = 1,400,000,000$
    - $E_0 = \frac{T}{50}$
    - $\lambda = 10^7$
29. What is the distance from planet Aramis to its Sun (in km)?
    - $T = 1,000,000$
    - $E_0 = 5T$
    - $\lambda = 10^7$

DRAFT

Minimal model. We consider here a simple model, where the sensitivity to social information of the subjects S is independent of the difference $D = X_p - M$ between their initial personal estimate X_p and the social information M (see Figure 2A for the actual observed dependence included in the full model). The interest of this model is that it can be solved analytically, shedding some light on the convergence of the collective estimation process. Considering the presence of experts, equation 2 (see main text) becomes:

$$\begin{aligned} X_i &= (1 - S_i)\, X_{p_i} + S_i\, M_i, \quad \text{with probability } 1-\rho, \\ X_i &= 0, \quad \text{with probability } \rho \text{ (“virtual experts”)}, \end{aligned} \tag{S1}$$

where X_i is individual i's log-transformed estimate, and the social information is $M_i = \frac{1}{\tau}\sum_{j=i-\tau}^{i-1} X_j$. The virtual experts give (the log-transform of) true value $V = 0$ and are very confident so that their sensitivity to social influence is $S = 0$.

The distribution of personal estimates X_p is assumed stable, with stability parameter $a \in [1, 2]$, and center (equal to the mean or the median for symmetric distributions considered here) m_p and width σ_p. $a = 1$ and $a = 2$ correspond respectively to the Cauchy and Gaussian distributions, while intermediate values of a correspond to Lévy distributions. Consequently, the distributions of group's advice M and of estimates after social influence X are also stable (by definition of stability) with the same distribution but characterized by different center and width. We define m_i as the center and σ_i as the width of the distribution of estimates after social influence (excluding the virtual experts). From equation S1, for the i-th iteration/individual, and after averaging over configurations (*i.e.* an infinite number of experiments), m_i obeys the following recursion equations:

$$m_i = (1 - S)\, m_p + S\,(1 - \rho)\, {m'}_{i-1}, \tag{S2}$$

where ${m'}_{i-1} = \frac{1}{\tau}\sum_{j=i-\tau}^{i-1} m_j$, and $S = \langle S_i \rangle$ is the mean sensitivity to social information. The term $(1 - \rho)\, {m'}_{i-1}$ is the mean social information provided by a fraction $(1 - \rho)$ of individuals and a fraction ρ of experts giving the estimate $V = 0$. For $i \to +\infty$, m_i and ${m'}_{i-1}$ converge to the fixed point solution m_∞ of equation S2

$$m_\infty = m_p\, \frac{1 - S}{1 - S(1 - \rho)}, \tag{S3}$$

which is independent of τ. This analytical prediction of the simple model is plotted in Figure 3A, and is in fair agreement with the experimental data. $m_p < 0$ (reflecting the human bias to underestimate large quantities) is inferred from the center/median of the experimental distribution of prior estimates X_p (Table S1).

Similarly, the width of the distribution of the i-th individual estimates (excluding the virtual experts) satisfies the recursion relation (before averaging over configuration)

$$\sigma_i^a = (1 - S_i)^a\, \sigma_p^a + S_i^a\,(1 - \rho){\sigma'}_i^a, \tag{S4}$$

where ${\sigma'}_{i-1}^a = \frac{1}{\tau^a}\sum_{j=i-\tau}^{i-1} \sigma_j^a$. If personal estimates are Cauchy distributed ($a = 1$), then σ_i follows the same dynamics as m_i, and the asymptotic width is

$$\sigma_\infty = \sigma_p\, \frac{1 - S}{1 - S(1 - \rho)}. \tag{S5}$$

This analytical prediction of the simple model is plotted in Figure 3B, showing a fair agreement with the experimental data. The width σ_p is also inferred from the experimental distribution of prior estimates X_p (Table S1).

Remarks:

- Note that this simple analytical model (along with preliminary data from the first experiment in Japan) was exploited to design the experiment in France, in order to characterize the parameters for which significative results could be observed experimentally: number of subjects, percentage of virtual experts, number of sequential iterations and questions, values of τ... We are confident that the present study demonstrates the fruitful impact of developing *in parallel* experiments and models in social, behavioral, and cognitive sciences.

- If the virtual agents provide the value $V \neq 0$ instead of the true answer $V = 0$ (experts) as in our experiments, the width of the estimate distribution after social influence remains unaffected, but the center of the estimate distribution becomes

$$m_\infty = \frac{m_p(1 - S) + \rho\, S\, V}{1 - S(1 - \rho)}. \tag{S6}$$

We note that m_∞ can be driven to the exact result ($m_\infty = 0$) if the fraction ρ of virtual agents provides an *incorrect* information taking the optimal value

$$V_0 = -\frac{m_p(1 - S)}{\rho\, S}. \tag{S7}$$

As explained in the main text, and since $m_p < 0$ due to the human bias to underestimate large quantities, a strictly positive $V = V_0$ can lead to a perfect collective performance. This optimal V_0 becomes naturally larger as the density of virtual agents decreases (see Figure S16 for the confirmation of this effect in the full model).

- For $\tau = 1$, the full dynamics can be computed analytically as equation S2 reduces to $m_i = (1-S)\, m_p + S\,(1-\rho)\, m_{i-1}$. Defining $z = (1 - \rho)\, S < 1$, one obtains

$$m_i = z^i\,(m_0 - m_\infty) + m_\infty, \tag{S8}$$

where m_0 is the initial condition. We hence find that m_i converges exponentially to its asymptotic value m_∞, a result which can be shown to remain true for any τ (m_i is then the sum of m_∞ and of τ exponentially decreasing terms).

- It is natural to assume that in our sequentially designed experiment, if enough subjects are considered, a convergence state would be reached where the distribution of estimates would be independent of the iteration step, as found in our models. Stable distributions satisfy this expectation: by definition of stability (see main text), the social information M, as the mean of previous estimates, also has the same distribution (up to its width), and the linear combination of equation S1 and 2 (main text) would self-consistently lead to a new estimate with again the same distribution.

DRAFT

Impact of τ. Our model predicts an effect of τ that is too small to be observed experimentally (much less than the experimental error bars in Figure 3). In our main model (like in the simpler solvable model developed above), the dynamics of estimates converges, although we cannot compute the convergence value analytically. The convergence process is enhanced, especially at $\rho = 0\%$, by the initial conditions, since we set off the initial estimate very far from the truth. The convergence value does not depend on τ, but the convergence speed (and hence the dynamics) does. It follows that the median of estimates over 20 successive subjects, and hence the collective performance in Figure 3A, depend on τ too. The effect of τ in Figure 3B can be understood intuitively: at $\rho = 0\%$, the larger τ, the more agents are likely to receive similar information, because of averaging effects. Therefore, subjects are also enticed to give more similar results, hence the smaller distribution width at $\tau = 3$ than at $\tau = 1$. This effect is the same when $\rho > 0\%$, but since experts are all providing the same information, the higher ρ, the more the experts' information takes over the agents' information, and hence the lesser the impact of τ.

Computation of the error bars. In this section, we explain our procedure to define upper and lower error bars in Figures 2A, 3 and 4 (main text), and in this appendix (for instance, in Figure S8A). In particular, these error bars should attempt to translate the variability of our results depending on the 29 questions presented to the subjects. In order to do so, we have defined a kind of bootstrap procedure for each point plotted in these figures.

We call x_0 the actual measurement of a quantity appearing in these figures by considering all 29 questions. Then, we generate the results of $N = 10000$ new effective experiments. For each effective experiment indexed by $n = 1, ..., N$, we randomly draw $Q = 29$ questions among the 29 questions actually 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 $C = 70\,\%$ 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 $C = 70\,\%$ 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 well adapted to the present case where the distribution of the x_n is not symmetric, which is expected when the measured quantity x_0 is intrinsically positive and potentially close to 0.

Note that when x_0 is the *average* of a quantity (like in Figures 2A and S8A), and if the law of large number applies (which should be marginally the case for $Q = 29$), the x_n would be symmetrically and Gaussian distributed around x_0. Our error bars would then be symmetric ($b_+ = b_-$) and would be only slightly larger than a standard error, which corresponds to $C \approx 68.3\,\%$ for a Gaussian distributed random variable, instead of $C = 70\,\%$. We have checked that for Figures 2A and S8A, our procedure indeed almost leads to identical error bars as the standard error $b = \lim_{N\to+\infty} \sum_{n=1}^{N} (x_0 - x_n)^2/N = \sigma/\sqrt{Q}$. σ is the usual standard deviation from question to question: $\sigma^2 = \sum_{q=1}^{Q} (x_0 - X_q)^2/Q$, where X_q is the quantity of interest measured for each question q.

Our procedure is particularly well suited for computing error bars for Figures 3 and 4, for which x_0 is a median or a semi-interquartile range, considering that there exists no robust definition of a standard error for such quantities when the x_n are not Gaussian distributed. In particular, the fact that the measured quantities are intrinsically positive generally produces smaller lower error bars than upper error bars ($b_+ > b_-$) in these figures.

DRAFT

	France	Japan	Japan
m_{p}	-0.28	-0.11	-0.3
σ_{p}	0.56	0.23	0.62
m	-0.18	-0.08	
σ	0.37	0.13	

Table S1. Experimental distributions of estimates. The three columns report respectively the median and width of the distribution of estimates (before – with index p – and after social influence) for the experiment performed in France, for the first experiment performed in Japan ("easier" questions, hence smaller center/median shift and width), and for the additional experiment in Japan, in which subjects were asked the same questions as in France (but no second estimate after social influence were collected). See also Figure S2.

	France	Japan
$\langle S \rangle$	0.45	0.7
Std(S)	3.1	3.68
Median(S)	0.34	0.59
$\langle -2 < S < 3 \rangle$	0.41	0.57
$\mathrm{Std}(-2 < \mathrm{S} < 3)$	0.34	0.58
$\mathrm{Median}(-2 < \mathrm{S} < 3)$	0.43	0.55
$\langle 0 \leq S \leq 1 \rangle$	0.37	0.46
$\mathrm{Median}(0 \leq \mathrm{S} \leq 1)$	0.31	0.48
$\langle 0 < S < 1 \rangle$	0.5	0.59
$\mathrm{Median}(0 < \mathrm{S} < 1)$	0.48	0.61

Table S2. We present statistical quantities characterizing the experimental distribution of sensitivities to social influence as obtained in France (blue) and Japan (red). $\langle x \rangle$ refers to the mean of the variable x (sometimes in the specified interval). One single data in France and Japan for which $S > 100$ has been removed from the analysis).

	France	Japan
m_{p}	-0.33	-0.11
σ_{p}	0.53	0.3
m_{g}	0.58	0.72
σ_{g}	0.3	0.38
P_1	0.04	0.07
α	0.36	0.47
β	0.07	0.3
$S_{\max}$	0.53	0.74

Table S3. Model parameters defined in the description of the full model, in the main text. The center m_{p} and width σ_{p} of the prior estimate distribution are inferred from Figure 1A and S2. m_{g} and σ_{g} are the mean and variance of the Gaussian part of the model distribution of sensitivities to social influence, and are inferred from Fig. 1B, by imposing that the model and experimental distributions of S lead to the same fraction of subjects in the five behavioral categories. The last four parameters are measured from Figure 2 and S8 (see the description of the model in the main text).

DRAFT

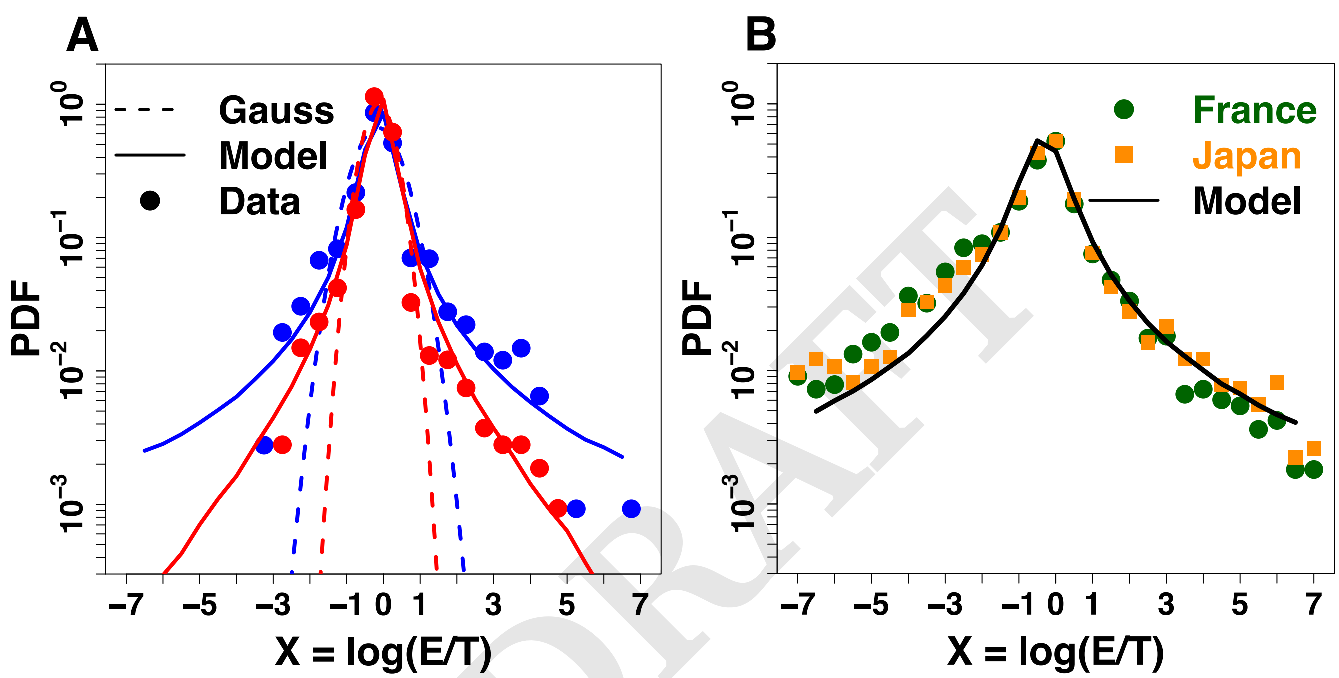

Fig. S2. A: PDF of log-transformed estimates, before (blue) and after (red) social influence, in the experiment performed in Japan (2520 estimates), for all questions mixed. The dashed and full lines are respectively a Gaussian fit and the prediction of the model. The distribution is more peaked than the one obtained in France, because the questions asked were generally "easier", in the sense that people had more cultural knowledge about them (see main text). B: Comparative distribution in Japan (4025 estimates) and France (5394 estimates), before social influence, with this time the same questions asked. Both distributions are very similar, supporting our claim that the difference in distributions is mainly due to the difference in questions difficulty. The black line is the distribution of estimates simulated by our model. One can notice that the cognitive bias to underestimate low values is common to both countries.

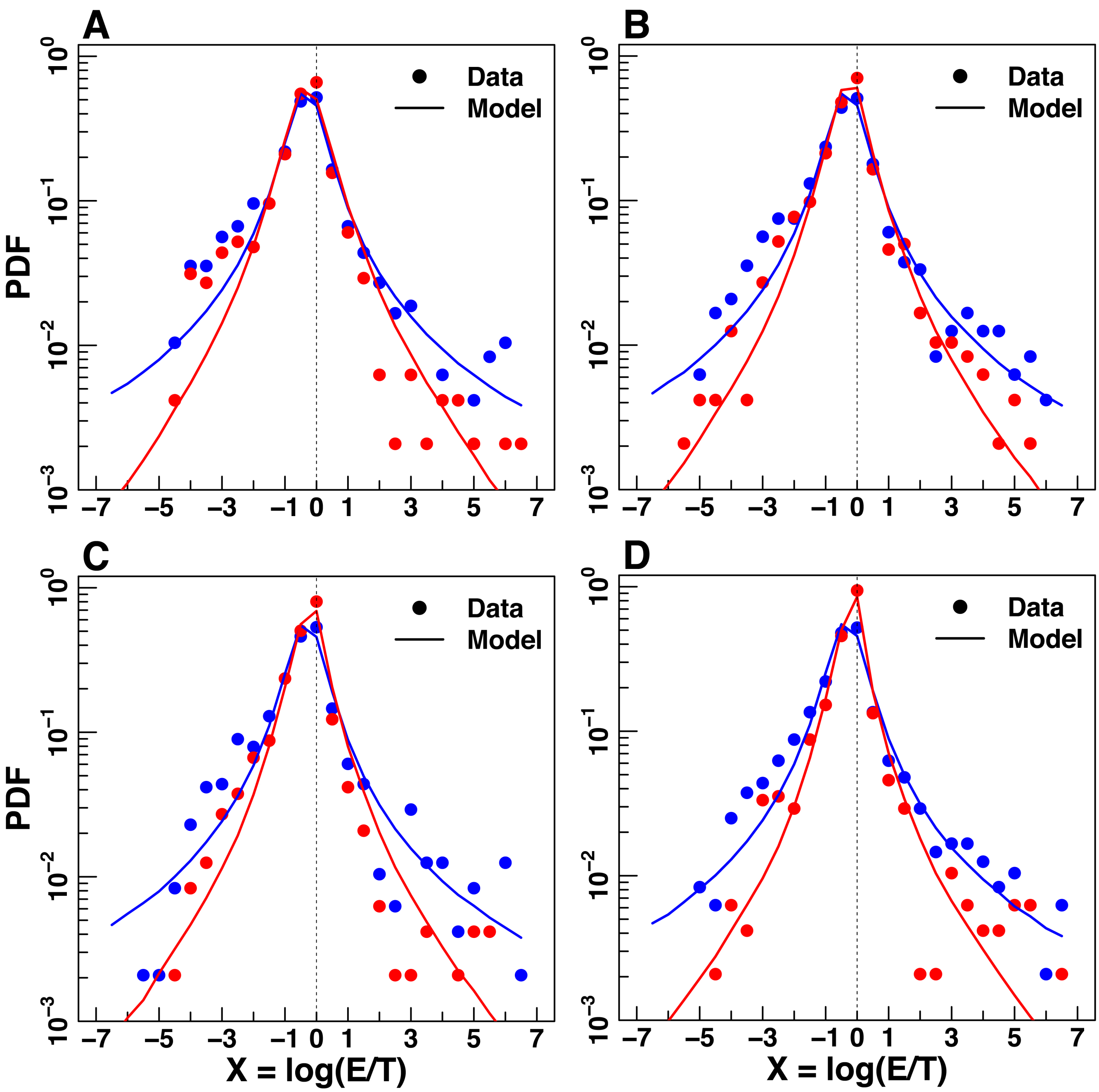

Fig. S3. Probability distribution function (PDF) of log-transformed normalized estimates, before (blue) and after (red) social influence, for all values of ρ: A. $\rho = 0\%$; B. $\rho = 20\%$; C. $\rho = 43\%$; D. $\rho = 80\%$. In our model, we start from a Cauchy distribution of the initial estimates, and restrict the estimates to $[-7, 7]$. Panel A shows that this restriction is enough to reproduce the experimental distribution of personal estimates (in blue), and to explain the sharpening after social influence (in red), when there are no experts. B, C and D (and Fig. 2 in the main text) show that this sharpening increases as more information is provided to the group (*i.e.* as ρ increases).

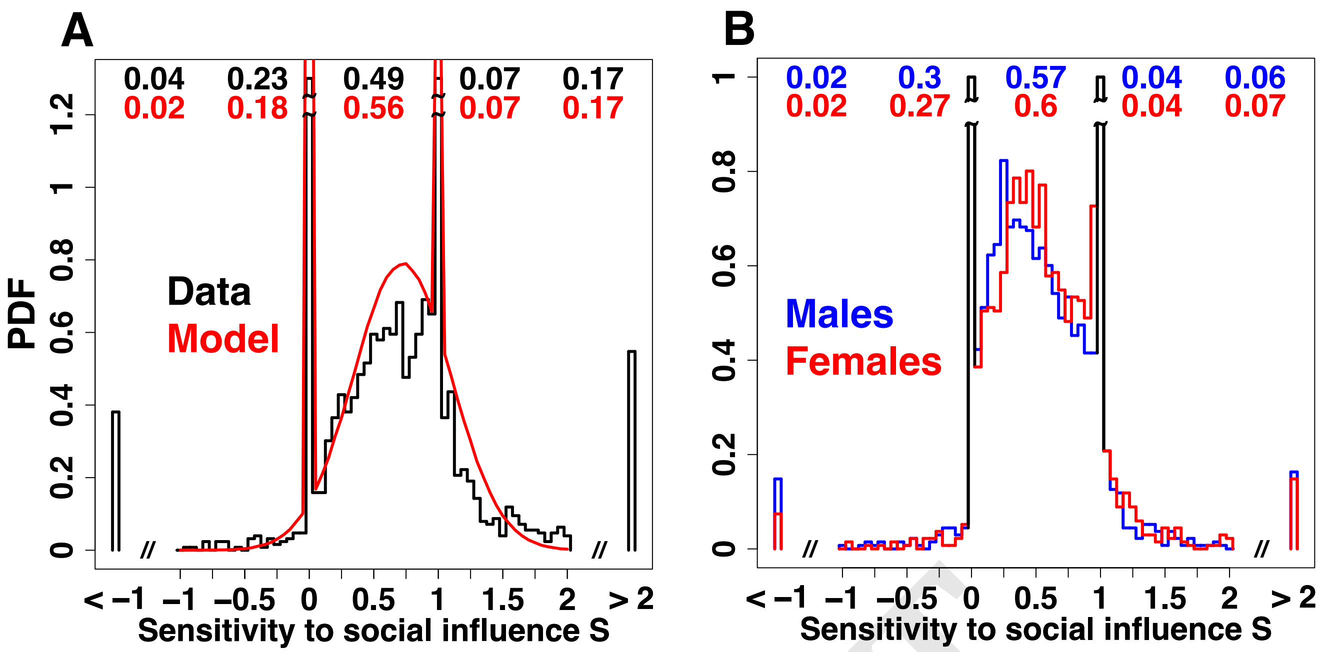

Fig. S4. A. PDF of the sensitivity to social influence S in the experiment performed in Japan. A smaller percentage of participants keep their opinion ($S = 0$) or compromise ($0 < S < 1$), compared with the experiment performed in France (Figure 1B in main text), while the percentage of subjects adopting others' opinion ($S = 1$) or overreacting to it ($S > 1$) is larger. B. Same PDF for males (blue) and females (red) in the experiment performed in France. The patterns are very similar, with a slight tendency for females to compromise more.

DRAFT

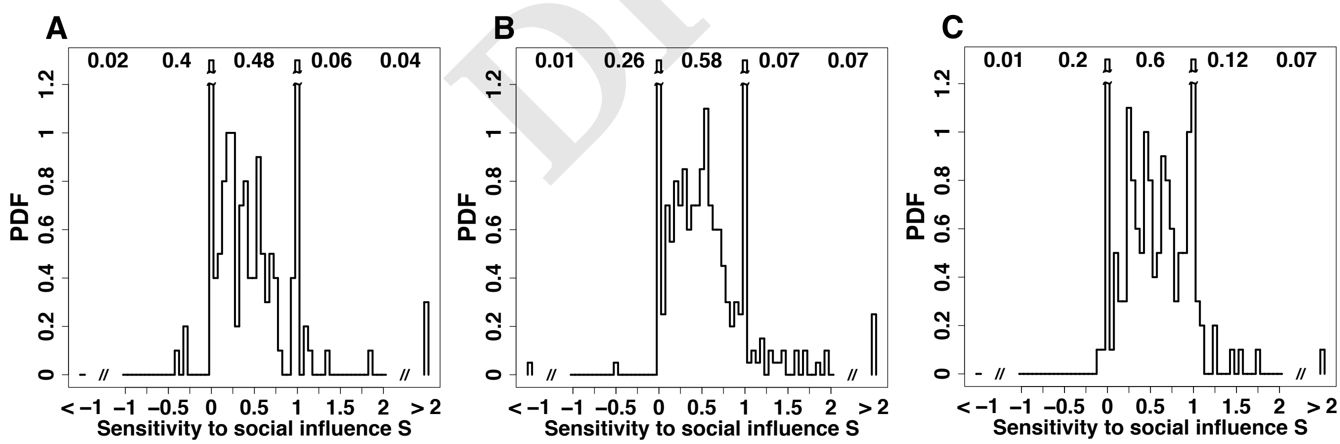

Fig. S5. PDF of the sensitivity to social influence S, for questions 25 to 29, in the following cases: A. *confident* subjects, defined as the quarter of each subgroup of 8 subjects minimizing $\langle |S_q| \rangle_q$, where q is the index for questions 1 to 24 (the subjects who were on average closest to $S = 0$); B. "*average*" subjects, defined as being neither confident nor follower; C. *followers*, defined as the quarter of the group minimizing $\langle |1 - S_q| \rangle_q$ (the ones who were on average closest to $S = 1$). In A, the peak at $S = 0$ is almost 7 times higher than the one at $S = 1$, whereas in B, it is less than 4 times, and in C, less than twice. Therefore, subjects who were characterized as confident from questions 1 to 24 remained highly confident, whereas subjects who were characterized as followers remained highly followers. For the sake of consistency, subjects who were characterized as "average" remained "average", in the sense that their distribution is very close to the global distribution (Figure 1B).

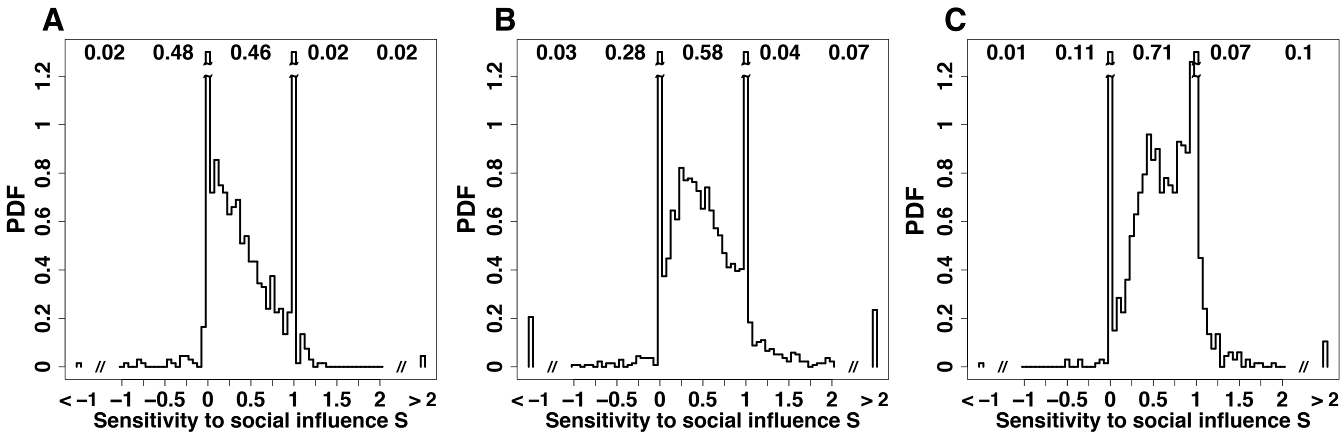

Fig. S6. PDF of the sensitivity to social influence S, for individuals identified as confident (A), "average" (B) or followers (C). We reclassified individuals in these 3 categories, according to the same definition as given in Figure S5, but this time for questions 1 to 29. The differences are even more apparent than in Figure S5.

DRAFT

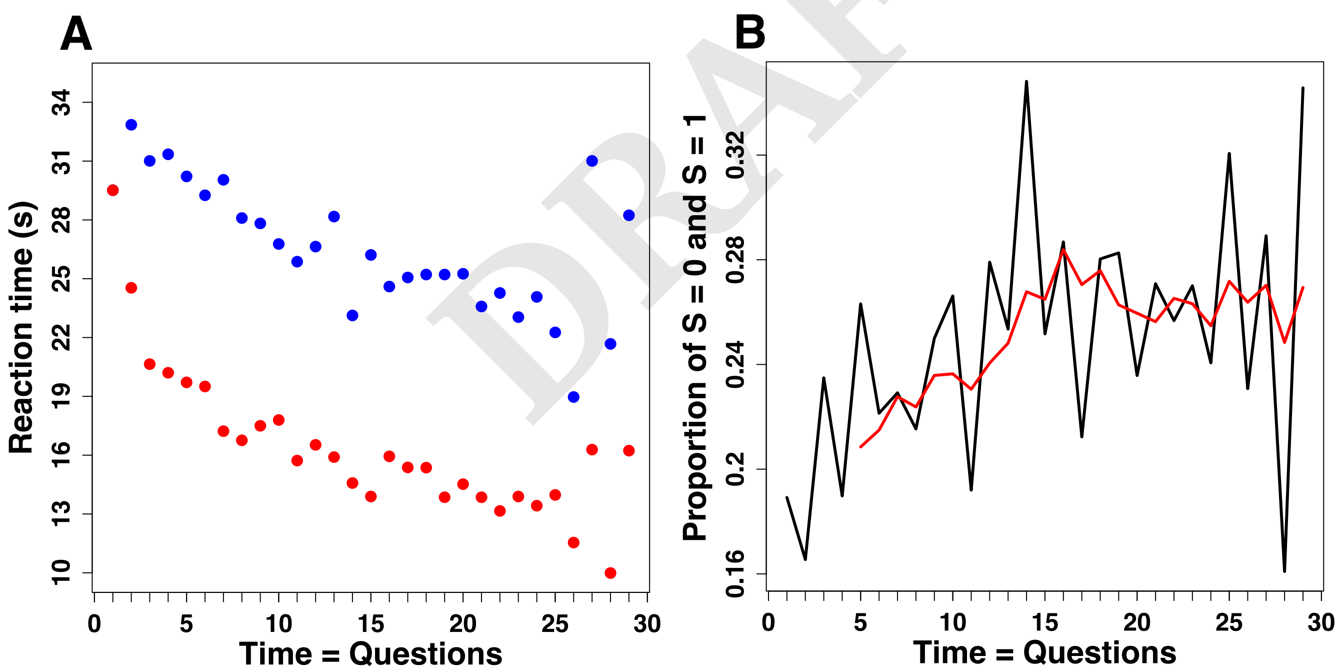

Fig. S7. A. Average reaction time (in seconds) for all subjects, before (blue) and after (red) social influence, against "time" (order of occurrence of a question during the experimental session). As the experiment proceeds, the subjects tend to take less and less time in providing their answers: they get more familiar with the procedure, but also, possibly, take less time for their answer out of increasing boredom. They also answer much faster after social influence, showing that subjects take much more time making a decision than revising it; B. Evolution of the proportion of subjects keeping their opinion ($S = 0$), or adopting that of the group ($S = 1$), with time. Data are in black, while the smoothed average over the 5 previous values is shown in red. As time passes, subjects tend to keep or adopt slightly more. Since keeping or adopting are mentally easier decisions than compromising, this suggests that subjects are getting gradually more tired/bored by the experiment.

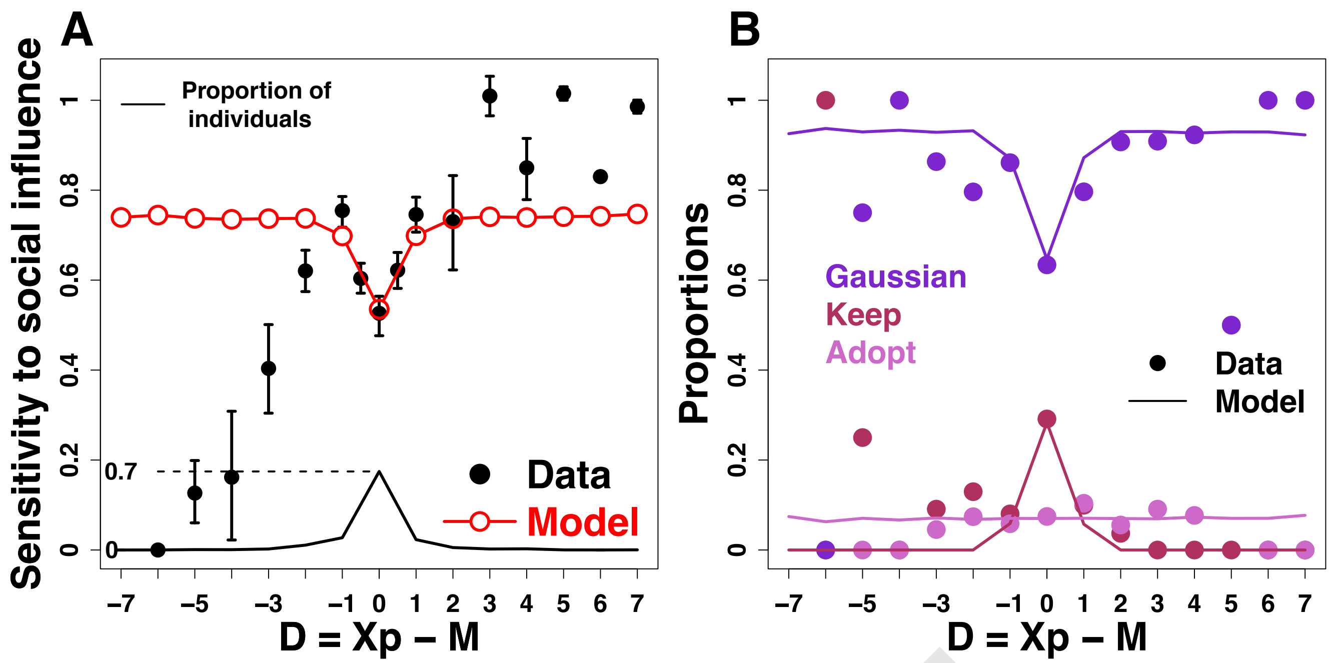

Fig. S8. A. Average sensitivity to social influence S against the distance between personal and group's estimate $D = X_p - M$, in the experiment performed in Japan. As in the experiment performed in France (Figure 2), there is a linear cusp relation for $|D| < \approx 1.5$ where most of the data lies: the farther the group's estimate from the personal estimate, the more subjects tend to trust the group (S gets closer to 1). The threshold $t \approx 1.5$ is lower than in the experiment performed in France ($t \approx 2.5$). This difference comes from differences in the questions asked to the participants (see also Figure S3). In Japan, the questions asked to the participants were of moderate difficulty compared to those asked in France, and much more answers were closer to each other. Note that the pattern appearing above the threshold is hardly significant, because they are only 3.6% of the data beyond 2 orders of magnitude $|D| > 2$. B. Proportion of subjects keeping their opinion (maroon), adopting the group's opinion (pink), or being in the Gaussian-like part of the distribution of S (compromising, contradicting, or overreacting; in purple), against D. Just as in France, we find that there is a probability transfer from the peak at $S = 0$ (keeping one's opinion), to the Gaussian-like part of the distribution of S. Therefore, the farther away the group's estimate from the personal estimate, the more the subjects are inclined to compromise.

DRAFT

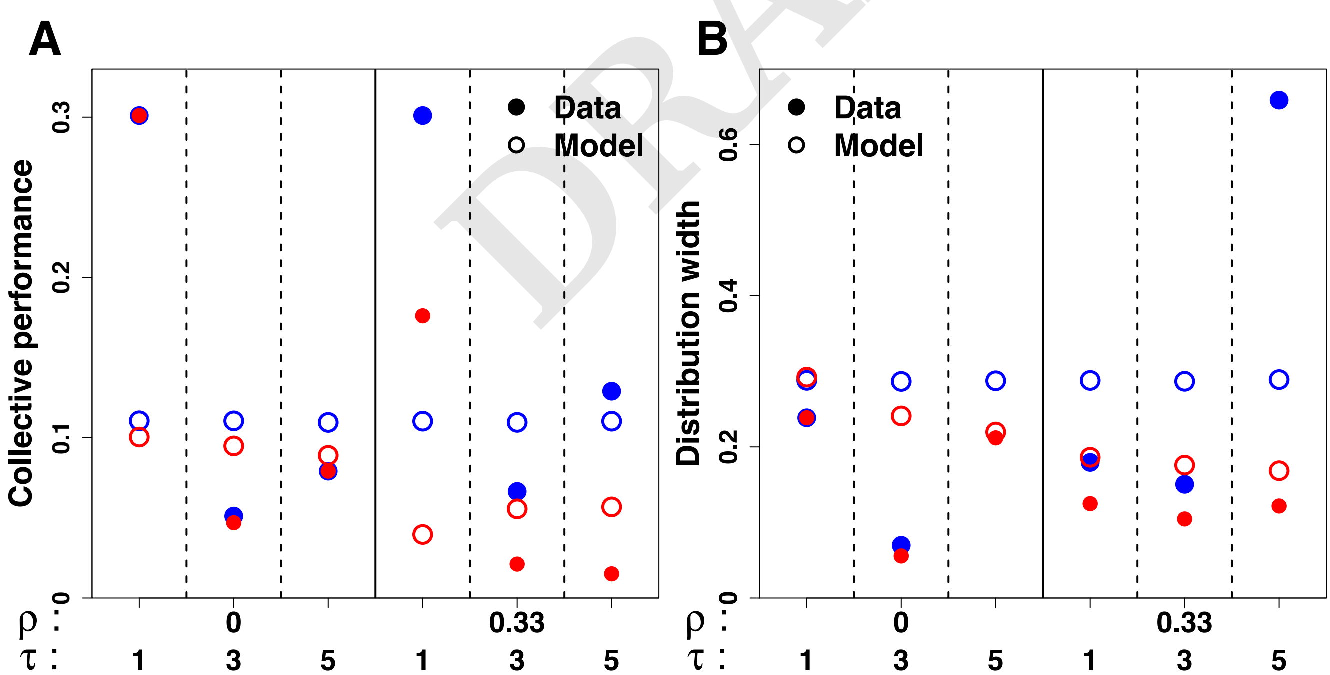

Fig. S9. A Collective performance (absolute value of the median of log-transformed estimates) in the experiment performed in Japan, before (blue) and after (red) social influence, for each couple of values (ρ, τ). At $\rho = 0\%$ (no virtual experts), the performance doesn't change after social influence, as expected (the center of the distribution doesn't move). However, at $\rho = 33\%$, there is an improvement in collective performance after social influence for all values of τ. It is not clear whether there is an effect of τ, because the questions asked were different in all conditions (ρ, τ); B. Width of the distribution of estimates in the experiment performed in Japan, before (blue) and after (red) social influence, for each couple of values (ρ, τ). The distribution width has on average decreased after social influence, even at $\rho = 0\%$. The decreasing is nonetheless clearer at $\rho = 33\%$. Thus, the patterns are similar to those observed in the experiment performed in France (Figure 3 in the main text).

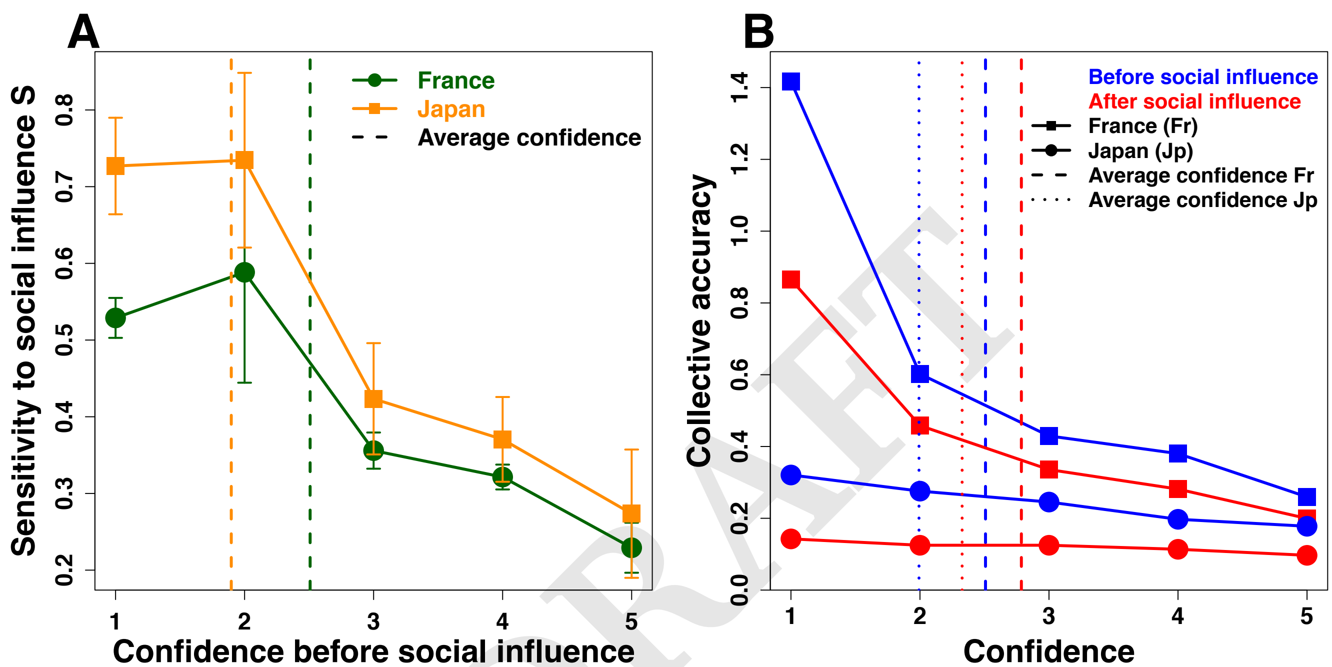

Fig. S10. A. Average sensitivity to social influence S against the confidence reported before social influence, in France (blue) and Japan (red). S is negatively correlated to confidence, meaning that the most confident subjects are also the less sensitive to social influence. The average sensitivity to social influence is higher in Japan than in France, and conversely, the average confidence is higher in France than in Japan. This is all the more surprising, because the questions were "harder" in France (see Figure S3). This suggests cultural differences in expression of confidence and attention to others' opinions. B. Collective accuracy (median of absolute values of log-transformed estimates), before (blue) and after social influence (red), for the experiments performed in France (squares) and Japan (circles). Collective accuracy is positively correlated with confidence, meaning that the most confident individuals are also the most accurate in their answers (answers closest to 0). Most likely, confident subjects perform well because they have some prior knowledge about the questions, which at the same time makes them less likely to follow the opinion of others. The average accuracy is better in Japan than in France, because of the difference in the questions' difficulty (see also Figure S3A). The average confidence and accuracy both improve after social influence.

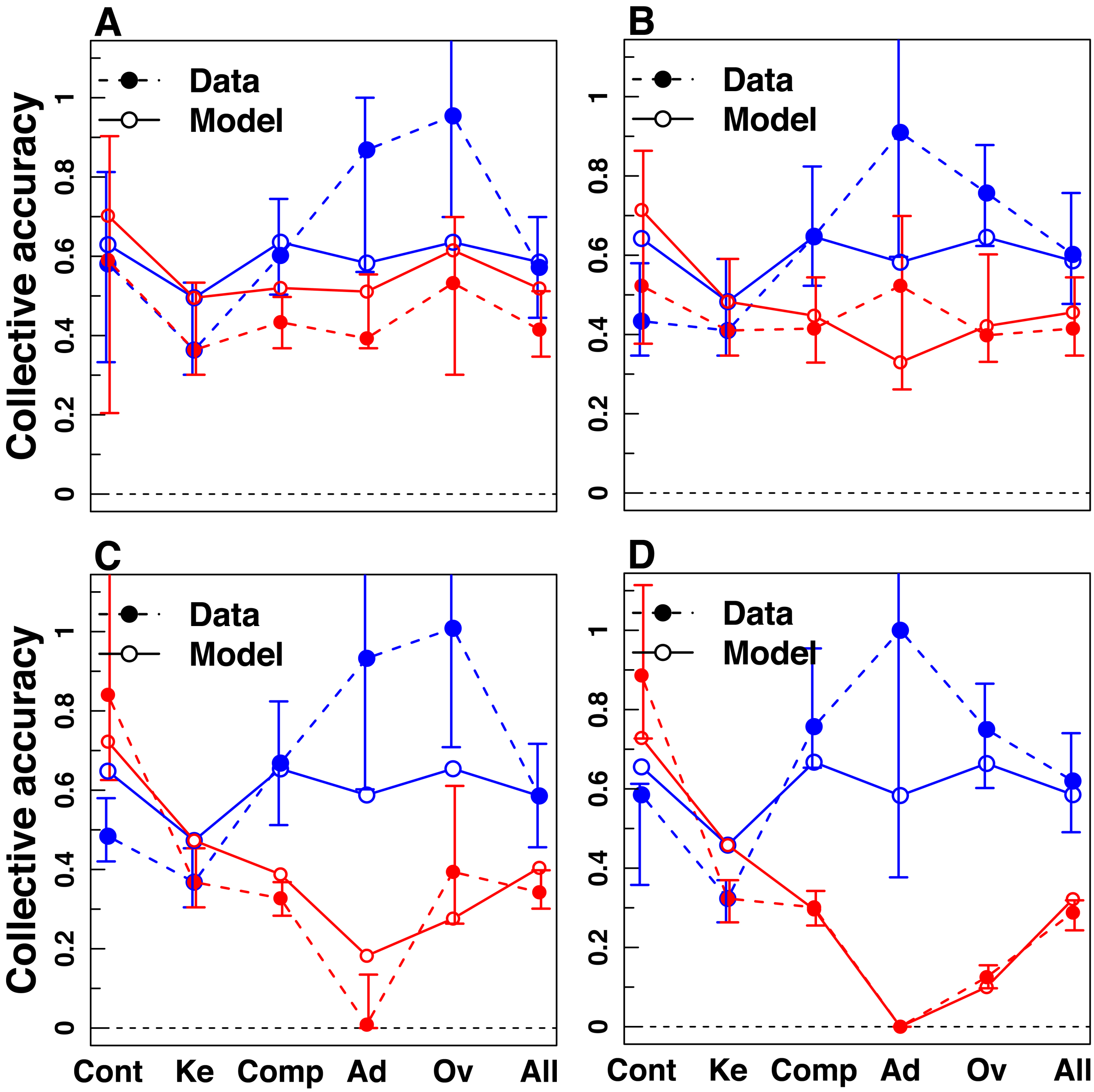

Fig. S11. Collective accuracy (median distance to the truth of individual estimates) before (blue) and after (red) social influence, for the 5 behavioral categories identified in Figure 1B and for the whole group (All): A. $\rho = 0\,\%$; B. $\rho = 20\,\%$; C. $\rho = 43\,\%$; D. $\rho = 80\,\%$. As ρ increases, the group improves its accuracy after social influence. Interestingly, adopting leads to the best improvement and accuracy after social influence for $\rho > 20\,\%$. Full circles correspond to experimental data, while empty circles represent the predictions of the model. This figure presents an alternative representation to that of Figure 4 in the main text

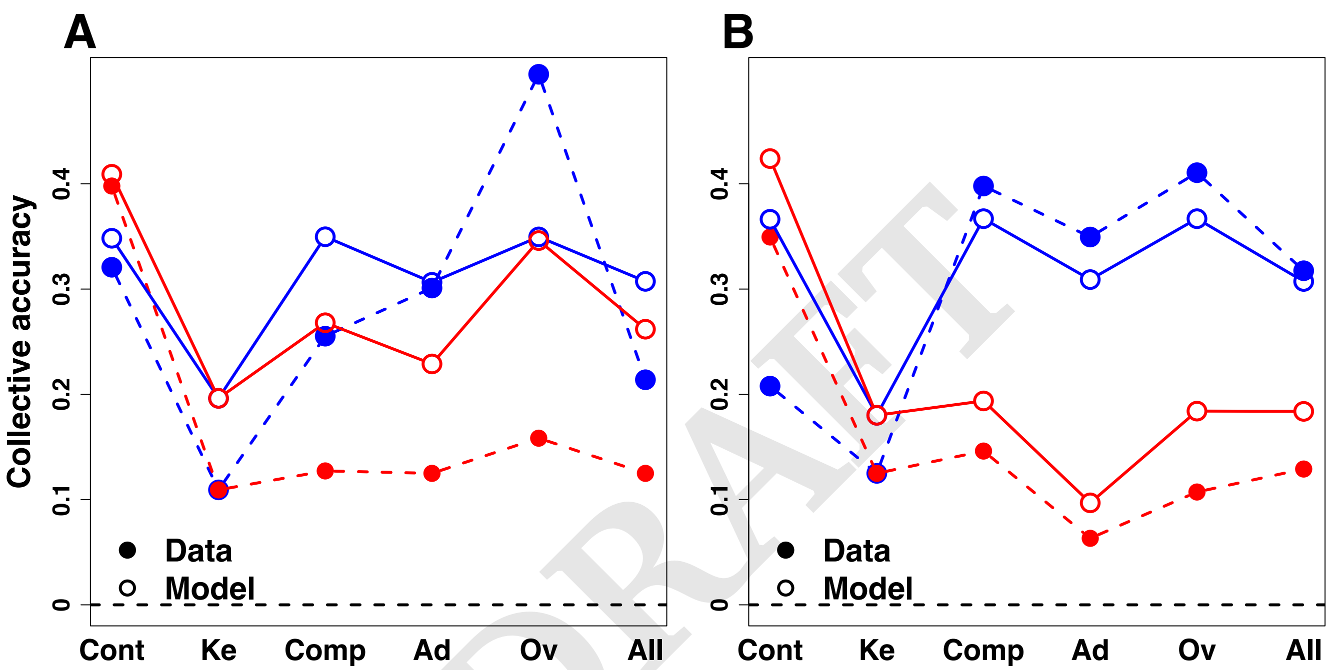

Fig. S12. Collective accuracy (median of absolute values of log-transformed estimates) for the experiment performed in Japan, before (blue) and after social influence (red), for the two values of ρ (percentage of experts): 0% (A) and 33% (B). As in France (Figure 4 of the main text), subjects who keep their opinion are the most accurate before social influence. In A, at $\rho = 0\%$, all behavioral categories lead approximately to the same accuracy after social influence, except contradicting, which leads to much worse accuracy. In B, at $\rho = 33\%$, adopting is the behavior leading to the best accuracy after social influence. The group's accuracy improves after social influence both without (A) and with (B) experts, but much more markedly in the presence of experts. However, it is difficult to compare A and B, because the questions asked were different. Again, the patterns are similar to those observed from the experiment performed in France (Figure 4 and S11).

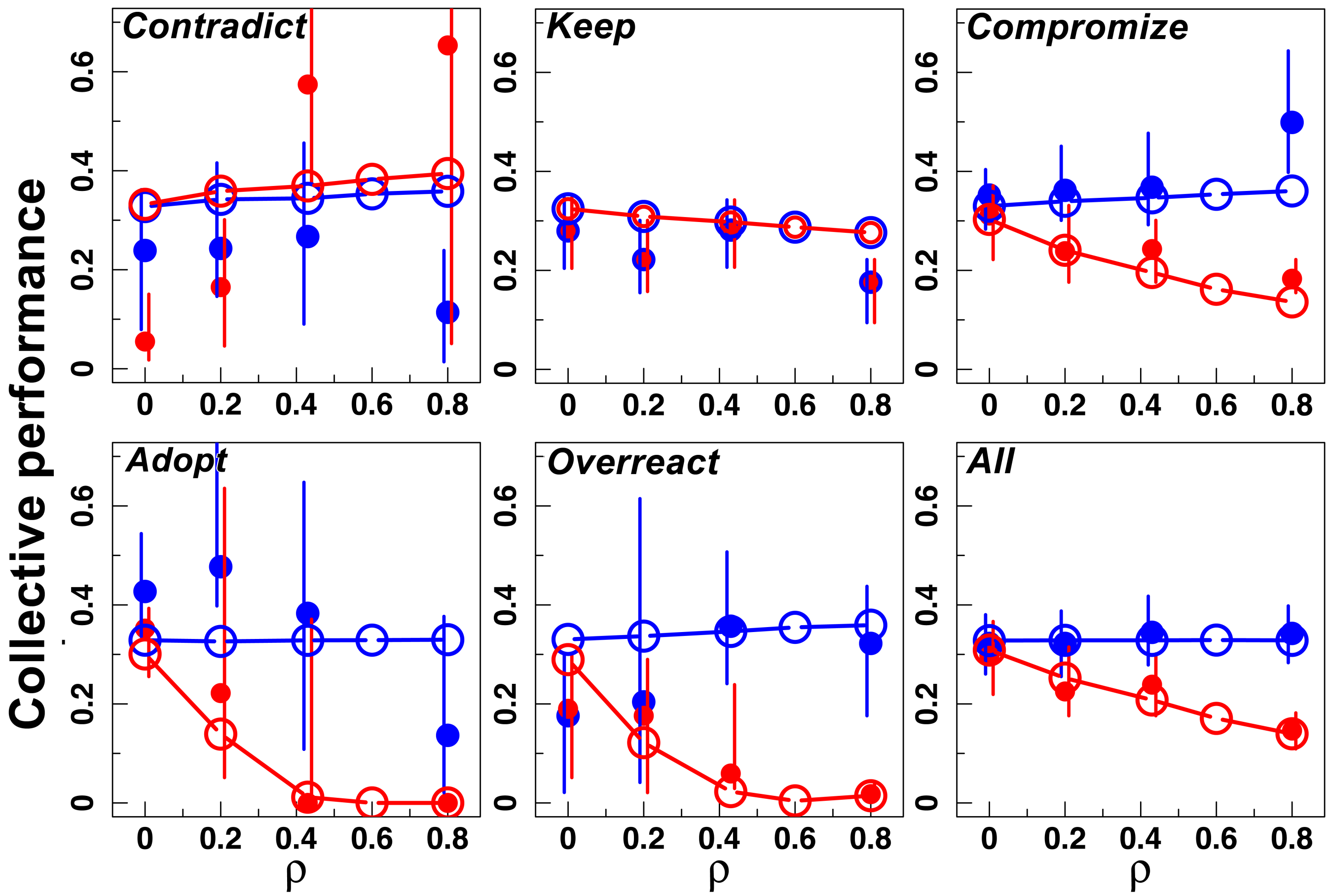

Fig. S13. Collective performance (absolute value of the median of log-transformed estimates) before (blue) and after (red) social influence against ρ, for the 5 behavioral categories identified in Figure 1B and for the whole group (All). Adopting leads to the sharpest improvement, and the best performance for $\rho \geq 40\,\%$. Full circles correspond to experimental data, while empty circles represent the predictions of the model (including for $\rho = 60\,\%$, a case not tested experimentally).

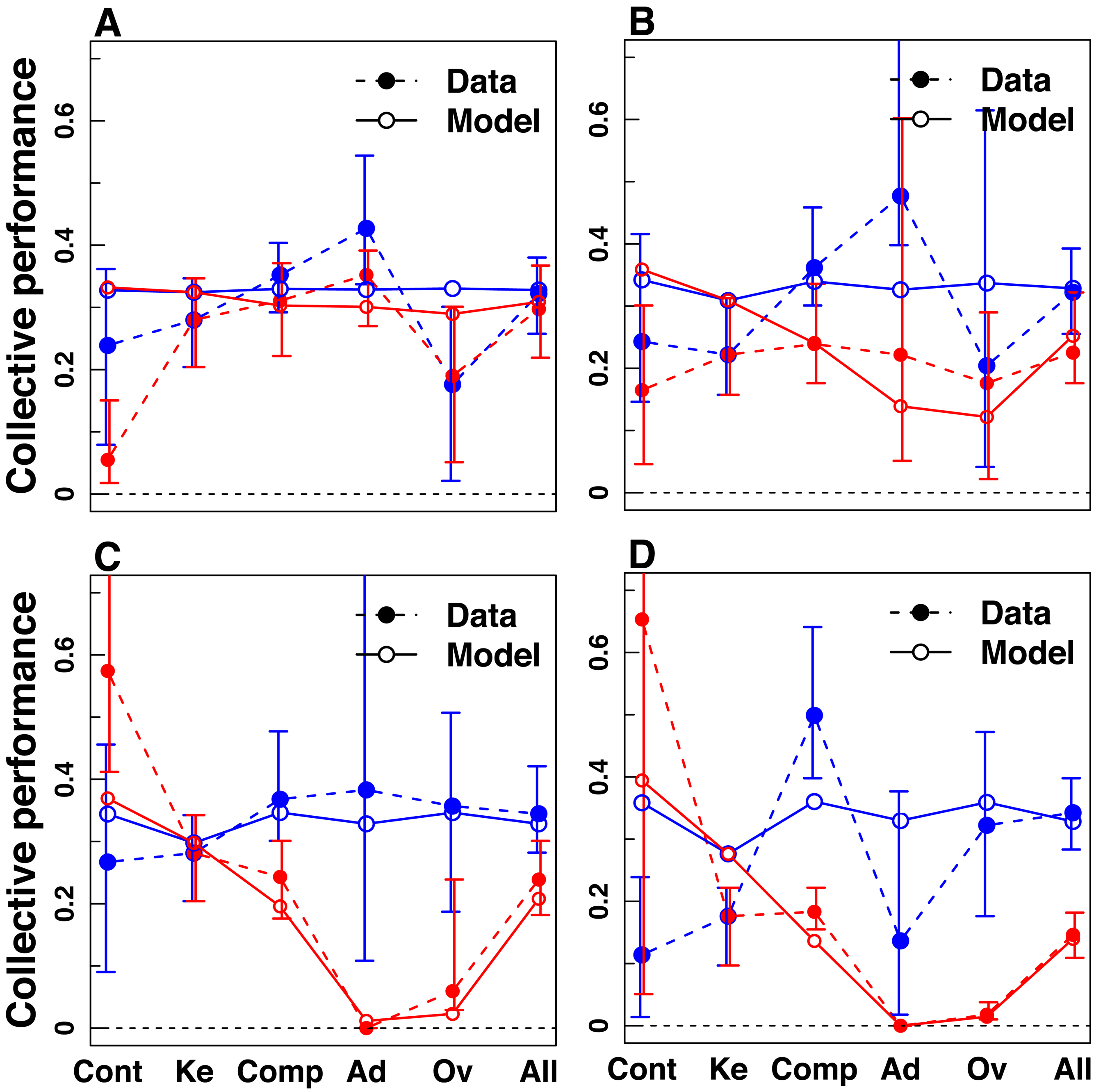

Fig. S14. Collective performance (absolute value of the median of log-transformed estimates), for the different values of ρ, before (blue) and after (red) social influence, in the experiment performed in France: A. $\rho = 0\%$; B. $\rho = 20\%$; C. $\rho = 43\%$; D. $\rho = 80\%$. The results for the five behavioral categories identified in Figure 1B, as well as the overall group (All), are presented. Experimental values correspond to full circles, and simulation values to empty circles. Similarly to the collective accuracy, the behavior leading to the best improvement in performance consists in adopting others' opinion, once virtual experts are introduced. Again, the improvement in collective performance after social influence increases with ρ.

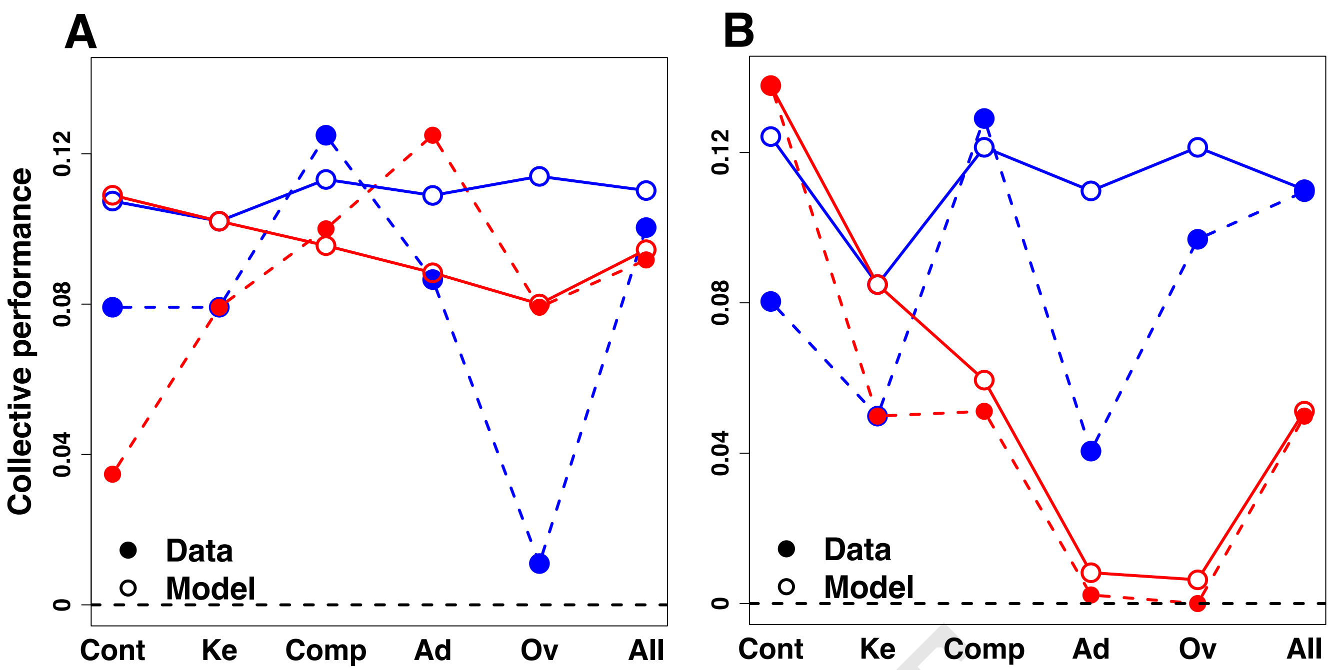

Fig. S15. Collective performance (absolute value of the median of log-transformed estimates) for the experiment performed in Japan, before (blue) and after social influence (red), for the two values of ρ (percentage of experts): 0% (A) and 33% (B). Adopting and overreacting also lead to the best improvement in collective performance after social influence. The improvement seems to increase with ρ, although we cannot tell if ρ is the only factor, because the questions asked in A and B were not the same.

DRAFT

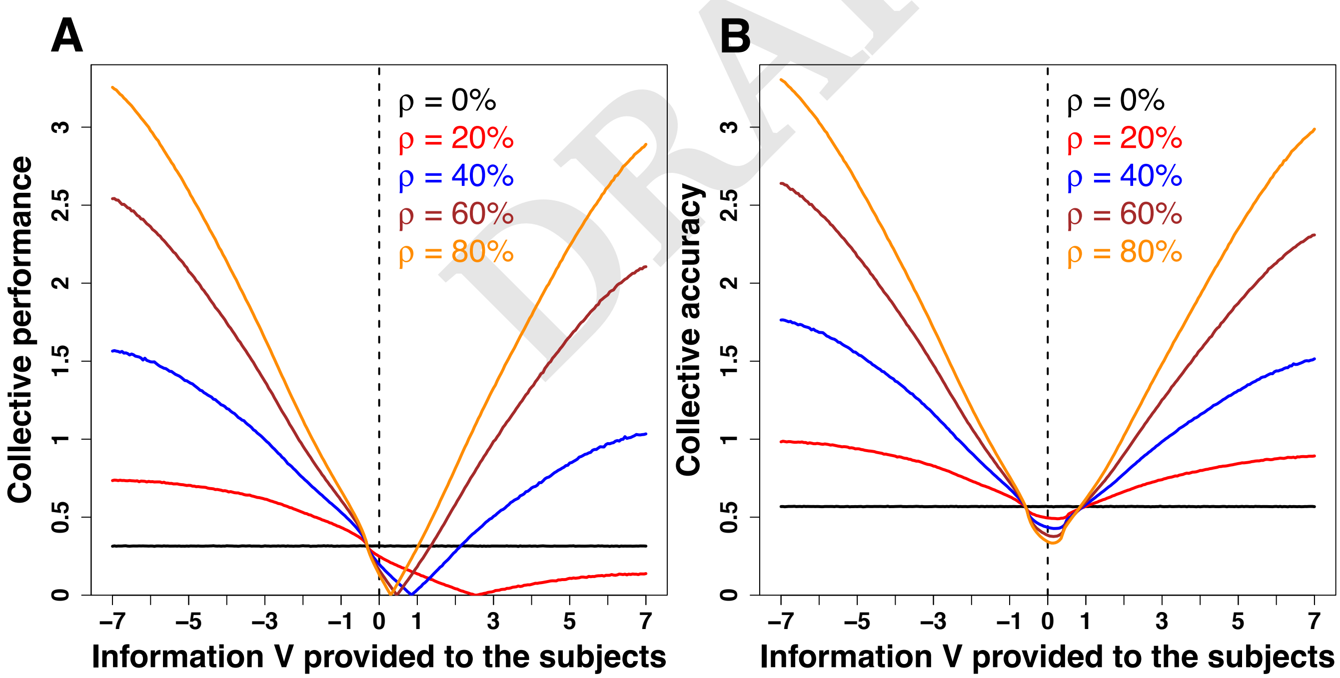

Fig. S16. A. Collective performance against the value of the information V (log-transformed value of the answer provided by the virtual experts), for different values of ρ, as obtained using the full numerical model. For all ρ, we find that the collective performance reaches a minimum at a strictly positive $V = V_0$ (decreasing as ρ increases), as predicted by the simple analytical model presented above. Hence, providing incorrect information greater than the true value can be more beneficial (the collective performance closer to 0) to the group performance than providing the truth, by compensating for the human natural cognitive bias to underestimate quantities ($m_p < 0$). For $\rho = 20\%$, 40%, 60% and 80%, the optimum collective performance is respectively reached for $V_0 = 2.5$, $V_0 = 0.8$, $V_0 = 0.4$ and $V_0 = 0.25$; B. Collective accuracy, against the value of the additional information V, for different values of ρ, as predicted by the full model. The pattern is very similar as in A, the optimum being reached for $V > 0$, for any ρ. Note that the optimum collective accuracy improves as ρ increases.